Structure-based Development of Secondary Amines as Aspartic Protease Inhibitors

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Pharmazie der Philipps-Universität Marburg

vorgelegt von

Jark Böttcher

aus Winsen (Luhe)

Marburg, 2008

Vom Fachbereich Pharmazie der Philipps-Universität Marburg
als Dissertation angenommen am: 21.11.2008

Erstgutachter: Prof. Dr. Gerhard Klebe
Zweitgutachter: Hochschuldozentin Dr. Wibke E. Diederich

Tag der mündlichen Prüfung am: 21.11.2008

Die Untersuchungen zur vorliegenden Arbeit wurden auf Anregung von Prof. Dr. Gerhard Klebe am Institut für Pharmazeutische Chemie des Fachbereichs Pharmazie der Philipps-Universität Marburg in der Zeit von November 2004 bis November 2008 durchgeführt.

The following scientific contributions are associated with this thesis:

1. Journal Articles

- **Jark Böttcher**, Andreas Blum, Stefanie Dörr, Andreas Heine, Wibke E. Diederich, and Gerhard Klebe. Targeting the Open Flap Conformation of HIV-1 Protease with Pyrrolidine-based Inhibitors. *ChemMedChem* **2008**, 3, (9), 1337-44.
- **Jark Böttcher**, Andreas Blum, Andreas Heine, Wibke E. Diederich, and Gerhard Klebe. Structural and Kinetic Analysis of Pyrrolidine-based Inhibitors of the Drug Resistant Ile84Val Mutant of HIV-1 Protease. *J. Mol. Biol.* **2008**, (383),2, 347-57
- Andreas Blum, **Jark Böttcher**, Benedikt Sammet, Torsten Luksch, Andreas Heine, Gerhard Klebe, and Wibke E. Diederich. Achiral Oligoamines as Versatile Tool for the Development of Aspartic Protease Inhibitors. *Bioorg. Med. Chem.* **2008**, 16, (18), 8574-86
- Andreas Blum, **Jark Böttcher**, Andreas Heine, Gerhard Klebe, and Wibke E. Diederich. Structure-Guided Design of C_2-Symmetric HIV-1 Protease Inhibitors Based on a Pyrrolidine Scaffold. *J. Med. Chem.* **2008**, 51, (7), 2078-2087.
- Andreas Blum, **Jark Böttcher**, Stefanie Dörr, Andreas Heine, Gerhard Klebe, and Wibke E. Diederich. Two solutions for the same problem:multiple binding modes of Pyrrolidine-based HIV-1 Protease Inhibitors. **in preparation**
- **Jark Böttcher**, Andreas Blum, Benedikt Sammet, Sascha Brass, Andreas Heine, Wibke E. Diederich, and Gerhard Klebe. Chasing Binding Modes in HIV Protease: From Seemingly Perturbed to Seemingly Relaxed Pose Without Altering Affinity. **in preparation**

2. Oral presentations

- **Jark Böttcher**, Andreas Blum, Wibke E. Diederich, and Gerhard Klebe. Structure-Based Design of Pyrrolidines as HIV Protease Inhibitors. *Frontiers in Medicinal Chemistry, Joint German-Swiss Meeting on Medicinal Chemistry,* Berlin **2007**
- **Jark Böttcher**, Oliver Rau, Yvonne Syha, Manfred Schubert-Zsilavecz, and Gerhard Klebe. Optimization of α-Substituted Pirinixic Acid Derivatives as Dual PPAR α/γ-Agonists. *Joint Meeting of the Czech, German and Hungarian Pharmaceutical Societies*, Marburg **2006**

3. Posters

- Andreas Heine, **Jark Böttcher**, Tina Ritschel, Andreas Blum, Benedikt Sammet, Simone R. Hoertner, Philipp Kohler, François Diederich, Wibke E. Diederich, and Gerhard Klebe. Structure-based drug design in HIV-1 protease- and tRNA-guanine transglycosylase inhibitor development. *XXI Congress and General Assembly of the international Union of Crystallography*, Osaka **2008**
- **Jark Böttcher**, Andreas Blum, Andreas Heine, Wibke E. Diederich, and Gerhard Klebe. Structural insights into Resistance: Pyrrolidine-based Inhibitors bound to crucial HIV-1 Protease Mutants. *Frontiers in Medicinal Chemistry, Joint German-Swiss Meeting on Medicinal Chemistry*, Berlin **2007**
- Andreas Blum, **Jark Böttcher**, Gerhard Klebe, and Wibke E. Diederich. Structure-Based Design of Pyrrolidines as HIV-Protease Inhibitors. Frontiers in Medicinal Chemistry, *Joint German-Swiss Meeting on Medicinal Chemistry,* Berlin **2007**
- Andreas Blum, **Jark Böttcher**, Gerhard Klebe, and Wibke E. Diederich. Development of 3,4-Disubstitueted Pyrrolidines as HIV-Protease Inhibitors. *Tetrahedron Symposium 2007: Challenges in Organic Chemistry*, Berlin **2007**
- **Jark Böttcher**, Andreas Blum, Edgar Specker, Sascha Brass, Andreas Heine, and Gerhard Klebe. Tracing a new binding competent conformation of the HIV-1 protease. *Frontiers in Medicinal Chemistry***,** *Annual Meeting***,** Frankfurt **2006**

- Wibke E. Diederich, Christof Gerlach, Andreas Blum, **Jark Böttcher**, Sascha Brass, Torsten Luksch, and Gerhard Klebe. Design and Synthesis of tailor-made compound libraries via a knowledge-based approach – A case study. *Abstracts of Papers, 232nd ACS National Meetin*g**,** San Francisco **2006**
- Nan-Si Chan, Sascha Brass, **Jark Böttcher**, Torsten Luksch, Gerhard Klebe, and Wibke E. Diederich. Aspartic Protease Inhibitors based on a 2,3,4,7,-Tetrahydro-1H-azepine scaffold. *Joint Meeting of the Czech, German and Hungarian Pharmaceutical Societies*, Marburg **2006**
- Sascha Brass, Nan-Si Chan, Torsten Luksch, **Jark Böttcher**, Gerhard Klebe, and Wibke E. Diederich**.** Synthesis of substitueted Tetrahydro-1H-azepines as non-peptidic inhibitors of aspartic proteases. *Doktoranden Tagung der DPhG*, Nürnberg **2006**
- Andreas Blum, **Jark Böttcher**, Gerhard Klebe, and Wibke E. Diederich. Functionalized Pyrrolidines – A New Class of HIV-Protease Inhibitors. *1st European Chemistry Congress*, Budapest **2006**
- Andreas Blum, **Jark Böttcher**, Gerhard Klebe, and Wibke E. Diederich. HIV-Protease Inhibitors based on a 3,4-Disubstitueted Pyrrolidines. *Frontiers in Medicinal Chemistry*, *Annual Meeting,* Frankfurt **2006**
- Andreas Blum, **Jark Böttcher**, Gerhard Klebe, and Wibke E. Diederich. Development of a 3,4-Disubstitueted Pyrrolidines as potent non-peptidic Inhibitors of HIV-Protease. *DPhG-Jahrestagung*, Mainz **2005**

Abbreviations

AIDS	acquired immunodeficiency syndrome
BACE-1	β-secretase
BisTris	1,3-*bis(tris*(hydroxymethyl)methylamino)propane,
BOC	di-*tert*-butyl dicarbonate
BSA	bovine serum albumin
DMPU	1,3-Dimethyl-3,4,5,6-tetrahydro-2(1H)-pyrimidinone
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EC number	enzyme commission number
EDC	*N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimid-hydrochlorid
EDTA	ethylenediaminetetraacetic acid
FDA	U.S. Food and Drug Administration
HAART	highly active anti-retroviral therapy
HIV	human immunodeficiency virus
HTS	high-throughput screening
LAH	lithium aluminium hydride
MES	2-(*N*-morpholino)ethanesulfonic acid
NMR	nuclear magnetic resonance
NRTI	nucleoside analog reverse transcriptase inhibitors
NtRTI	nucleotide analog reverse transcriptase inhibitors
PDB	protein data bank
Plm	Plasmepsin
PR_{I50V}	Ile50Val HIV-1 protease mutant
PR_{I84V}	Ile84Val HIV-1 protease mutant
PR_{WT}	wild-type HIV-1 protease

rmsd	root mean square deviation
RNA	ribonucleic acid
RT	room temperature
RP-MPLC	reversed phase medium pressure liquid chromatography
SAR	structure-activity-relation
SBDD	structure-based drug design
THF	tetrahydrofuran
VdW	Van der Waals
WHO	World Health Organization

1. Introduction 1

1.1. Backround 1
1.2. The human immunodeficiency virus 1
1.3. HAART 4
1.4. HIV-1 protease 4
  1.4.1. Structural aspects of HIV-1 protease 5
  1.4.2. Functional aspects of HIV-1 protease 7
  1.4.3. Substrate specificity of HIV-1 protease 7
1.5. Inhibitors of HIV-1 protease 10
  1.5.1. Peptidomimetic Inhibitors 11
  1.5.2. Non-pepdidic Inhibitors 13
1.6. Resistance development 16
1.7. Motivation and incipient studies 18
1.8. References 23

2. Targeting the Open Flap Conformation of HIV-1 Protease with Pyrrolidine-based Inhibitors 29

2.1. Introduction 29
2.2. Results and Discussion 32
  2.2.1. Synthesis 32
  2.2.2. Biological data 33
  2.2.3. Structural Analysis 34
    2.2.3.1. Binding mode of α 36
    2.2.3.2. Binding mode of β 37
2.3. Discussion 39
2.4. Summary and Conclusion 40
2.5. Experimental Section 41
2.6. References 44

3. Structure-Guided Design of C_2-Symmetric HIV-1 Protease Inhibitors Based on a Pyrrolidine Scaffold 49

3.1. Introduction 49
3.2. Chemistry 51

3.3. Results and Discussion 54
3.4. Summary and Conclusion 64
3.5. Experimental Section 65
3.6. References 69

4. Structural and Kinetic Analysis of Pyrrolidine-based Inhibitors of the Drug Resistant Ile84Val Mutant of HIV-1 Protease 73
4.1. Introduction 73
4.2. Results 76
  4.2.1. Kinetic characterization 76
  4.2.2. Structural analysis 79
    4.2.2.1. Crystal structures of 8 81
    4.2.2.2. Crystal structures of 9 84
4.3. Discussion 87
4.4. Summary and Conclusion 89
4.5. Experimental Section 90
4.6. References 92

5. Two Solutions for the same Problem: Multiple Binding Modes of Pyrrolidine-based HIV-1 Protease Inhibitors 97
5.1. Introduction 97
5.2. Results and Discussion 99
  5.2.1. Chemistry 99
  5.2.2. Biological evaluation 101
  5.2.3. Structural Analysis 102
    5.2.3.1.Binding mode in the orthorhombic form 104
    5.2.3.2.Binding mode in the hexagonal form 105
5.3. Discussion 106
5.4. Summary and Conclusion 108
5.5. Experimental Section 109
5.6. References 110

6. Achiral Oligoamines as Versatile Tool for the Development of Aspartic Protease Inhibitors 113
6.1. Introduction 113
6.2. Results 117
  6.2.1. Chemistry 118
  6.2.2. Kinetic characterization 119
  6.2.3. Structural analysis 121
    6.2.3.1. Binding mode of 6d 123
    6.2.3.2. Binding mode of 9b 124
6.3. Discussion 125
6.4. Summary and Conclusion 127
6.5. Experimental Section 128
  6.5.1. Enyzme assays 128
  6.5.2. Structural Analysis 129
6.6. References 130

7. Chasing Binding Modes in HIV Protease: From Seemingly Perturbed to Seemingly Relaxed Pose Without Altering Affinity 136
7.1. Introduction 136
7.2. Results and Discussion 137
7.3. Summary and Conclusion 141
7.4. References 143

8. Summary / Zusammenfassung 145
8.1. Summary 145
8.2. Zusammenfassung 150

1 Introduction

1.1 Background

Nearly three decades ago, the first cases of the acquired immunodeficiency syndrome (AIDS) were reported in the United States.[1] An anomalously high number of patients suffered from numerous unusual life-threatening opportunistic infections, e.g. an otherwise rarely occurring form of pneumonia (*P. carinii*) indicating that the immune response in these patients was suppressed. This newly occurring epidemic seemed to cause a gradual destruction of the patient’s immune response and was therefore termed early on “acquired immunodeficiency syndrome” (AIDS). Before the discovery of AIDS being caused by infection with the human immunodeficiency virus (HIV), most of the patients died within two years.[2] Today, AIDS has reached pandemic proportions: The world health organization (WHO) estimates that about 33.2 million people live with the disease worldwide and that about 2.5 million people get newly infected every year. Approximately 2 million people died of AIDS-related conditions last year.[3] In the past 20 years, chemotherapy against HIV infection has been tackled intensively by modern drug discovery and development. Several stages in the viral replication cycle (chapter 1.2) have been targeted to reduce the viral load thus delaying the progression to AIDS. However, an entire remedy of the infection or a vaccination is still an unaccomplished goal.

1.2 The human immunodeficiency virus

HIV is a lentivirus and belongs to the class of retroviruses. In lentiviruses the genetic information is encoded in the viral single-stranded positive-sense enveloped RNA and upon host cell entry the viral genome is transcribed into double-stranded DNA. Lentiviruses are characterized by a long incubation period between the initial infection and the final outbreak of the disease. The HI-viron is schematically represented in Figure 1. It has a spherical shape and a diameter of about 100nM. The viron’s envelope consists of a lipid bilayer, originally derived from an infected human cell. The HIV envelope contains the surface and transmembrane glycoprotein gp160, composed of the subdomains gp120 and gp41 which are essential for the adsorption and penetration processes. The viral matrix protein (p17) serves as a structural protein, the capsid protein (p24) builds up the nucleocapsid, which contains two identical strands of single-stranded RNA. The (+)-ssRNA is protected by a high amount of the nucleocapsid protein. In addition to these proteins, HIV-1 encodes at least six regulatory proteins (Tat, Rev, Nef, Vif, Vpr, Vpu) and three viral enzymes, the HIV protease, the reverse transcriptase, and the integrase. Two variants of HIV are differentiated, the HIV-1 and HIV-2 type. Whereas HIV-1 is widely spread in America, Europe, and Asia, the HIV-2 type is mainly observed in West Africa.

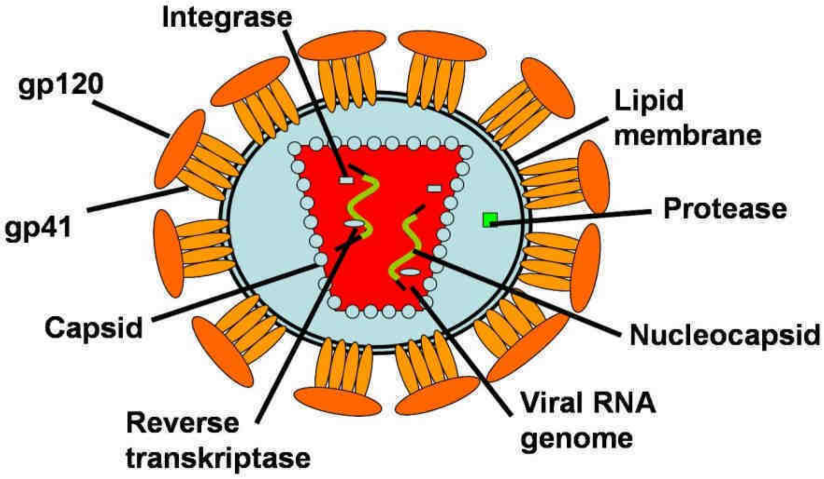


Figure 1: Schematic representation of the HI-viron, selected proteins are indicated and labeled

The HI-virus replicates inside human cells. The replication cycle is schematically represented in Figure 2. The HI virons bind with their envelope glycoprotein (gp120) to the human CD4 receptors present on the surface of host T-cells (1). For cell entry, HIV additionally requires one of two chemokine receptors, either CCR5 or CXCR4 (2).[4] Via binding of gp120 to CD4 and a co-receptor, the insertion of the second glycoprotein gp41 facilitates the fusion of the viral and the host cell membrane. Inside the cell, the viral RNA, which is still protected in the nucleocapsid, is released and uncoated (3).[5] The single-stranded viral RNA is then transcribed into double-stranded DNA by the reverse transcriptase (4)[6] and transported to the nucleus, where it is integrated into the host cell genome by the viral integrase (5).[7] In this state, the viral DNA is referred to as provirus and the host cell is now latently infected with HIV. By activation of the immune cell, the virus's genetic information is consecutively transcribed into RNA molecules using the protein machinery of the cell (6) and then translated into HIV proteins (7).[8] These expressed components assemble near the cell membrane (8) and than undergo the process of budding and release from the host cell (9).[9] The viral protease plays a crucial role in the maturation process of the virons. In this process, the *gag* and *pol* polypeptide chains are cleaved into their functional units (10).

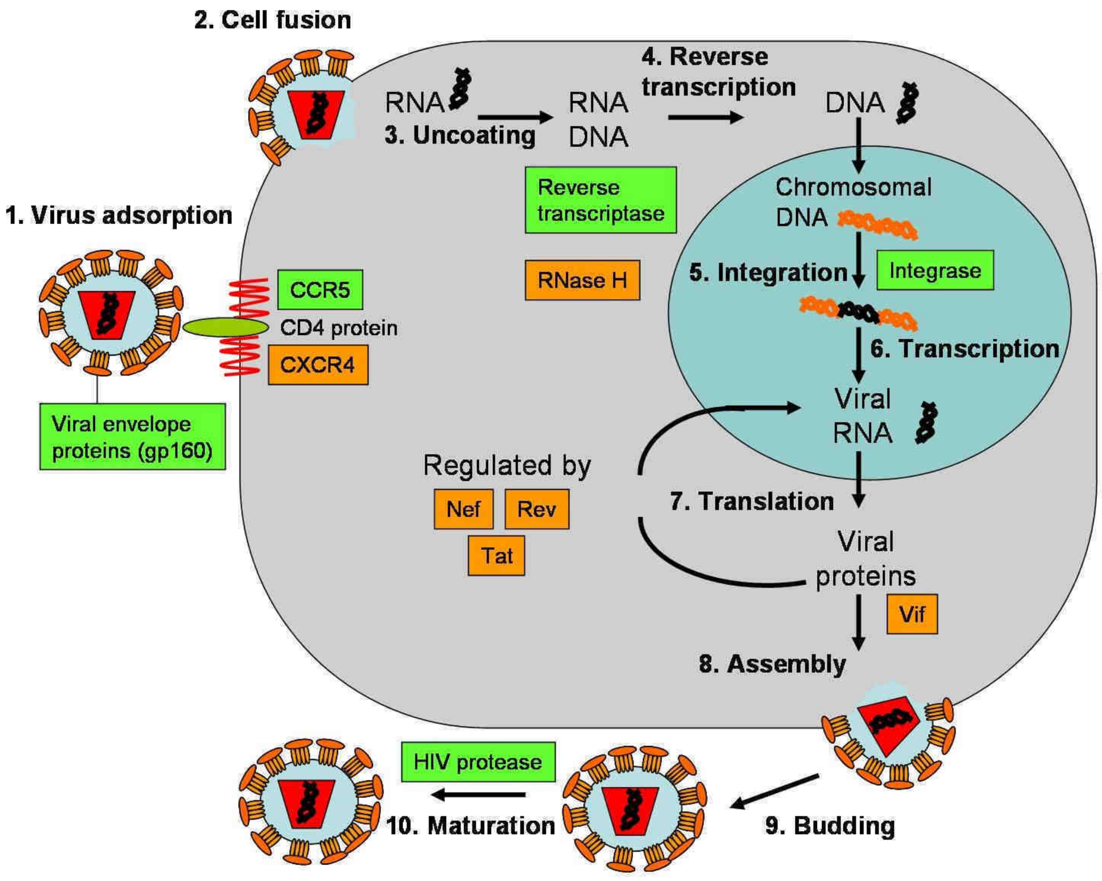

2. Cell fusion
RNA
3. Uncoating
RNA
DNA
4. Reverse transcription
DNA
Reverse transcriptase
RNase H
Chromosomal DNA
5. Integration
Integrase
6. Transcription
Viral RNA
1. Virus adsorption
CCR5
CD4 protein
CXCR4
Viral envelope proteins (gp160)
Regulated by
Nef
Rev
Tat
7. Translation
Viral proteins
Vif
8. Assembly
HIV protease
10. Maturation
9. Budding

Figure 2: Replication cycle of HIV, numbering corresponds to the text. Current targets of approved drugs are colored in green (gp160, reverse transcriptase, HIV protease, integrase and chemokine receptor 5 (CCR 5)), and possible additional targets are colored in orange (chemokine receptor 4 (CXCR 4), RNase H, Tat, Rev, Nef, and Vif).[10]

Several steps of the HIV life cycle have been identified as promising drug targets. Today, inhibitors of the viral enzymes reverse transcriptase, protease, and integrase as well as inhibitors of the cell entry process are available.[10]

1.3 HAART

The development of the Highly Active Anti-Retroviral Therapy (HAART) for treatment of the HIV infection is one of the great success stories of modern medicine: In the past 25 years, the mortality of HIV-positive patients could be reduced dramatically and changed AIDS from a nearly completely fatal illness into a nowadays often manageable chronic illness.[11] HAART not only significantly prolongs but also improves the quality of the patient's life.[12, 13] The first approved drug for the treatment of HIV was an inhibitor of the viral reverse transcriptase *Zidovudine (AZT).* The class of reverse-transcriptase inhibitors can be divided into nucleoside (NRTI) and nucleotide (NtRTI) inhibitors on the one hand and non-nucleoside inhibitors on the other hand.[14] In this content several different lead structures had been approved, however, the rapid occurrence of drug resistant strains demanded the development of new drugs possessing a different mode of action. The HIV-protease was identified early on as promising target and several inhibitors of this protease have meanwhile been approved. All recommended HAART regimens generally include at least one non-nucleoside reverse-transcriptase inhibitor or HIV protease inhibitor combined with two nucleoside or nucleotide reverse-transcriptase inhibitors. In 2003, the entry inhibitor enfuvirtide[15] was approved and very recently raltegravir[16] the first inhibitor of the viral integrase has been launched to the market. However, despite the success of the antiretroviral therapy, it still suffers from major drawbacks. Many of the applied drugs possess relatively short half-lifes, low bioavailability, poor permeability and often cause severe side effects. Besides this, the development of drug-resistant strains of HIV increasingly hampers an effective therapy (chapter 1.6) thus demanding the development of new drugs with improved properties.

1.4 HIV-1 protease

By comparing the genomic sequence of HIV-1 with that of other retroviruses, *Ratner et al.* postulated that the genome encodes for a protease.[17] This was confirmed by the identification of the Asp-Thr-Gly motif which is characteristic for retroviral aspartic proteases.[18] It has been shown that inactivation of the viral protease by either mutation or chemical inhibition leads to immature, noninfectious virus particles.[19, 20] Since that discovery, unprecedented efforts were made to solve the three-dimensional structure of this promising target enzyme. Hampered by the fact that the protease is only a minor component of HI-virons, a synthetic or recombinant access to the protein material was required.[21, 22] Both techniques were finally successful and enabled the determination of the structure by NMR and X-ray techniques.[23-25]

1.4.1 Structural aspects of HIV-1 protease

Today, HIV protease is probably one of the most studied and hence best characterized enzymes: More than 250 crystal structures are assigned to its EC number (EC 3.4.23.16) in the publicly available protein data bank (PDB).[26] Successful crystallization is reported under a variety of different crystallization conditions and the deposited apo and complex structures adopt several different space groups. The most frequently observed ones are shown in Table 1.

number of protein structures	space group	cell dimensions (Å)		
		a	b	c
82	$P2_12_12$	56-61	85-90	46-47
66	$P2_12_12_1$	51-53	57-61	61-63
58	$P6_1$/ $P6_122$	61-64	61-64	81-85
8	$P4_1$/ $P4_12_12$	45-50	45-50	100-110

Table 1: Most frequently observed space groups in crystal structures of HIV-1 protease.

First protein crystals obtained in absence of any ligand exhibited the space group $P4_12_12$ and the corresponding structures were determined and published in 1989 (PDB ID: 3PHV, 2HPV).[23, 24] These apo structures confirmed models postulating HIV-protease to be active only in its dimeric form. However, in these complexes the binding cavity was in a very open form and the flaps were about 7Å apart from the catalytically active aspartates. This open flap conformation was attributed to be a result of a kinetic trap during crystallization stabilized by extensive crystal contacts to symmetry-related molecules. The first crystal structures of protein ligand complexes (PDB ID: 4HVP ($P2_12_12_1$),[27] 5HVP ($P2_12_12$)[28] and 9HVP ($P6_1$)[29]) unveiled the flap region being in a closed conformation covering the catalytic dyad, thus leading to the active site in a closed, tunnel-shaped conformation. However, the solution structure of the HIV protease monomer resembles the open flap conformation. Additionally, NMR-experiments revealed high flap flexibility and molecular dynamics suggest a similar energy content for the open and the closed flap conformations in the unbound state.[30, 31]

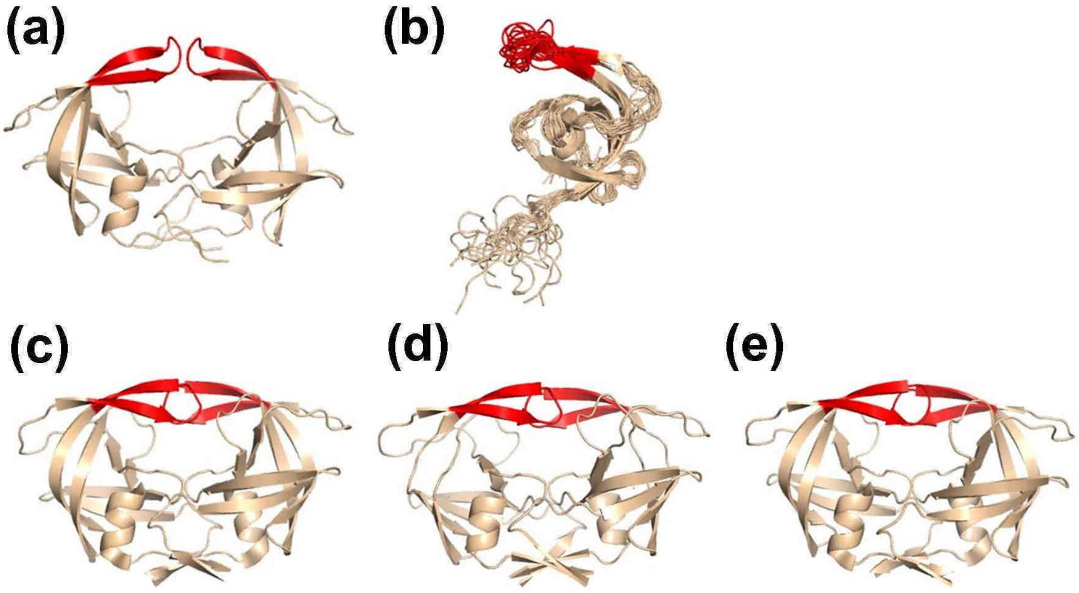
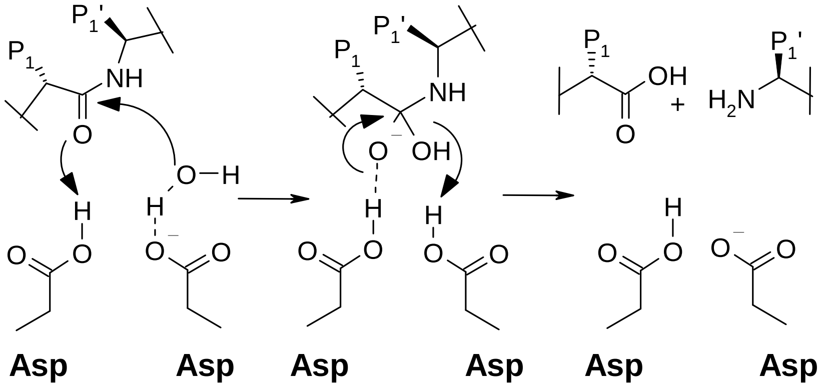


Figure3: HIV protease structures, the protein is represented as cartoon in wheat and the flexible flap region is colored in red. (a) Apo crystal structure in space group $P4_12_12$ (pdb ID: 2PC0). (b) NMR ensemble of 20 structures (pdb ID: 1Q9P). (c-e) X-ray structures of HIV protease inhibitor complexes (c) $P2_12_12_1$ (PDB ID: 2NMW), (d) $P2_12_12$ (PDB ID: 4PHV), and (e) $P6_1$ (PDB ID: 1HPV).

Besides this high flexibility all studies confirmed common structural features: The protease is active as C_2-symmetric homodimer and each monomer consists of 99 amino acids. The secondary structure of each monomer includes one α-helix and two antiparallel β sheets. Even though each monomer contains two cysteine residues, the dimer does not contain any disulfide bonds but is stabilized by non-covalent interactions. The interface of the two monomers is mainly formed by hydrophobic interactions and each monomer contributes one of the two catalytically active aspartates. The active site triad (Asp25-Thr26-Gly27) is located in a loop region stabilized by a network of hydrogen bonds called fireman's grip, in which the Thr26 hydroxyl group is addressed by the main-chain NH of the same residue of the other chain and vice versa. The carboxylate groups are orientated nearly coplanar and are in close contact. A conserved water molecule mediates the interactions of a substrate or inhibitor to the β-hairpins (flap) covering the active site.

1.4.2 Functional aspects of HIV-1 protease

The detailed catalytic mechanism of HIV-1 protease is up to now not fully understood.[32] The similar structural features compared to non-retroviral aspartic protease, however, suggest a similar mechanism. The catalytic mechanism of non-viral aspartic proteases has been extensively studied by kinetic methods, affinity labeling, and X-ray crystallography. Most studies are consistent with a general acid-base mechanism postulated by *Suguna et al* in which it is assumed that one aspartate is in the deprotonated and the other in the protonated state (Scheme 1).[33]

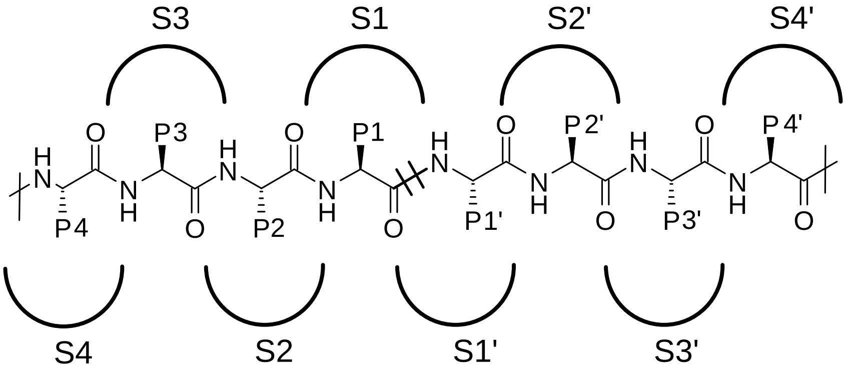


Scheme 1: Schematic representation of the catalytic mechanism of aspartic proteases.[33]

Activation of a water molecule facilitates the nucleophilic attack at the scissile peptide bond leading to a tetrahedrally coordinated *gem*-diol intermediate, which collapses under cleavage of the peptide bond thus releasing the hydrolysis products.

1.4.3 Substrate specificity of HIV-1 protease

Aspartic proteases in general recognize 6-10 amino acids of their natural polypeptide substrates. The standard nomenclature defines the substrate residues as e.g. P3, P2, P1, P1’, P2’, P3’ and the corresponding recognition pockets as e.g. S3, S2, S1, S1’, S2’, S3’ as shown in Scheme 2.[34]



Scheme 2: Nomenclature of the protease's subsites according to Berger and Schlechter, [34] the scissile peptide bond is indicated by crossing lines.

The natural substrates of HIV protease are the *gag* and *pol* polyproteines. At least nine different cleavage sites are recognized by the protease, which are listed in Table 1.[35] Even though the protease is C_2-symmetric, the amino acid sequences of the substrates are asymmetric taking the scissile amide bond as center. In addition to the asymmetry, the substrates share only little sequence homology.

	Cleavage domains	**P4**	**P3**	**P2**	**P1**	**P1'**	**P2'**	**P3'**	**P4'**
***gag* polyprotein**	MA-CA	Ser	Gln	Asn	Tyr	Pro	Ile	Val	Gln
	CA-p2	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met
	P2-NC	Ala	Thr	Ile	Met	Met	Gln	Arg	Gly
	p2-p1	Gln	Ala	Asn	Phe	Leu	Gly	Lys	Ile
	p1-p6	Pro	Gly	Asn	Phe	Leu	Gln	Ser	Arg
***pol* polyprotein**	TF-PR	Ser	Phe	Asn	Phe	Pro	Gln	Ile	Thr
	PR-RT	Thr	Leu	Asn	Phe	Pro	Ile	Ser	Pro
	RT-RH	Ala	Glu	Thr	Phe	Tyr	Val	Asp	Gly
	RH-IN	Arg	Lys	Ile	Leu	Phe	Leu	Asp	Gly

Table 2: Sequences of the cleavage sites within the HIV-1 *gag* and *pol* polyproteins that are cleaved by HIV-1 protease. In column 2 the proteins released after cleavage of the substrates are indicated: matrix protein (MA), capsid (CA), nucleocapsid (NC), trans frame peptide (TF), protease (PR), reverse transcriptase (RT), RNAse H (RH), integrase(IN), and the structural proteins (p1/p2/p6).

In the P1 and P1’ position preferably hydrophobic amino acids and in P2/P2’position either smaller hydrophobic or polar amino acids are recognized. In more distal sites the recognition is non-uniform and residues with various physicochemical properties are tolerated. The mode, in which the different substrates interact with the HIV protease, has been elucidated on structural basis by the group of *C.A. Schiffer*. Several crystal structures of substrate analogue peptides in complex with an inactive Asp25Asn mutant were determined and one example, the complex of a synthetic decapeptide cleavage site (IRKIL-FLDGI) originating from the *pol*-polyprotein, is depicted in Figure 3.[36]

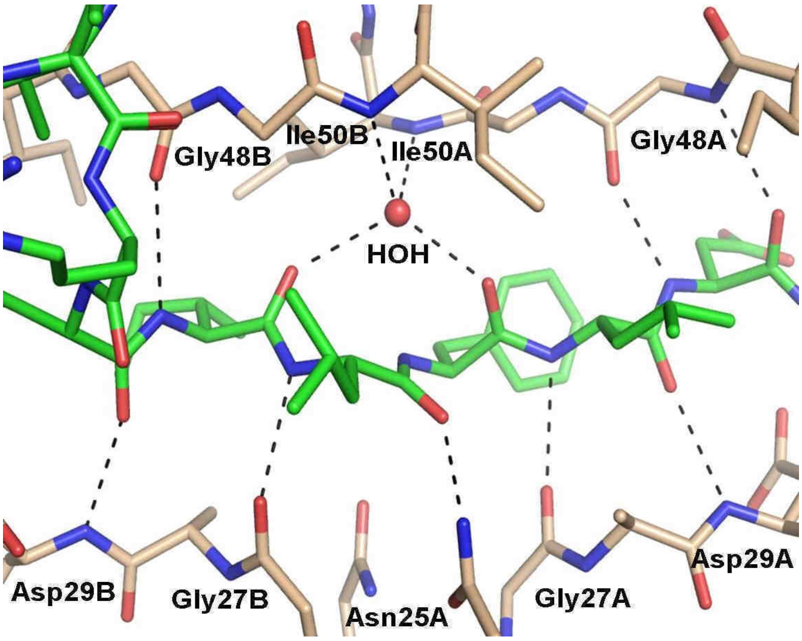


Figure 4: Crystal structure of a substrate analogue decapeptide, shown in green, color-coded by atom type, in complex with the HIV protease (pdb ID: 1KJH). The protein is represented in sticks and in wheat color-coded by atom type. Hydrogen bonds are indicated by dashed lines and the flap water is illustrated as red sphere.

The synthetic peptide is bound in an extended conformation and several hydrogen bonds to the peptide recognition motif comprising main-chain NHs of Asp29, Gly48, and main-chain carbonyl oxygen atoms of Gly27, and Gly48 of chain A and B, respectively, are formed. The substrate forms a parallel or antiparallel β-sheet with each of the flaps. The carbonyl oxygen atom of the scissile amide bond establishes a hydrogen bond to the amide nitrogen of one of the introduced asparagines (Asn25A). A tetrahedrally coordinated water molecule mediates polar contacts of the carbonyl oxygen atoms of the amide functionalities adjacent to the cleavage site to the flap Ile 50 NHs. In case of the mentioned complex, about 1200Å^2 of the peptide's surface is buried in the protease's binding pocket. In addition to the main-chain interactions, intensive hydrophobic interaction as well as additionally formed hydrogen bonds of the side-chains contribute to the overall affinity.

1.5 Inhibitors of HIV-1 protease

Currently 10 HIV protease inhibitors are approved for the treatment of the HIV infection and applied within the highly active antiviral therapy (HAART)(Table 1).[10]

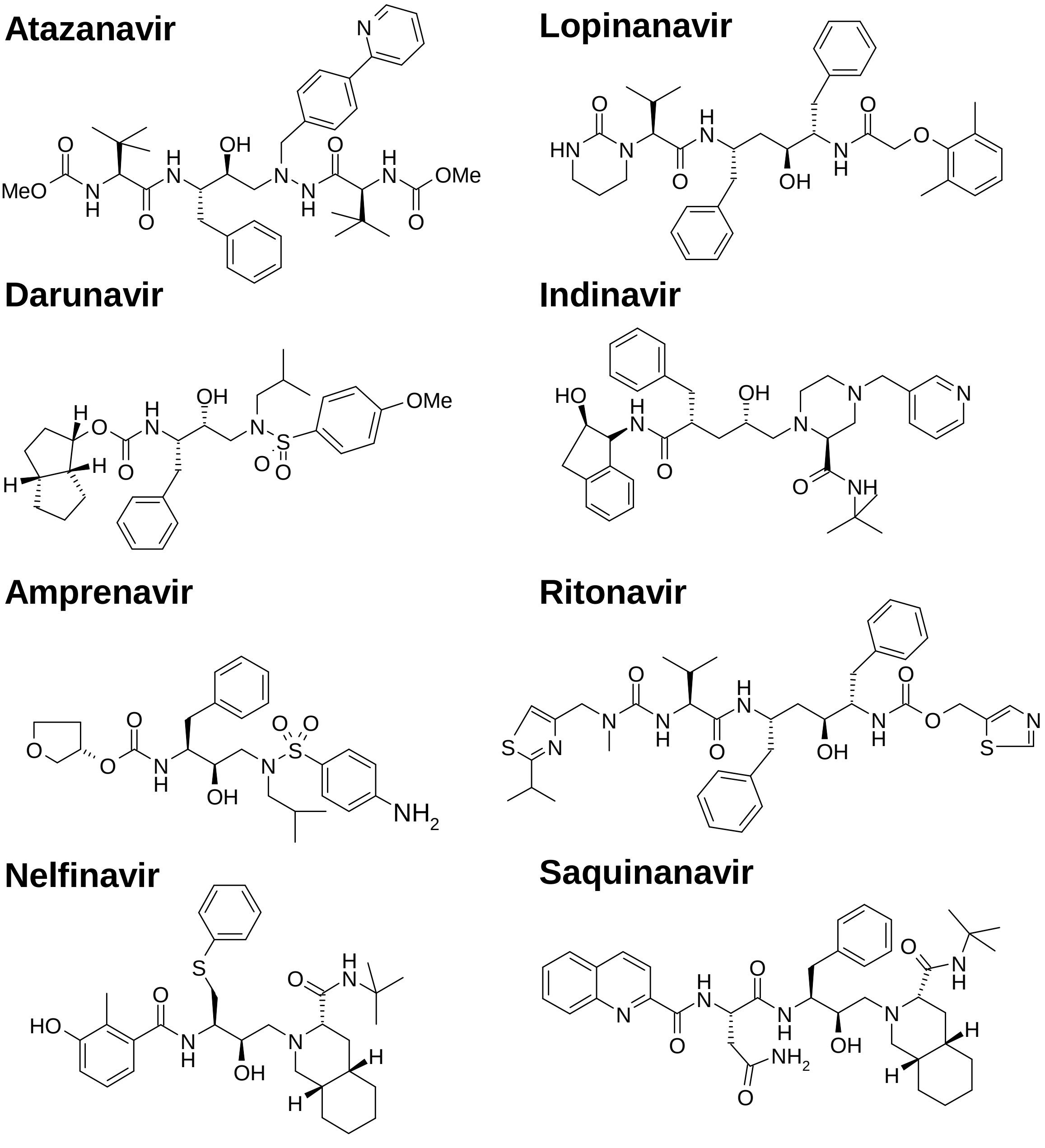
Name	Trade name	Company	Launched
Saquinavir	Invirase	Hoffmann-La Roche	1995
Indinavir	Crixivan	Merck	1996
Ritonavir	Norvir	Abbott, GlaxoSmithKline	1996
Nelfinavir	Viracept	Agouron, Pfizer	1997
Amprenavir	Agenerase, Prozei	Vertex	1999
Lopinavir	Kaletra, Aluvia	Abbott	2000
Atazanavir	Reyataz, Zrivada	Bristol–Myers Squibb, Novartis	2003
Fosamprenavir	Lexiva, Telzir	Vertex, GlaxoSmithKline	2003
Tipranavir	Aptivus	Boehringer Ingelheim	2005
Darunavir	Prezista	Tibotec	2006

Table 3: Licensed protease inhibitors, United States, 2007, listed by the releasing date.[10]

Initially, inhibitor development was mainly guided by the peptidic nature of the substrates. As so called substrate analogues they have a linear character and bear a non-cleavable isostere of the peptide bond in their central core. Several isosteres including statine, norstatine, phosphinate, reduced amide, dihydroxyethylene, α-keto amide, and silicon-based groups have been utilized. However, the continuously increasing resistance towards approved drugs (chapter 1.6) made the development of novel inhibitors bearing new scaffolds referred to as non-peptidic inhibitors necessary. Both classes of inhibitors, the substrate analogues as well as the non-peptidic inhibitors, will be presented in more detail in the two following chapters.

1.5.1 Peptidomimetic Inhibitors

All approved inhibitors exceed the affinities of the natural substrates by several orders of magnitude. The peptidomimetic compounds achieve this amongst others by their structural similarity to the tetrahedral *gem*-diol intermediate. The approved peptidomimetic inhibitors possess either a hydroxyethylene or hydroxyethylamine moiety to interact with the catalytic dyad Asp25A and Asp25B (scheme 3).


Atazanavir
Lopinanavir
Darunavir
Indinavir
Amprenavir
Ritonavir
Nelfinavir
Saquinanavir

Scheme 3: *Approved peptidomimetic HIV-1 protease inhibitors*

NMR experiments suggest that in the complexed state, similar to the apo HIV protease, one aspartate is in the protonated whereas the second one is in the deprotonated state.[37] The crystal structures of the inhibitors in complex with the HIV-1 protease reveal a similar binding situation for all representatives. In addition to the interactions with the catalytic dyad, water-mediated polar interactions are established to the flexible flap region. A water molecule, termed flap water, is similar to substrate HIV protease complexes, tetrahedrally coordinated by the acceptor groups flanking the transition state isoster and the NHs of the flap residues Ile50A and Ile50B. The binding mode of saquinavir as one representative of the class of peptidomimetic inhibitors is represented in Figure 5.

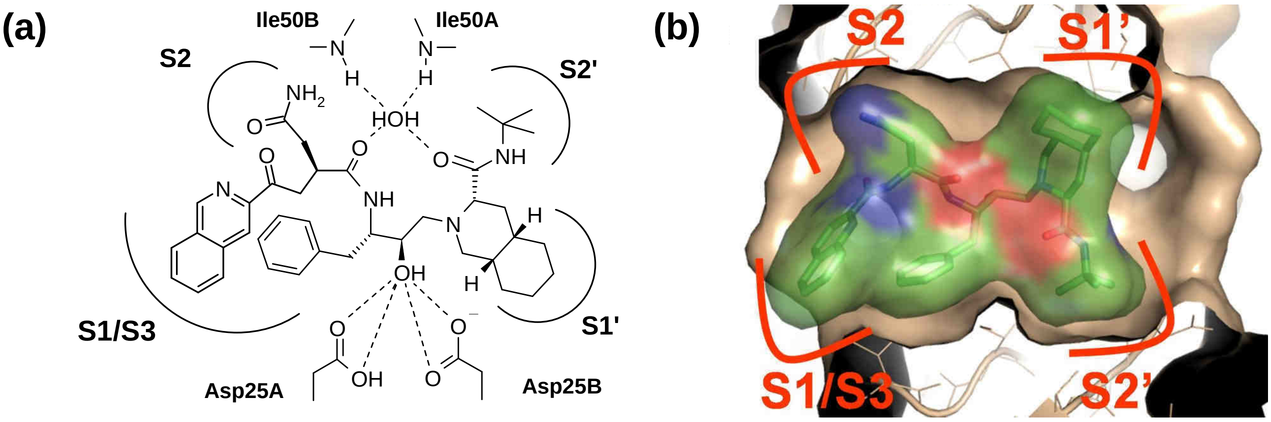
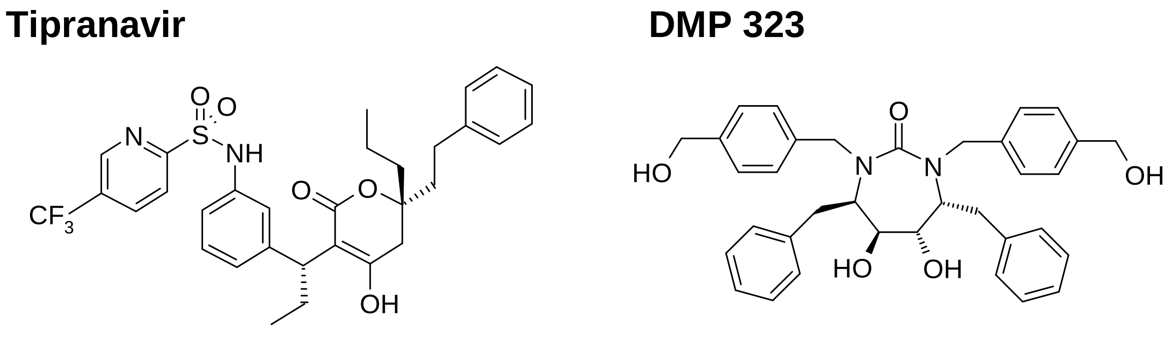


Figure 5: (a) Schematic representation of the binding mode of Saquinavir. Hydrogen bonds are indicated by dashed lines (b) View from the top of the protease with the flap region clipped off. The ligand skeleton and surface are color-coded in green and the protein is represented in wheat. The occupied specificity pockets are indicated and labeled in red.

The sub-pockets of the protease are addressed similarly to the natural substrates, successively from S3 to S3'. Like the natural substrates Saquinavir addresses the S1 and S1' with hydrophobic moieties; benzyl in case of the S1 pocket and a bulky decahydroquinoline moiety in case of the S1' pocket. Although the S2 and S2' pocket exhibits a predominantly hydrophobic environment, most of the approved inhibitors possess hydrophilic substituents, an asparagin side-chain in case of saquinavir. More distal binding pockets are not as well defined and are reached in diverse fashion by the different peptidomimetic inhibitors. Saquinavir reaches the S3 pocket according to the natural substrates with its quinoline ring system, whereas for example Atazanavir simultaneously addresses the S1 and S3 pocket with its large phenyl pyridine substituent. Other inhibitors like Amprenavir omit this binding pocket.

1.5.2 Non-peptidic Inhibitors

The class of non-peptidic HIV protease inhibitors comprises mostly cyclic compounds of which the dihydropyrones and the cyclic ureas have gained pronounced attention (Scheme 4).[38]

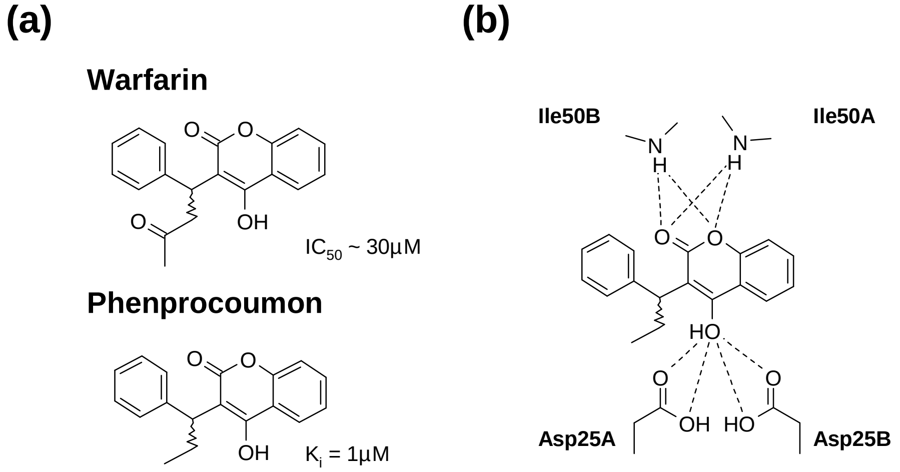

Tipranavir

DMP 323

Scheme 4: Non-peptidic HIV-1 protease inhibitors

Structural analysis of the complexes between Tipranavir or DMP 323 and the HIV-1 protease revealed that in both cases the structural water, present in substrate as well as peptidomimetic inhibitor bound complexes, is repelled from the complex. Whereas in case of the dihydropyrones this replacement of the flap water was a happenstance, it was intended in the original design concept of the cyclic ureas. Based on entropic considerations, this displacement should be energetically favorable. Additionally, improved selectivity towards other mammalian aspartic proteases can be expected by this mode of binding, because the water-mediated flap interactions are unique for HIV protease.[39] Both inhibitor developments are prime examples of modern drug discovery projects and will be described in detail in the following.

Tipranavir was developed from a non-peptidic coumarin template that had been discovered by high-throughput screening.[40] Upjohn (today Pfizer) screened a compound collection in a fluorescence-based HIV-1 protease assay and 4-hydroxycoumarin (Warfarin) was identified as a weak inhibitor (IC50 ~ 30µM). After the identification of this initial lead, a more focused screening led to the identification of phenprocoumon (Marcumar®) exhibiting a K_i of 1µM. The determination of the crystal structure in complex with the HIV protease enabled further rational drug design.



Scheme 5: (a) Initial screening hits used for the development of Tipranavir. (b) Schematic representation of the binding mode of phenprocoumon observed in the crystal structure in complex with the HIV protease.

Based on the complex structure, iterative cycles comprising synthesis, biological evaluation, computational methods and structural analysis were performed to optimize the initial lead structure.[41] Optimizations involved the benzopyran ring which was replaced by a 5,6-dihydropyrone ring system hence facilitating the introduction of further substituents thought to address additional enzymatic sub-pockets. Additional favorable binding interactions were gained by introduction of a *p*-trifluoromethyl sulfonamide in *meta*-position of the phenyl ring of the core structure. The binding mode of the finally approved inhibitor Tipranavir is represented in Figure 6.

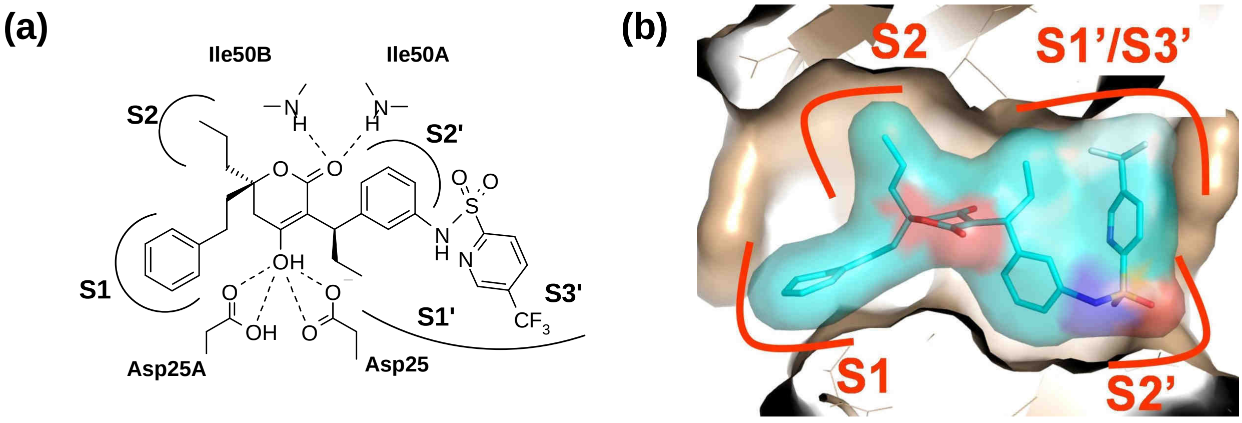


Figure 6: (a) Schematic representation of the binding mode of Tipranavir. Hydrogen bonds are indicated by dashed lines. (b) View from the top of the protease with the flap region clipped off. The ligand skeleton and surface are color-coded in light blue and the protein is represented in wheat. The occupied specificity pockets are indicated and labeled in red.

The two hydrophobic substituents at the central 5,6-dihydropyrone ring system address the S1 and S2 pockets of the enzyme. The third substituent occupies with its ethyl group the S1' pocket whereas the phenyl moiety points to the S2' pocket. From this orientation, the S3' pocket is reached by the *para*-trifluoromethyl substituent. Additional to the polar interactions formed to the catalytic dyad and the flap isoleucins, the inhibitor addresses parts of the peptide recognition motif with its sulfonamide functionality establishing hydrogen bonds to the main-chain NHs of Asp29A and Asp30A. Tipranavir possesses a binding affinity of $K_i = 8$ pM.

Whereas the staring point for the development of Tipranavir was a high-throughput screening, a chance-based method, the corresponding starting point for the development of the cyclic ureas was based on rational methods. Studies by DuPont Merck using the structural information obtained from extensive work on peptidic lead structures as well as the observation of a conserved water molecule mediating the interaction to the flap region were exploited to generate a vector model hypothesis facilitating the search for suitable lead structures in publicly available crystallographic databases.[42] Consecutively, molecular modeling predictions suggested a simple core ring system which incorporated the hydrogen-bonding equivalents of the flap water molecule into the inhibitor scaffold (Figure 7a). Early C_2-symmetric lead compounds already exhibited low nanomolar affinity and finally led to the first clinical candidate, DMP 323.[43] The design strategy was confirmed by crystal structure analysis and revealed that the inhibitors are bound symmetrically in the active site of the protease (Figure 7). The urea oxygen atom accepts two hydrogen bonds from the flap Ile50A and Ile50B and the diol of the ring system interacts with the two catalytic aspartic acids, Asp25A and Asp25B.

The recognition pockets S1, S2, S1', and S2' of the HIV protease are addressed separately by the corresponding substituents branching from the rigid ring system. In case of DMP323, the S1 and S1' pockets are occupied by benzyl moieties and the S2 and S2' pockets with *p*-(hydroxymethyl)-benzyl groups. The hydroxyl functionalities address the main-chain NHs of Asp29 and Asp30 of the A and B chain at the far end of the pocket, respectively.

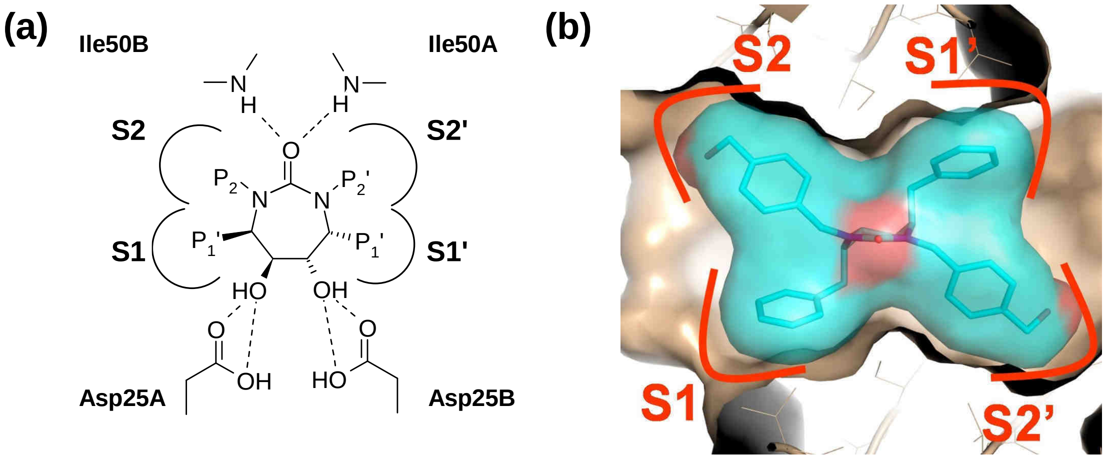


Figure 7: (a) Schematic representation of the binding mode of DMP323 (pdb ID: 1QBS). Hydrogen bonds are indicated by dashed lines (b) View from the top of the protease with the flap region clipped off. The ligand skeleton and surface are color-coded in light blue and the protein is represented in wheat. The occupied specificity pockets are indicated and labeled in red.

DMP 323, however, failed in Phase I clinical trials due to highly variable pharmacokinetics. Although it was not the only clinical candidate from this series, up to now no cyclic urea has made it to the market, either due to a poor pharmacokinetic profile or due to high susceptibility to crucial active site mutations.[44]

1.6 Resistance development

The emergence of drug resistant strains strongly complicates the treatment of the HIV infection. Like many other RNA viruses, the replication of HIV is highly error prone. The repeatedly occurring resistance is a direct consequence of the high genetic diversity of the virus caused by the error-prone reverse transcriptase (~ 1error per 10.000 bases) and the high *in vivo* replication rate (10^8-10^9 replications per day) of the virus. It has been estimated that up to 10^5 point mutations occur daily within the viral population of an infected individual. This high mutation rate leads to dynamic distributions of related but non-identical genomes termed viral quasispecies.[45] In this mixture of HIV strains the molecular targets of the antiviral drugs are present with various single-point amino acid substitutions. HAART is aimed to impede the viral replication, however, this is not achieved completely. Under the selection pressure of antiviral drugs, resistant HIV variants showing reduced susceptibility to the applied drug treatment regime are singled out.

In case of the HIV protease, one has to discriminate between major and minor mutations.[46] Major mutations are initially selected by drug treatment and occur at residues forming direct interactions upon drug binding. These mutations lead to structural changes in the substrate cleft and reduce the binding affinity of the approved drug molecule.[47, 48] Minor mutations consecutively improve the replicative fitness of virus variants carrying these major mutations. The effects of these non-active site mutations are diverse and involve effects on dimer stability, inhibitor binding kinetics as well as long-range structural perturbations.[49] Mutations which are unambiguously associated with protease inhibitor therapy are listed in Table 4.

	D30	**V32**	**V33**	**M46**	**I47**	**G48**	**I50**	**I54**	**L76**	**V82**	**I84**	**N88**	**L90**
ATV			F	IL	V	V	L	VALM		AF	V	S	M
DRV		I	F		VA		V	LM	V	F	V		M
FPV		I	F	IL	VA		V	LM	V	F	V	S	M
IDV		I		IL	V			VALM	V	AFTS	V		M
LPV		I	F	IL	VA		V	VALM	V	AFTS	V		M
NFV	N		F	IL	V	V		VALM		AFTS	V	DS	M
SQV						V		VALM		TF	V	S	M
TPV		I	F	IL	V			VA		TFSL	V		M

Table 4: Selected mutations associated with protease inhibitor resistance. [50]

The different types of point mutations can occur in nearly one half of the amino acid positions of the protease.[51] In Figure 8, possible positions of point mutations that occur during treatment with the peptidomimetic inhibitor Saquinavir are illustrated as red spheres.

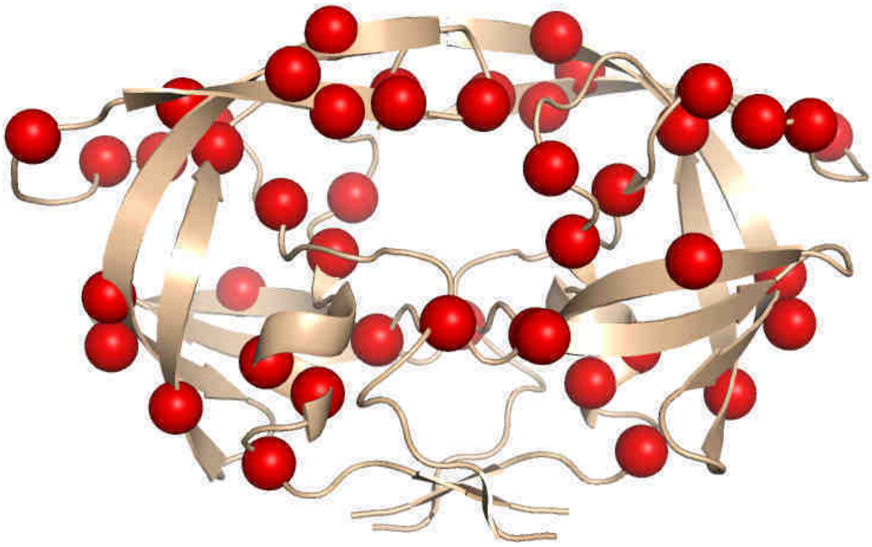


Figure 8: (a) Schematic representation of single point mutations observed under treatment with Saquinavir.

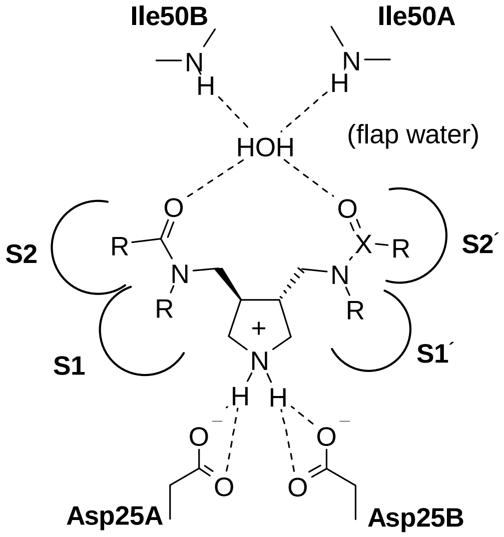
The reason for the drop in affinity of a certain inhibitor confronted with a certain mutation is manifold and specific for every single case. However, increasing cross-resistance to several or even all the approved inhibitors has been observed thus demanding a continuous development of new inhibitors. These inhibitors should exhibit a different mutation profile compared to that of the marketed drugs to circumvent the development of cross resistance.

1.7 Motivation and incipient studies

Because a similar mode of action of all approved inhibitors has resulted in pronounced cross resistance, the need of a steadfast and continuous search for new inhibitors is evident. An increased structural diversity of inhibitor scaffolds could be a possible strategy to at least diminish the accelerated development of multidrug resistant variants. Therefore, the objective of this thesis was to elucidate the suitability of cyclic and acyclic amines to serve as novel core structures for the structure-based design of HIV-protease inhibitors utilizing molecular and structural biology methods. In collaboration with medicinal chemists, using SAR, X-ray crystallography and modeling tools, initial lead structures should be developed and optimized with respect not only to affinity but also regarding their resistance profile towards crucial point mutants.

In the late 1990s, substituted piperidines were discovered in a HTS-screening at Hoffmann-La Roche and further developed as inhibitors of the aspartic protease renin. [52, 53] Based on this discovery, cyclic amines have been utilized as non-peptidic core-structure for the development of inhibitors of β-secretase and plasmepsin, which, like renin, belong to the class of pepsin-like aspartic proteases.[54-56] In a previous study in our group, substituted pyrrolidines were successfully developed as HIV-1 protease inhibitors in cooperation with Bayer AG.[57, 58]

The design concept of these inhibitors combined key structural elements from classical peptidomimetics (chapter 1.3) with a non-peptidic heterocyclic core structure comprising an endocyclic amino function. This functionality was intended to address the catalytic aspartates Asp25A and Asp25B and is implemented in the selected pyrrolidine ring system (Scheme 6)



Scheme 6: General structure of 3,4-dimethyleneamino-pyrrolidine based inhibitors and the binding mode intended in their initial design concept. Desirable hydrogen bonds are indicated by dashed lines

Furthermore, the pyrrolidine ring was equipped with two acceptor moieties, either carbonyl- or sulfone groups, supposed to form hydrogen bonds to the flap water molecule present in substrate or peptidomimetic inhibitor complexes. To estimate the inhibitory potential of the developed core structure, it was decorated with side-chains already optimized for HIV-protease inhibitors. Docking studies suggested the *S*,*S* enantiomer to be preferred in addressing the binding pocket of the protease appropriately. However, due to synthetic ease, an achiral synthetic route was initially chosen rendering the putative inhibitors as racemic mixture. The most potent derivative of the series, compound **1,** exhibited a K_i of 1,5 µM towards HIV-1 protease. The subsequent crystal structure analyses of **1** in complex with the protease revealed the binding mode depicted in Figure 9.

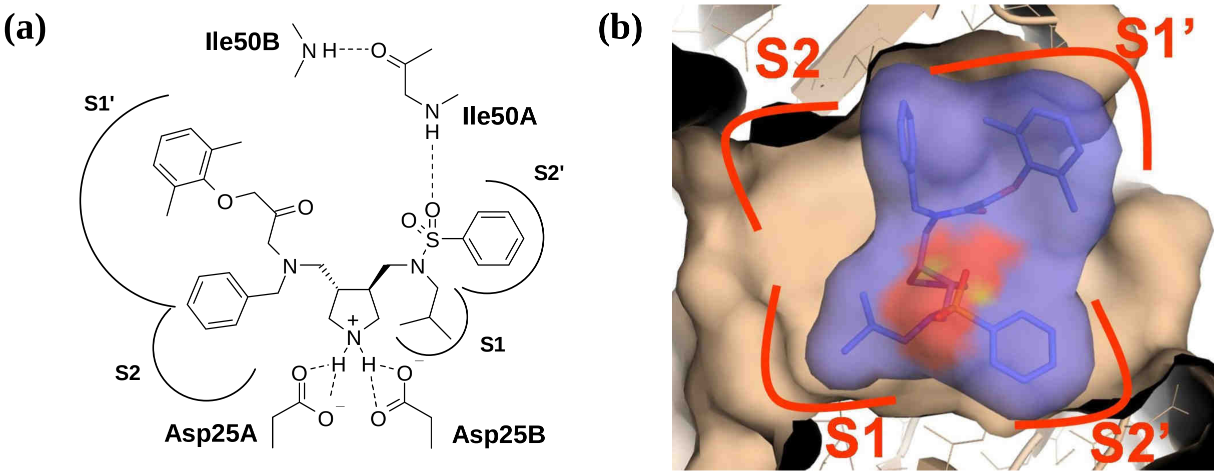


Figure 9: (a) Schematic representation of the binding mode of the most potent 3,4-dimethyleneamino-pyrrolidine derivative. Hydrogen bonds are indicated by dashed lines (b) View from the top of the protease with the flap region clipped off. The ligand skeleton and surface are color-coded in magenta and the protein is represented in wheat. The occupied specificity pockets are indicated and labeled in red.

In agreement with the docking result, the central pyrrolidine moiety was found at the pivotal position between both catalytic aspartates (Figure 9). However, apart from this successful prediction, significant deviations from the proposed docking modes were observed. Whereas the initial design concept suggested the *S*,*S* enantiomer to be the more active one, the crystal structure revealed the *R*,*R* enantiomer being bound in the active site of the protease. Surprisingly, the flap water, which was intended to be addressed by the polar acceptor groups, was repelled from the complex and the inhibitor adopted an up to now unique binding mode: It forms a direct hydrogen bond to the backbone NH group of Ile50B with one of its sulfonyl oxygen atoms. The carbonyl oxygen atom of the inhibitor’s amide group does not form any polar contacts to the enzyme. The Ile50A NH, lacking its usual hydrogen bond partner, the flap water, finds a surrogate in the main-chain carbonyl oxygen atom of Ile50B. The sub-pocket occupancy deviates strongly from the binding mode of the inhibitors presented in Chapter 1.4-1.5 and the binding mode intended in the design concept. Compared to peptidomimetic HIV protease inhibitors (chapter 1.5.1) a rotation of about 90° of the inhibitor is observed. Therefore, a successive occupation of the sub-pockets S1/S2 and S1'/S2' with each of the substituents at the pyrrolidine ring is impossible. Thus, the *iso*-butyl substituent of the sulfonamide part of the inhibitor addresses the S1 pocket whereas the connected phenyl moiety occupies the S2' pocket. For the two remaining substituents, the *N*-benzyl and 2,6-dimethylphenoxy moieties, a distinct pocket assignment is hardly possible. Both substituents share an enlarged space, formally assigned to the S1' pocket, and the *N*-benzyl group slightly streaks the S2 pocket. The hydrophobic 2,6-dimethylphenoxy group is solvent-exposed and may contribute to the overall high perturbations on protein side. Taking all the observed structural features into account, the overall affinity for the racemic mixture of 1,5 µM is remarkable and encouraged us to further optimize the initial scaffold.

The thorough interpretation of the obtained complex structure was the starting point for this theses. However, for the further optimization of the initial scaffold, it is essential to gain insights whether the misplaced and bulky *o,o'*-dimethylphenoxy substituent is responsible for the perturbed binding mode or whether this is caused by the orientation of the acceptor moieties with respect to the cyclic amino functionality. A small series lacking either this sterical demanding side-chain, the benzyl moiety, or both substituents has been designed and synthesized in order to answer this crucial question (Table 5).

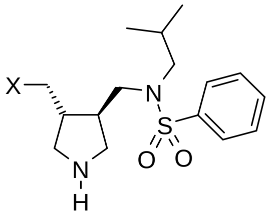
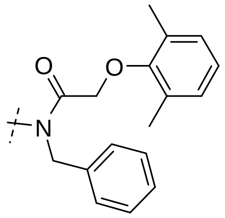
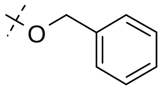
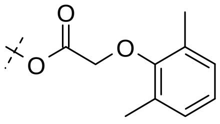
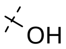
		
	1 (K_i = 1,5µM)	**2 (**K_i = 52µM)
		
	3 (K_i = 58µM)	**4 (**IC_{50} ~1mM)

Table 5: Chemical structure and affinity data of clipped 3,4-dimethyleneamino-pyrrolidine derivatives. Listed data correspond to racemic mixtures.

The binding affinities of derivatives lacking either one or two side-chains decrease indicating that the former substituents contribute positively to the overall binding affinity. Even though all compounds exhibited only weak affinity, it was possible to obtain a cocrystal structure of derivative **2** with a resolution of 1.75 Å. As expected, the flap water molecule is repelled from the complex and the sulfonyl group forms a direct hydrogen bond to Ile50NH. The flap distortion remains unchanged whereas the unmodified side-chains swap their positions (Figure 10).

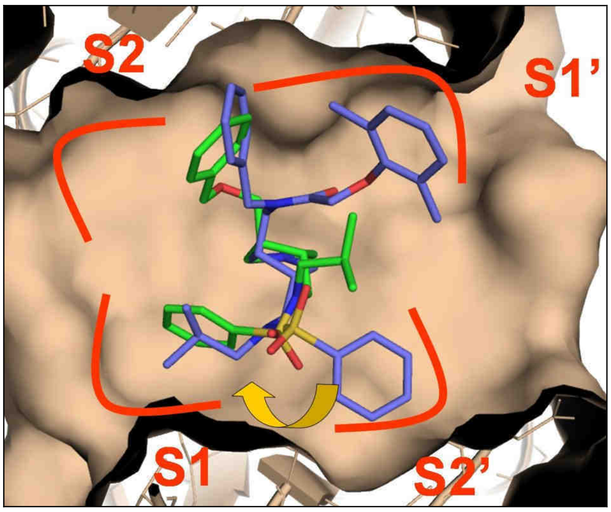


Figure 10 : C_α-superposition of the cocrystal structures of **1** (blue, color-coded by atom type) and **2** (green, color-coded by atom type). The protein surface is schematically represented in wheat and the occupied specificity pockets are indicated and labeled in red.

The benzene sulfonamide group occupies the S1 pocket formerly addressed by the *iso*-butyl moiety which is now located at the inner part of the S2’ pocket. The benzyl ether moiety nearly retains the former position of the benzyl substituent. These observations suggest that instead of the bulky dimethylphenoxy side-chain the spacial orientation of the cyclic amino functionality and the connected acceptor groups are the driving forces for protein deformations.

The results obtained by these preliminary studies forced us to revise our initial design concept. The following chapters represent in chronological order the further development of HIV protease inhibitors based on a pyrrolidine scaffold. All studies have been perused in cooperation with Dr. Andreas Blum, who was amongst others responsible for the synthesis of potential inhibitors. Inhibitor molecules presented in this theses have also been synthesized by

Stefanie Dörr (Chapter 5) and Dipl.-Chem. Benedikt Sammet (Chapter 6 and 7). The reaction pathways for the synthesis of all inhibitors relevant to this thesis have been implemented for reader benefits. Kinetic investigations of the inhibitors towards Plasmpepsins, presented in Chapter 6 have been performed by Dr. Torsten Luksch.

1.8 References

1. Pneumocystis pneumonia—Los Angeles. *MMWR Morb. Mortal. Wkly. Rep.* **1981**, 30, 250-252.
2. Rothenberg, R.; Woelfel, M.; Stoneburner, R.; Milberg, J.; Parker, R.; Truman, B., Survival with the acquired immunodeficiency syndrome. Experience with 5833 cases in New York City. *N. Engl. J. Med.* **1987**, 317, (21), 1297-1302.
3. AIDS epidemic update : December 2007. *UNAIDS* **2007**.
4. Chan, D. C.; Kim, P. S., HIV Entry and Its Inhibition. *Cell* **1998**, 93, (5), 681-684.
5. Auewarakul, P.; Wacharapornin, P.; Srichatrapimuk, S.; Chutipongtanate, S.; Puthavathana, P., Uncoating of HIV-1 requires cellular activation. *Virology* **2005**, 337, (1), 93-101.
6. Kati, W. M.; Johnson, K. A.; Jerva, L. F.; Anderson, K. S., Mechanism and fidelity of HIV reverse transcriptase. *J. Biol. Chem.* **1992**, 267, (36), 25988-25997.
7. Maele, B. V.; Debyser, Z., HIV-1 Integration: an Interplay Between HIV-1 Integrase, Cellular and Viral Proteins. *AIDS Reviews* **2005**, 7, (1), 26-43.
8. Morrow, C. D.; Park, J.; Wakefield, J. K., Viral gene products and replication of the human immunodeficiency type 1 virus. *Am. J. Physiol. Cell Physiol.* **1994**, 266, (5), C1135-1156.
9. Gottlinger, H. G., The HIV-1 assembly machine. *AIDS* **2001**, 15 (5), 13-20.
10. Flexner, C., HIV drug development: the next 25 years. *Nat. Rev. Drug Discovery* **2007**, 6, (12), 959-966.
11. Pomerantz, R. J.; Horn, D. L., Twenty years of therapy for HIV-1 infection. *Nat. Med.* **2003**, 9, (7), 867-873.
12. Bartlett, J. A.; DeMasi, R.; Quinn, J.; Moxham, C.; Rousseau, F., Overview of the effectiveness of triple combination therapy in antiretroviral-naive HIV-1 infected adults. *AIDS* **2001**, (15), 1369-1377.

13. Gulick, R. M.; Mellors, J. W.; Havlir, D.; Eron, J. J.; Meibohm, A.; Condra, J. H.; Valentine, F. T.; McMahon, D.; Gonzalez, C.; Jonas, L.; Emini, E. A.; Chodakewitz, J. A.; Isaacs, R.; Richman, D. D., 3-Year Suppression of HIV Viremia with Indinavir, Zidovudine, and Lamivudine. *Ann Intern Med* **2000**, 133, (1), 35-39.
14. Spence, R. A.; Kati, W. M.; Anderson, K. S.; Johnson, K. A., Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Science* **1995**, 267, (5200), 988-993.
15. Schneider, S. E.; Bray, B. L.; Mader, C. J.; Friedrich, P. E.; Anderson, M. W.; Taylor, T. S.; Boshernitzan, N.; Niemi, T. E.; Fulcher, B. C.; Whight, S. R.; White, J. M.; Greene, R. J.; Stoltenberg, L. E.; Lichty, M., Development of HIV fusion inhibitors. *J. Pept. Sci.* **2005**, 11, (11), 744-753.
16. Deeks, S. G.; Kar, S.; Gubernick, S. I.; Kirkpatrick, P., Raltegravir. *Nat. Rev. Drug Discovery* **2008**, 7, (2), 117-118.
17. Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K. J.; Starcich, B.; Josephs, S. F.; Doran, E. R.; Rafalski, J. A.; Whitehorn, E. A.; Baumeister, K.; Ivanoff, L.; Petteway, S. R.; Pearson, M. L.; Lautenberger, J. A.; Papas, T. S.; Ghrayeb, J.; Chang, N. T.; Gallo, R. C.; Wong-Staal, F., Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **1985**, 313, (6000), 277-284.
18. Toh, H.; Ono, M.; Saigo, K.; Miyata, T., Retroviral protease-like sequence in the yeast transposon Ty 1. *Nature* **1985**, 315, (6021), 691-691.
19. Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A. F.; Scolnick, E. M.; Sigal, I. S., Active Human Immunodeficiency Virus Protease is Required for Viral Infectivity. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, 85, (13), 4686-4690.
20. McQuade, T. J.; Tomasselli, A. G.; Liu, L.; Karacostas, V.; Moss, B.; Sawyer, T. K.; Heinrikson, R. L.; Tarpley, W. G., A synthetic HIV-1 protease inhibitor with antiviral activity arrests HIV-like particle maturation. *Science* **1990**, 247, (4941), 454-456.
21. Cheng, Y.-S. E.; McGowan, M. H.; Kettner, C. A.; Schloss, J. V.; Erickson-Viitanen, S.; Yin, F. H., High-level synthesis of recombinant HIV-1 protease and the recovery of active enzyme from inclusion bodies. *Gene* **1990**, 87, (2), 243-248.
22. Schnolzer, M.; Kent, S. B., Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* **1992**, 256, (5054), 221-225.
23. Navia, M. A.; Fitzgerald, P. M. D.; McKeever, B. M.; Leu, C.-T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P., Three-dimensional structure

of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **1989**, 337, 615-620.

24. Lapatto, R.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M., X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature* **1989**, 342, 299-302.
25. Yamazaki, T.; Nicholson, L. K.; Torchia, D. A.; Stahl, S. J.; Kaufman, J. D.; Wingfield, P. T.; Domaille, P. J.; Campbell-Burk, S., Secondary structure and signal assignments of human-immunodeficiency-virus-1 protease complexed to a novel, structure-based inhibitor. *Eur. J. Biochem.* **1994**, 219, (1-2), 707-712.
26. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M., The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, 112, (3), 535-542.
27. Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B.; Wlodawer, A., Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **1989**, 246, (4934), 1149-1152.
28. Fitzgerald, P. M. D.; McKeever, B. M.; VanMiddlesworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C.-T.; Herber, W. K.; Dixon, R. A. F.; Darke, P. L., Crystallographic Analysis of a Complex between Human Immunodeficiency Virus Type 1 Protease and Acetyl-Pepstatin at 2.0-Å Resolution. *J. Biol. Chem.* **1990**, 265, (24), 14209-14219.
29. Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N. N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M., Design, activity, and 2.8 Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease. *Science* **1990**, 249, (4968), 527-533.
30. Ishima, R.; Torchia, D. A.; Lynch, S. M.; Gronenborn, A. M.; Louis, J. M., Solution Structure of the Mature HIV-1 Protease Monomer: insight into the tertiary fold and stability of a precursor. *J. Biol. Chem.* **2003**, 278, (44), 43311-43319.

31. Wang, W.; Kollman, P. A., Computational study of protein specificity: The molecular basis of HIV-1 protease drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, (26), 14937-14942.
32. Ashraf Brik; Wong, C.-H., HIV-1 Protease: Mechanism and Drug Discovery. *ChemInform* **2003**, 34, (18).
33. Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. R., Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis*: Implications for a mechanism of action. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, 84, (20), 7009-7013.
34. Schechter, I.; Berger, A., On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, 27, (2), 157-162.
35. Rawlings, N. D.; Morton, F. R.; Barrett, A. J., MEROPS: the peptidase database. *Nucleic Acids Res.* **2006**, 34, D270-272.
36. Prabu-Jeyabalan, M.; Nalivaika, E.; Schiffer, C. A., Substrate Shape Determines Specificity of Recognition for HIV-1 Protease: Analysis of Crystal Structures of Six Substrate Complexes. *Structure* **2002**, 10, (3), 369-381.
37. Wang, Y. X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A., Solution NMR Evidence That the HIV-1 Protease Catalytic Aspartyl Groups Have Different Ionization States in the Complex Formed with the Asymmetric Drug KNI-272. *Biochemistry* **1996**, 35, (31), 9945-9950.
38. Chrusciel, R. A.; Strohbach, J. W., Non-Peptidic HIV Protease Inhibitors. *Curr. Top. Med. Chem.* **2004**, 4, 1097-1114.
39. Swain, A. L.; Miller, M. M.; Green, J.; Rich, D. H.; Schneider, J.; Kent, S. B. H.; Wlodawer, A., X-Ray Crystallographic Structure of a Complex Between a Synthetic Protease of Human Immunodeficiency Virus 1 and a Substrate-Based Hydroxyethylamine Inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, 87, (22), 8805-8809.
40. Thaisrivongs, S.; Tomich, P. K.; Watenpaugh, K. D.; Chong, K.-T.; Howe, W. J.; Yang, C.-P.; Strohbach, J. W.; Turner, S. R.; McGrath, J. P.; et al., Structure-Based Design of HIV Protease Inhibitors: 4-Hydroxycoumarins and 4-Hydroxy-2-pyrones as Non-peptidic Inhibitors. *J. Med. Chem.* **1994**, 37, (20), 3200-3204.
41. Suvit Thaisrivongs, J. W. S., Structure-based discovery of tipranavir disodium (PNU-140690E): A potent, orally bioavailable, nonpeptidic HIV protease inhibitor. *Peptide Science* **1999**, 51, (1), 51-58.

42. Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; al., e., Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **1994**, 263, (5145), 380-384.

43. Lam, P. Y. S.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; DeLucca, G. V.; Eyermann, C. J.; Chang, C. H.; Emmett, G.; Holler, E. R.; Daneker, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S. P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. M.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Jackson, D. A.; Erickson-Viitanen, S.; Hodge, C. N., Cyclic HIV Protease Inhibitors: Synthesis, Conformational Analysis, P2/P2' Structure-Activity Relationship, and Molecular Recognition of Cyclic Ureas. *J. Med. Chem.* **1996**, 39, (18), 3514-3525.

44. Rodgers, J. D.; Lam, P. Y. S.; Johnson, B. L.; Wang, H.; Ko, S. S.; Seitz, S. P.; Trainor, G. L.; Anderson, P. S.; Klabe, R. M.; Bacheler, L. T.; Cordova, B.; Garber, S.; Reid, C.; Wright, M. R.; Chang, C. H.; Erickson-Viitanen, S., Design and selection of DMP 850 and DMP 851: the next generation of cyclic urea HIV protease inhibitors. *Chem. Biol.* **1998**, 5, 597-608.

45. Coffin, J. M., HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* **1995**, 267, (5197), 483-489.

46. D’Aquila, R. T.; Schapiro, J. M.; Brun-Vézinet, F.; Clotet, B.; Conway, B.; Demeter, L. M.; Grant, R. M.; Johnson, V. A.; Kuritzkes, D. R.; Loveday, C.; Shafer, R. W.; Richman, D. D., Drug Resistance Mutations in HIV-1. *Top. HIV Med.* **2002**, 10, (5), 21-25.

47. Hong, L.; Zhang, X. C.; Hartsuck, J. A.; Tang, J., Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance *Protein Sci.* **2000**, 9, (10), 1898-1904.

48. Swairjo, M. A.; Towler, E. M.; Debouck, C.; Abdel-Meguid, S. S., Structural Role of the 30's Loop in Determining the Ligand Specificity of the Human Immunodeficiency Virus Protease. *Biochemistry* **1998**, 37, (31), 10928-10936.

49. Muzammil, S.; Ross, P.; Freire, E., A Major Role for a Set of Non-Active Site Mutations in the Development of HIV-1 Protease Drug Resistance. *Biochemistry* **2003**, 42, (3), 631-638.

50. HIV drug resistance database, Stanford University, http://hivdb.stanford.edu/hiv.

51. Ohtaka, H.; Muzammil, S.; Schön, A.; Velazquez-Campoy, A.; Vega, S.; Freire, E., Thermodynamic rules for the design of high affinity HIV-1 protease inhibitors with adaptability to mutations and high selectivity towards unwanted targets. *Int. J. Biochem. Cell Biol.* **2004**, 36, (9), 1787-1799.

52. Güller, R.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Hirth, G.; Jenny, C.; Kansy, M.; Montavon, F.; Müller, M.; Oefner, C.; Stadler, H.; Vieira, E.; Wilhelm, M.; Wostl, W.; Märki, H. P., Piperidine-Renin Inhibitors Compounds with Improved Physicochemical Properties. *Bioorg. Med. Chem. Lett.* **1999**, 9, (10), 1403-1408.

53. Vieira, E.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Güller, R.; Hirth, G.; Märki, H. P.; Müller, M.; Oefner, C.; Scalone, M.; Stadler, H.; Wihelm, M.; Wostl, W., Substituted Piperidines - Highly Potent Renin Inhibitors due to Induced Fit Adaptation of the Active Site. *Bioorg. Med. Chem. Lett.* **1999**, 9, (10), 1397-1402.

54. John, V.; Beck, J. P.; Bienkowski, M. J.; Sinha, S.; Heinrikson, R. L., Human β-Secretase (BACE) and BACE Inhibitors. *Journal of Medicinal Chemistry* **2003**, 46, (22), 4625-4630.

55. Prade, L.; Jones, A. F.; Boss, C.; Richard-Bildstein, S.; Meyer, S.; Binkert, C.; Bur, D., X-ray Structure of Plasmepsin II Complexed with a Potent Achiral Inhibitor. *Journal of Biological Chemistry* **2005**, 280, (25), 23837-23843.

56. Hof, F.; Schütz, A.; Fäh, C.; Meyer, S.; Bur, D.; Liu, J.; Goldberg, D. E.; Diederich, F., Starving the Malaria Parasite: Inhibitors Active against the Aspartic Proteases Plasmepsins I, II, and IV. *Angewandte Chemie, International Edition* **2006**, 45, (13), 2138-2141.

57. Specker, E.; Böttcher, J.; Brass, S.; Heine, A.; Lilie, H.; Schoop, A.; Muller, G.; Griebenow, N.; Klebe, G., Unexpected Novel Binding Mode of Pyrrolidine-Based Aspartyl Protease Inhibitors: Design, Synthesis and Crystal Structure in Complex with HIV Protease. *ChemMedChem* **2006**, 1, (1), 106-17.

58. Specker, E.; Böttcher, J.; Lilie, H.; Heine, A.; Schoop, A.; Müller, G.; Griebenow, N.; Klebe, G., An Old Target Revisited: Two New Privileged Skeletons and an Unexpected Binding Mode For HIV-protease Inhibitors. *Angew. Chem., Int. Ed.* **2005**, 44, (20), 3140-4.

2. Targeting the Open Flap Conformation of HIV-1 Protease with Pyrrolidine-based Inhibitors *

2.1 Introduction

Acquired Immune Deficiency Syndrome (AIDS) is caused by infection with the human immunodeficiency virus (HIV). Despite the immense efforts in combating the epidemic, the WHO estimates that currently approximately 40 million people are infected worldwide and thus AIDS remains one of the most serious health problems nowadays. Potent drugs targeting several stages in the viral life-cycle have been developed and approved.[1] The combination of inhibitors of the viral transcriptase and protease, termed highly active antiretroviral therapy (HAART), is currently the most effective treatment for HIV-infected patients.[2, 3] The introduction of HAART has increased the quality of the patient’s life dramatically, however, eradication of the virus still remains an unaccomplished goal.

HIV protease is a viral aspartic proteinase that processes the viral polyprotein gene products *gag* and *pol* into their functional units. The enzyme consists of 99 amino acids and is only active in dimeric form. It has been shown that inhibition of the protease leads to immature, non-infective virons, consequently making the enzyme an attractive antiretroviral drug target.[4] Unprecedented efforts in drug development made HIV protease to one of the most studied and hence best characterized enzymes: More than 240 crystal structures are assigned to its EC number (EC 3.4.23.16) in the publicly available protein data bank (PDB).[5] The first protein crystals were grown in absence of any ligand and exhibited the space group $P4_12_12$.

---

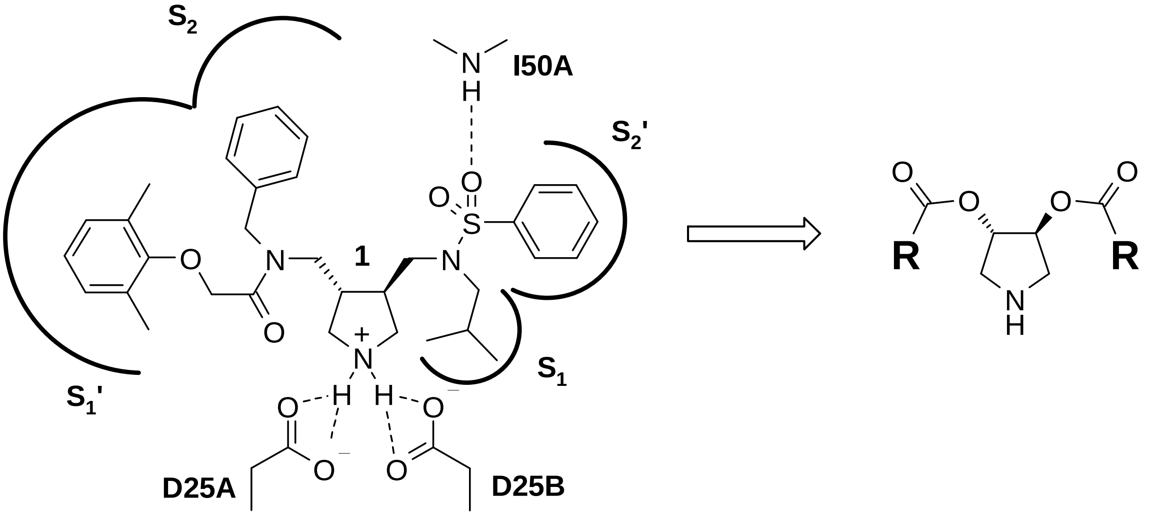
* Taken from original publication, Jark Böttcher, Andreas Blum, Stefanie Dörr, Andreas Heine, Wibke E. Diederich, Gerhard Klebe. Targeting the Open Flap Conformation of HIV-1 Protease with Pyrrolidine-based Inhibitors. *ChemMedChem* **2008**, 3, (9),1337-44.

The corresponding structures were determined and published in 1989 (PDB IDs: 3PHV, 2HVP) and revealed an expanded active site.[6, 7] The flaps covering the binding pocket exhibited a distance of more than 7Å to the catalytically active aspartates. Because the flap region possesses extensive crystal contacts to symmetry-related molecules, the observed flap conformation was attributed to be a result of a kinetic trap during crystallization. Indeed, further structural analyses of HIV protease inhibitor complexes unveiled that the flap region of the enzyme covers the catalytic dyad, thus leading to the active site in a closed, tunnel-shaped, conformation. The first crystal structures of protein ligand complexes (PDB IDs: 4HVP, 5HVP and 9HVP)[8-10] hosting peptidomimetic inhibitors revealed a conserved binding mode: The ligands address the catalytic dyad via a secondary hydroxyl group and the two flaps form a parallel and an antiparallel β-sheet with the ligand. Additionally, water-mediated interactions to two pivotal acceptor moieties of the ligand are formed by the Ile50 NHs being situated at the tip of the flaps. The crystal structure of substrate HIV protease complexed with an inactive Asp25Asn mutant showed a similar interaction pattern upon substrate binding (PDB ID: 1F7A).[11] The closed flap conformation has also been observed in case of the tethered dimer apo-structure in space group $P6_1$ (PDB ID: 1G6L).[12]

The open and the closed flap conformations were analyzed by molecular dynamics and for both similar energy contents were calculated in the unbound state.[13] NMR experiments also revealed high flap flexibility, and the solution structure of the HIV protease monomer resembles the open flap conformation.[14] Considering the multiple conformational states of the protease of nearly equal or close-by energy content, inhibitor design should be focused on addressing the protein in various states, either in the open, closed or intermediate conformations of the flap. Particularly in the drug development of kinase inhibitors, high affinity ligands have been developed against different states of the enzymes.[15]

The development of HIV-protease inhibitors targeting the closed conformation was extremely successful and resulted in nine FDA-approved drugs.[16] However, the clinical efficacy of these highly affinic inhibitors is strongly hampered by the increasing resistance development. The low fidelity of the viral reverse transcriptase and the fast replication rate lead to a high mutation rate thus providing HIV with a high degree of adaptability.[17, 18] This resulted already in resistant strains exhibiting decreased susceptibility towards all approved inhibitors, thus impeding an effective antiviral therapy. Therefore, the development of new drugs preferably possessing a different mode of binding thus circumventing the occurrence of cross-resistance is essential.

Most of the inhibitors developed so far address the catalytic dyad via a hydroxy functionality. Recently, cyclic amines have been developed as novel anchoring group.[19, 20] In a previous study, the pyrrolidine derivative **1** was designed and the racemic mixture exhibited an affinity of 1.5 µM towards HIV-1 protease. To elucidate its binding mode, a cocrystal structure was determined and revealed the *R*,*R*-enantiomer to bind with a unique interaction profile (Scheme 1). The endocyclic nitrogen binds to the catalytic dyad. *Poisson-Boltzmann* calculations suggest the amine being protonated and both aspartates being deprotonated, resulting in strong electrostatic interactions.[21] An asymmetric flap interaction pattern is observed (Scheme 1), resulting in an unsatisfactory occupation of the sub-pockets and strong deformations of the protein structure.



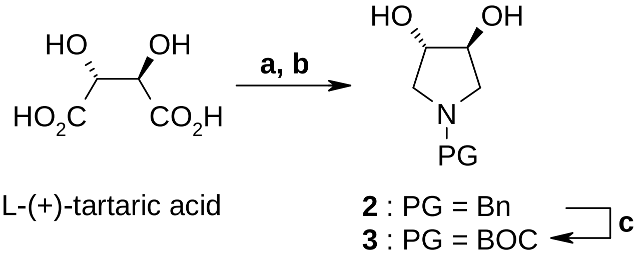
Scheme 1: Design of C_2-symmetric inhibitors starting from the cocrystal structure of **1**.

Both observations might explain the only moderate affinity obtained by a ligand of such size. Based on an in-depth analysis of the crystal structure with **1**, the following hypothesis was formulated: The pyrrolidine is a suitable anchoring group, however, the interactions to the flap region and the subsite occupancy should be improved. As an initial approach, symmetric pyrrolidine diesters with the required stereochemistry indicated in the crystal structure with **1** were designed, fulfilling the essential pharmacophore requirements. These compounds were decorated with two hydrophobic moieties to occupy the protease’s subpockets and equipped with acceptor groups attached closer to the central pyrrolidine ring compared to **1** in order to address the flap in a geometrically better suited way (Scheme 1).

2.2 Results and Discussion

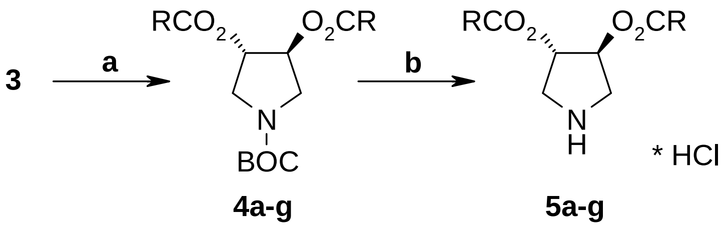
2.2.1 Synthesis

3,4-Difunctionalized chiral pyrrolidines are accessible starting from tartaric acid. The synthesis of the 3*S*,4*S*-pyrrolidine diols **2** and **3** has already been described in previously (Scheme 2).[22] Condensation of L-(+)-tartaric acid with benzylamine gave rise to the cyclic imide, which was further reduced utilizing $LiAlH_4$ in refluxing THF. The benzyl protecting group was replaced by the BOC-group via catalytic hydrogenation in presence of BOC_2O,[23] yielding the core structure **3** (Scheme 2).



Scheme 2: Synthesis of chiral pyrrolidine-diol **3** from tartaric acid:[22] a) $BnNH_2$, xylene, reflux, 81%; b) $LiAlH_4$, THF, reflux, 65%; c) H_2, Pd/C, BOC_2O, EtOH, 90%.

A series of diesters was prepared via condensation of **3** with the corresponding acid chlorides or the respective carboxylic acids after activation with EDC, giving rise to the enantiomerically pure diesters **4a-g** in high yields. The BOC-group was removed by treatment with HCl in Et_2O furnishing the inhibitors **5a-g** as hydrochlorides (Scheme 3). The resulting inhibitor series comprises alkyl moieties of different size (**5a-c**), the benzoyl derivative **5d** and arylacetic acid derivatives (**5e-g**).

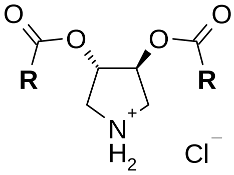


4a Isobutyl (81%)	**5a** Isobutyl (85%)
4b Neopentyl (91%)	**5b** Neopentyl (87%)
4c Cylcohexyl (77%)	**5c** Cylcohexyl (76%)
4d Phenyl (84%)	**5d** Phenyl (79%)
4e Benzyl (66%)	**5e** Benzyl (91%)
4f CH_2-1-Naphthyl (78%)	**5f** CH_2-1-Naphthyl (92%)
4g CH_2-2-Naphthyl (74%)	**5g** CH_2-2-Naphthyl (89%)

Scheme 3: Preparation of pyrrolidine-based inhibitors **5a-g**: a) RCOCl, NEt_3, CH_2Cl_2 or RCO_2H, EDC, CH_2Cl_2; b) HCl, Et_2O.

2.2.2 Biological data

The affinities of the pyrrolidine diesters **5a-g** were determined in a fluorescence-based assay towards three HIV-1 protease variants: the wild type originating from the BH10 isolate (PR_{WT}) and the active site point mutants Ile50Val (PR_{I50V}) and Ile84Val (PR_{I84V}). The affinities are given in Table 1. In case of PR_{WT} and PR_{I50V}, similar structure-activity relationships are observed. Whereas for the alkyl (**5a-c**) and phenyl (**5d**) substituted derivatives no affinity could be detected, the arylacetic acid derivatives (**5e-g**) exhibit a two-digit micromolar affinity. A more pronounced structure-activity relationship can be observed in case of PR_{I84V}: For all derivatives inhibition of PR_{I84V} was observed. The benzoyl-substituted inhibitor **5d** exhibits an affinity of 170 µM. In case of the alkyl derivatives **5a-c**, the affinity increases with the sterical demand of the substituents ranging from 280 µM for **5a** to 29 µM for **5c**. The arylacetic acid derivatives **5e-g** show the highest affinities in the series, particularly compounds **5e** and **5f** with a K_i of 3.0 µM and 4.5 µM, respectively. To elucidate the binding mode of this class of inhibitors, the most potent representatives **5e-g** were selected for cocrystallization experiments.



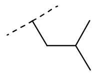
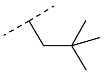
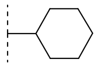
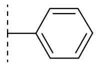
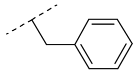
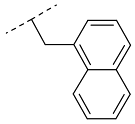
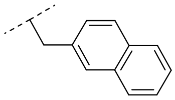
	R	**$\mathrm{PR}_{\mathrm{WT}}$** K_i [µM]	**$\mathrm{PR}_{\mathrm{I50V}}$** K_i [µM]	**$\mathrm{PR}_{\mathrm{I84V}}$** K_i [µM]
5a		n.i.	n.i.	280
5b		n.i.	n.i.	43
5c		n.i.	n.i.	29
5d		n.i.	n.i.	170
5e		18	25	3.0
5f		20	41	4.5
5g		20	15	15

Table 1. K_i-values of the inhibitors towards the wild type and mutant HIV proteases; n. i.: No inhibition observed.

2.2.3 Structural Analysis

Crystallization of HIV protease inhibitor complexes is well established in our laboratory, peptidomimetic and pyrrolidine-type inhibitors could be crystallized in complex with the protease exhibiting binding affinities ranging from low nanomolar to two-digit micromolar range. Surprisingly, the crystallization of the pyrrolidine diesters **5e-g** in complex with $\mathrm{PR}_{\mathrm{WT}}$ failed under the routinely applied standard crystallization conditions, however, utilizing the same conditions for $\mathrm{PR}_{\mathrm{I84V}}$ were successful. In presence of compound **5f**, large octahedral crystals were obtained and data were collected directly after transferring the crystals to cryo-protection conditions.

Remarkably, the grown crystals exhibited tetragonal symmetry in space group $P4_12_12$, typically adopted by apo-HIV protease for crystallization. In this space group, the asymmetric unit contains HIV-1 protease in its monomeric form. The functional protease dimer occupies a site coinciding with the symmetry element of a twofold axis. For clarity the amino acids of the protein chain A and its symmetry equivalent B are labeled as 1A to 99A and 1B to 99B, respectively. Due to the symmetry, all interactions within the A and the B chain are identical. For the description of contacts to other symmetry related monomers, the chain identifiers C and D are used (Figure 1).

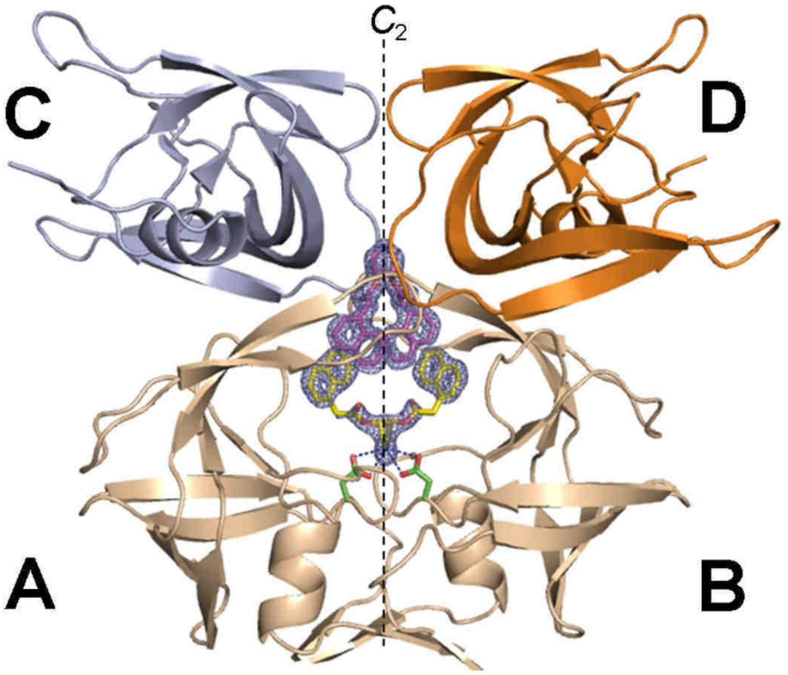


Figure 1: Crystal structure of **5f** (color-coded by atom type, **α** yellow, **β** magenta) in complex with HIV protease. The protein backbone trace is schematically illustrated in wheat for monomers A and B which form the functional dimer. It is related via the indicated crystallographic C_2-axis. Within this dimer the catalytic D25A and D25B are displayed with green carbon atoms. Additional selected protein monomers from crystal packing are shown as backbone trace in blue for C and orange for D. The F_o-F_c density for the ligands is displayed at a σ level of 2.5 as blue mesh. Figures 1-4 were created using PyMol 0.99.[24]

The observed protein conformation exhibits high similarity to the apo structures found in this space group, e.g. a C_α-rmsd of 0.18 Å is calculated with respect to PDB-entry 2PC0.[25] The notation of the protein subpockets, usually applied to HIV protease-inhibitor complexes, cannot be applied in our case, because these pockets are only regularly evolved once the flap is closed. Two molecules of **5f** are observed in the large active site cavity, comprising an area encompassed by the catalytic dyad and the flaps in the open conformation. Complex structures of HIV protease revealing a two to one inhibitor to protein ratio have been previously described, however, in contrast to the complex of **5f**, in these studies the second inhibitor molecule is found on the protein surface.[26, 27]

Both C_2-symmetric inhibitor molecules are centered on the C_2-axis (Figure 1). However, for both molecules two different binding modes are observed, which are not related by any symmetry. Consecutively, in the following both binding modes will be described separately: The ligand situated at the catalytic dyad is referred to as **α**, whereas the ligand next to the flap region is named **β**.

2.2.3.1 Binding mode of α

As intended, **5f** forms with its endocyclic amino functionality, hydrogen bonds to the catalytic dyad Asp25A (2.9Å, 2.8Å) and Asp25B (Figure 2).

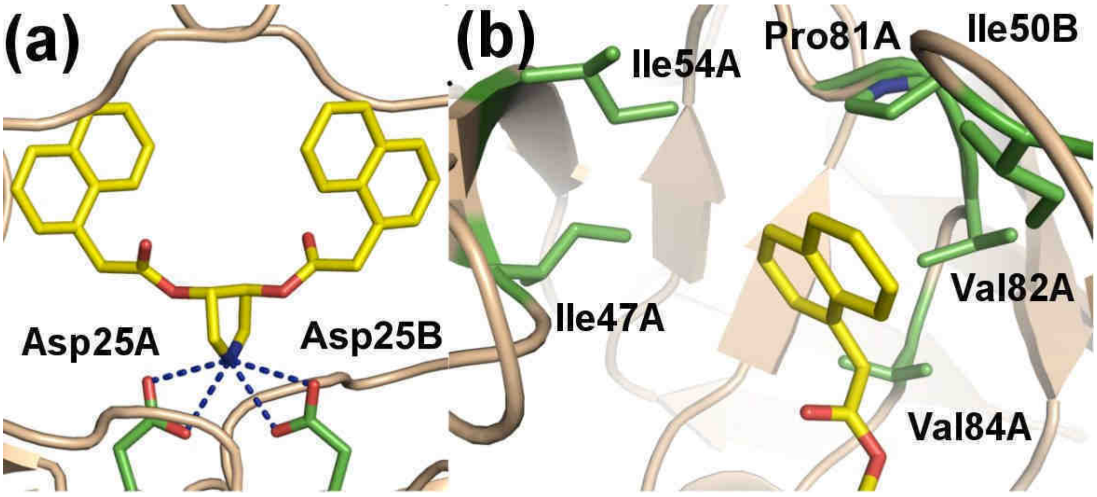


Figure 2: Binding mode of **α**, shown in yellow, color-coded by atom type; the protein backbone trace is schematically illustrated in wheat. Ligand **β** is omitted for clarity. (a): Polar interactions to the catalytic dyad are indicated by dashed lines; (b): Protein sidechains in contact with the naphthyl moiety are shown in green, color-coded by atom type.

These are the only polar interactions which can be observed; notably, the carbonyl ester groups thought to address the flap remain unsatisfied. The substituents at the pyrrolidine ring exhibit axial conformations. The naphthyl moieties establish numerous hydrophobic interactions to several residues of the widely opened protease binding pocket (Gly27, Ala28, Val32, Ile47, Ile50, Ile54, Thr80, Pro81, Val82, Val84 of chain A and B, respectively). Neglecting the second inhibitor molecule **β** occupying part of the binding pocket, ligand **α** would have to be classified as only poorly buried (62%). However, considering contacts to **β**, this value increases to 87%, leading to a pronounced burial of **α**. The naphthyl moieties of **α** also form contacts to the symmetry-related molecules C and D via Van der Waals contacts to the side-chain of Gln61 (Figure 4a). However, this contact leads only to a minor increase of the burial (88%).

2.2.3.2 Binding mode of β

The second inhibitor molecule **β** exclusively interacts with the flap region of the protease (Figure 3).

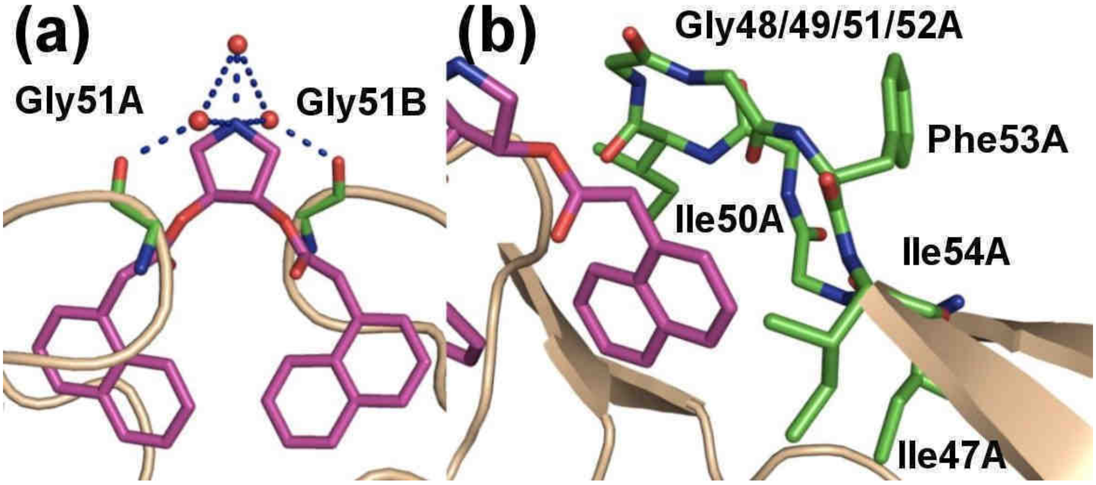


Figure 3: Binding mode of **β**, shown in magenta, color-coded by atom type, the protein backbone trace is schematically illustrated in wheat. Ligand **α** is omitted for clarity. (a): Polar interactions of the pyrrolidine-nitrogen indicated by dashed lines, waters by red spheres; (b): Amino acids in contact with the naphthyl moiety are shown in green, color-coded by atom type.

Van der Waals contacts are formed to the amino acids Ile47, Gly48, Gly49, Ile50, Gly51, Gly52, Phe53, and Ile54. In contrast to **α**, the substituents at the pyrrolidine ring now exhibit equatorial conformation. The amino functionality of the pyrrolidine ring forms a water-mediated hydrogen-bond network to the backbone carbonyls of Gly51A and Gly51B. Furthermore, the side-chain amide functionalities of Gln61C and Gln61D, respectively, establish hydrogen bonds to the ester carbonyl oxygen of the ligand (3.0 Å). The corresponding main-chain carbonyls of Gln61C and Gln61D contribute to the above described water-mediated hydrogen-bond interaction network involving Gly51A and Gly51B (Figure 4b).

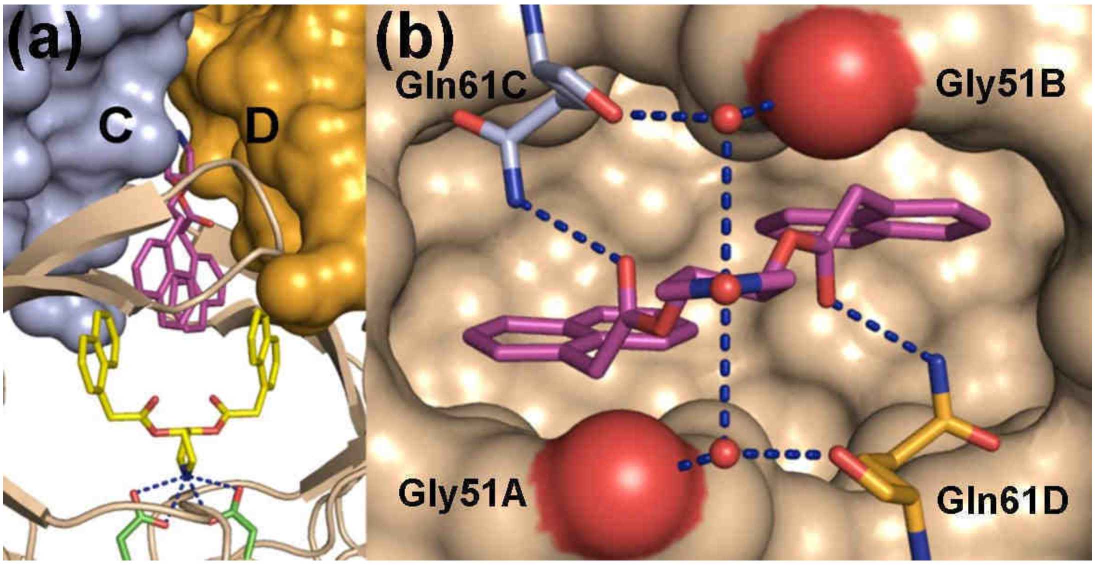


Figure 4: Interactions of **β** shown in magenta, color-coded by atom type. (a): The protein backbone trace of A and B is schematically illustrated in wheat, ligand **α** is shown in yellow, color-coded by atom type. The monomers C (blue) and D (orange) are shown in surface representation. (b): Polar interactions are indicated by dashed lines, waters by red spheres. The monomers A and B are shown in wheat surface representation except for the backbone carbonyl oxygen of Gly51 in red. The interacting Gln61 of chains C (blue) and D (orange) are color-coded by atom type. Ligand **α** is omitted for clarity.

Regarding only contacts formed by the inhibitor molecule **β** to the protease chains A and B, approximately half of the inhibitor's surface is buried (53%). However, additional contacts to the second inhibitor molecule **α** increase this value to 78%. The inhibitor molecule **β** also experiences numerous contacts to the further two symmetry-related HIV protease molecules C and D in the crystal packing. Additional Van der Waals contacts are formed to amino acids

Pro39, Ile62, Leu63, and Ile72 of chain C and D, respectively. Taking all contacts to interaction partners into account, ligand **β** is almost completely buried (96%).

2.3 Discussion

The described complex, obtained by crystallization in presence of **5f**, is the first HIV protease cocrystal structure showing the open-flap conformation of the enzyme in which the inhibitor addresses the catalytic dyad. In case of a previously described metallacarborane cocrystal structure of the HIV protease in an open flap conformation, the inhibitor molecules bind to the “upper” part of the active site, and no interaction of the catalytic dyad to the inhibitor molecules was observed.[28] A crystallization strategy starting with the crystal form of the apo protein and using soaking to prepare the complex likely prevents folding of the flap upon the accommodated ligand.[29] However, in our case the open form has been obtained from a cocrystallization experiment where the complex forms already in solution. Under such conditions any constraints from the crystal packing environment are avoided. Regarding the exceptional geometry of the cocrystal structure of **5f**, the question arises whether it resembles the binding situation in solution.

All inhibitors of our series show improved affinity against PR_{I84V} compared to the wild-type and PR_{I50V}. In the crystal structure of **5f**, the side-chain of Val84 is in Van der Waals contact to the methylene group adjacent to the ester carbonyl group of **α** (Figure 2b). The replacement of valine 84 by isoleucine would result in decreased surface complementarities and should concomitantly lead to a decrease in affinity towards PR_{WT} which is experimentally observed. The rigid structure of the ligand core and the stereochemistry at both chiral centers presets the orientation of the introduced substituents. This hampers accommodation of the ligand at this position, thus explaining the low affinity of the even more rigid benzoyl derivative **5d**, which lacks the flexible methylene linkers. Comparing the affinity of PR_{WT} and PR_{I50V}, no major differences are observed, which is in agreement with the only small contribution of Ile50 to protein ligand surface contacts in case of binding **α**. According to the observed structure-activity relationship, ligands with larger substituents exhibit improved affinity. This is in good agreement with the large binding pocket, which allows for hydrophobic interactions only. Summarizing all these observations, the structure-activity relationship seems to confirm the binding mode of **α** being the biologically relevant one.

In case of **β**, neither the side-chain of Val84 nor Ile50 is in contact with the inhibitor, thus the improved affinity towards PR_{I84V} is most likely not attributed to **β**. The naphthyl moieties establish contacts to the main-chain atoms of the flap residues, the side-chain of Ile54 being the only exception, creating a comparable polar environment. Taking only chains A and B into account, no direct polar interactions are established, and the burial of the ligand remains rather incomplete. Major contribution to its burial arises from interactions with chains C and D present in the crystal packing. However, they are not present in solution, leaving an unsatisfactory binding situation for **β**. It is likely that the presence of **β** might favor crystallization of this complex by stabilizing the flap region and by interactions to neighboring protein molecules. Both ligands **α** and **β** interact extensively via π-stacking of their naphthyl moieties with each other. In addition, both ligands as a bulk fill the large binding cavity. Such extensive inter-ligand interactions are only possible in case of inhibitor **5f**, providing a possible explanation, why crystallization only with this particular ligand has been successful.

2.4 Summary and Conclusion

Starting from the known inhibitor **1**, a novel type of pyrrolidine-based ligand skeleton was designed. The short and high yielding synthesis was performed via a *chiral pool* approach starting from L-(+)-tartaric acid. The resulting inhibitors **5a-g** show up to low micromolar affinity against HIV protease variants. A cocrystal structure of **5f** in complex with PR_{I84V} could be determined. Surprisingly, two inhibitor molecules **α** and **β** with mutually facing binding modes and extensive interligand contacts are accommodated in the active site. In the observed binding mode **α**, the pyrrolidine nitrogen addresses the catalytic dyad as intended by our structure-guided inhibitor design. The observed structure-activity relationship can be convincingly explained considering the binding mode of **α** as the relevant one under biological conditions as well. The second inhibitor molecule **β** is only used to stabilize the flap region and presumably supports assembly and crystal growth. The open-flap conformation of HIV protease has already been postulated as promising to be exploited as drug target.[30]

The novel pyrrolidine-based ligands **5** are the first reported inhibitors that accommodate the protein in this conformation at least in the crystalline state. A comparable binding mode with a closed flap would be impossible, taking the rigidity of the ligand core and the required reorganization of the protein into account. The remaining binding cavity volume would be insufficient to accommodate ligands of the size and shape of **5a-g**. The cocrystal structure of **5f** provides a valuable novel starting point for further development of HIV protease inhibitors possessing a different mode of binding compared to known drugs.

2.5 Experimental Section

Purification and crystallization of the HIV protease variants: HIV protease variants were expressed from Escherichia coli and purified as previously described.[31] The HIV protease inhibitor complex was crystallized at 18°C by the sitting-drop vapor diffusion method using a 1:1 ratio. Crystals were obtained by cocrystallization of the enzyme with an inhibitor concentration of 1mM, final DMSO concentration of 10%. 1 µL of the well buffer (0.1 M BisTris, pH 6.5, 3.0 M NaCl) was mixed with 1µL protein solution (50 mM NaAc, pH 6.5, 1 mM EDTA, 1 mM DTT) with an HIV protease concentration of 7 mg/mL. Crystals were obtained within a week and had octahedral shape. For cryo-protection the crystals were briefly soaked in mother liquor containing 25% glycerol.

Kinetic Assay: Enzymatic assays were performed in 172 µL assay buffer (100 mM MES, 300 mM KCl, 5 mM EDTA, 1 mg/mL BSA, pH 5.5) by the addition of substrate dissolved in 4 µL DMSO, distinct inhibitor concentrations dissolved in 4 µL DMSO and 20 µL HIV-1 protease in assay buffer to a final volume of 200 µL (final DMSO concentration 4%). The fluorogenic anthranilyl-HIV protease substrate (Abz-Thr-Ile-Nle-(*p*-NO_2-Phe)-Gln-Arg-NH_2) was purchased from Bachem. The hydrolysis of the fluorogenic substrate was recorded as the increase in fluorescence intensity (excitation wavelength 337 nm, emission wavelength 410 nm) over a time interval of 10 min during which the signal increased linearly with time.[32] The kinetic parameters of PR_{WT} (K_m = 14.6 µM), PR_{I50V} (K_m = 139 µM) and PR_{I84V} (K_m = 70 µM) were determined by the method of Lineweaver and Burk. IC_{50}-values were calculated using nonlinear regression curves for single site competitive binding analysis using the program Grafit. K_i values were calculated from the following equation: $K_i = IC_{50}\,[1 + (S/K_m)]^{-1}$ assuming a competitive binding mechanism with a one to one ratio.[33] The overall error of the measurement estimated to be ±40%.

Data collection, phasing and refinement: The data set was collected using a Rigaku R-AXIS IV image plate detector using Cu K_α radiation from an in-house Rigaku RU-H3R rotating anode. Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000.[34] The structure was determined by the molecular replacement method with Phaser,[35] the apo HIV protease structure (PDB ID: 2PC0) was used as the search model. The structure refinement was continued with SHELXL-97,[36] for each refinement step at least 10 cycles of conjugate gradient minimization were performed, with restraints on bond distances, angles and B-values. Intermittent cycles of model building were done with the program COOT.[37] The coordinates have been deposited in the PDB (http://www.rcsb.org/pdb/) with access code 3BC4.

	5f
Resolution (Å)	25-1.82
space group	$P4_12_12$
cell dimensions (Å)	a, b = 46.3
	c = 101.4
highest resolution shell (Å)	1.85 -1.82
no. of measured reflections	101919
no. of independent reflections	10446
completeness (%)[a]	99.3 [93.7]
I/σ[a]	42.2 [10.3]
R_{sym} (%)[a]	5.3 [15.9]
resolution in refinement (Å)	25-1.82
R_{cryst} (%) $F > 4\sigma F_o$; F_o	18,7; 19.3
R_{free} (%) $F > 4\sigma F_o$; F_o	24.1; 25.0
mean B-factor ($Å^2$) (protein)	22.5
mean B-factor ($Å^2$) ligand **α**	29.1
mean B-factor ($Å^2$) ligand **β**	21.3
mean B-factor ($Å^2$) water	28.2
Ramachandran plot	
most favorate geometry (%)	96.8
additionally allowed (%)	3.2
generously allowed (%)	-
disallowed (%)	-
RMSD bonds (Å)	0.007
RMSD angles (°)	2.0

Table 2. X-ray data processing and refinement for the PR_{I84V} complex of derivative **5f**. Values in brackets refer to the highest resolution shell.

2.6 References

1. Pomerantz, R. J.; Horn, D. L., Twenty years of therapy for HIV-1 infection. *Nature Medicine* **2003**, 9, (7), 867-873.
2. Bartlett, J. A.; DeMasi, R.; Quinn, J.; Moxham, C.; Rousseau, F., Overview of the effectiveness of triple combination therapy in antiretroviral-naive HIV-1 infected adults. *AIDS* **2001**, 15, (11), 1369-1377.
3. Gulick, R. M.; Mellors, J. W.; Havlir, D.; Eron, J. J.; Meibohm, A.; Condra, J. H.; Valentine, F. T.; McMahon, D.; Gonzalez, C.; Jonas, L.; Emini, E. A.; Chodakewitz, J. A.; Isaacs, R.; Richman, D. D., 3-Year Suppression of HIV Viremia with Indinavir, Zidovudine, and Lamivudine. *Annales of Internal Medicine* **2000**, 133, (1), 35-39.
4. Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A. F.; Scolnick, E. M.; Sigal, I. S., Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, 85, (13), 4686-4690.
5. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M., The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, 112, (3), 535-542.
6. Lapatto, R.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M., X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature* **1989**, 342, 299-302.
7. Navia, M. A.; Fitzgerald, P. M. D.; McKeever, B. M.; Leu, C.-T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P., Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **1989**, 337, 615-620.
8. Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B.; Wlodawer, A., Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **1989**, 246, (4934), 1149-1152.

9. Fitzgerald, P. M. D.; McKeever, B. M.; VanMiddlesworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C.-T.; Herber, W. K.; Dixon, R. A. F.; Darke, P. L., Crystallographic Analysis of a Complex between Human Immunodeficiency Virus Type 1 Protease and Acetyl-Pepstatin at 2.0-Å Resolution. *J. Biol. Chem.* **1990**, 265, (24), 14209-14219.

10. Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N. N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M., Design, activity, and 2.8 Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease. *Science* **1990**, 249, (4968), 527-533.

11. Prabu-Jeyabalan, M.; Nalivaika, E.; Schiffer, C. A., How Does a Symmetric Dimer Recognize an Asymmetric Substrate? A Substrate Complex of HIV-1 Protease. *J. Mol. Biol.* **2000**, 301, (5), 1207-1220.

12. Pillai, B.; Kannan, K. K.; Hosur, M. V., 1.9 Å X-Ray Study Shows Closed Flap Conformation in Crystals of Tethered HIV-1 PR. *Proteins: Structure, Function, and Genetics* **2001**, 43, (1), 57-64.

13. Rick, S. W.; Erickson, J. W.; Burt, S. K., Reaction Path and Free Energy Calculations of the Transition Between Alternate Conformations of HIV-1 Protease. *Proteins: Structure, Function, and Genetics* **1998**, 32, (1), 7-16.

14. Ishima, R.; Torchia, D. A.; Lynch, S. M.; Gronenborn, A. M.; Louis, J. M., Solution Structure of the Mature HIV-1 Protease Monomer: insight into the tertiary fold and stability of a precursor. *J. Biol. Chem.* **2003**, 278, (44), 43311-43319.

15. Liu, Y.; Gray, N. S., Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* **2006**, 2, (7), 358-364.

16. Wlodawer, A.; Erickson, J. W., Structure-Based Inhibitors of HIV-1 Protease. *Annu. Rev. Biochem* **1993**, 62, (1), 543-585.

17. Ohtaka, H.; Muzammil, S.; Schön, A.; Velazquez-Campoy, A.; Vega, S.; Freire, E., Thermodynamic rules for the design of high affinity HIV-1 protease inhibitors with adaptability to mutations and high selectivity towards unwanted targets. *The International Journal of Biochemistry & Cell Biology* **2004**, 36, (9), 1787-1799.

18. D’Aquila, R. T.; Schapiro, J. M.; Brun-Vézinet, F.; Clotet, B.; Conway, B.; Demeter, L. M.; Grant, R. M.; Johnson, V. A.; Kuritzkes, D. R.; Loveday, C.; Shafer, R. W.; Richman, D. D., Drug Resistance Mutations in HIV-1. *Topics in HIV Medicine* **2002**, 10, (5), 21-25.

19. Specker, E.; Böttcher, J.; Lilie, H.; Heine, A.; Schoop, A.; Müller, G.; Griebenow, N.; Klebe, G., An Old Target Revisited: Two New Privileged Skeletons and an Unexpected Binding Mode For HIV-Protease Inhibitors. *Angew. Chem. Int. Ed.* **2005**, 44, (20), 3140-3144.

20. Specker, E.; Böttcher, J.; Brass, S.; Heine, A.; Lilie, H.; Schoop, A.; Müller, G.; Griebenow, N.; Klebe, G., Unexpected Novel Binding Mode of Pyrrolidine-Based Aspartyl Protease Inhibitors: Design, Synthesis and Crystal Structure in Complex with HIV Protease. *ChemMedChem* **2006**, 1, (1), 106-117.

21. Czodrowski, P.; Sotriffer, C. A.; Klebe, G., Atypical Protonation States in the Active Site of HIV-1 Protease: A Computational Study. *Journal of Chemical Information and Modeling* **2007**, 47, (4), 1590-1598.

22. Rocha Gonsalves, A. M. d. A.; Serra, M. E. S.; Murtinho, D.; Silva, V. F.; Matos Beja, A.; Paixão, J. A.; Ramos Silva, M.; Alte da Veiga, L., Pyrrolidine-based amino alcohols: novel ligands for the enantioselective alkylation of benzaldehyde. *J. Mol. Catal. A: Chem.* **2003**, 195, (1-2), 1-9.

23. Rodriguez Sarmiento, R. M.; Wirz, B.; Iding, H., Chemoenzymatic preparation of non-racemic *N*-Boc-pyrrolidine-3,4-dicarboxylic acid 3-ethyl esters and their 4-hydroxymethyl derivatives. *Tetrahedron: Asymmetry* **2003**, 14, (11), 1547-1551.

24. DeLano, W. L., DeLano Scientific, San Carlos, CA, (USA). **2002**.

25. Heaslet, H.; Rosenfeld, R.; Giffin, M.; Lin, Y.-C.; Tam, K.; Torbett, B. E.; Elder, J. H.; McRee, D. E.; Stout, C. D., Conformational flexibility in the flap domains of ligand-free HIV protease. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2007**, 63, (Part 8), 866-875.

26. Brynda, J.; Řezáčová, P.; Fábry, M.; Hořejší, M.; Štouračová, R.; Souček, M.; Hradílek, M.; Konvalinka, J.; Sedláček, J., Inhibitor binding at the protein interface in crystals of a HIV-1 protease complex. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2004**, 60, 1943-1948.

27. Kovalevsky, A. Y.; Liu, F.; Leshchenko, S.; Ghosh, A. K.; Louis, J. M.; Harrison, R. W.; Weber, I. T., Ultra-high Resolution Crystal Structure of HIV-1 Protease Mutant Reveals Two Binding Sites for Clinical Inhibitor TMC114. *J. Mol. Biol.* **2006**, 363, (1), 161-173.

28. Cígler, P.; Kozísek, M.; Rezácová, P.; Brynda, J.; Otwinowski, Z.; Pokorná, J.; Plesek, J.; Grüner, B.; Dolecková-Maresová, L.; Mása, M.; Sedlácek, J.; Bodem, J.; Kräusslich, H.-G.; Král, V.; Konvalinka, J., From nonpeptide toward noncarbon protease inhibitors: Metallacarboranes as specific and potent inhibitors of HIV protease. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, (43), 15394-15399.

29. Logsdon, B. C.; Vickrey, J. F.; Martin, P.; Proteasa, G.; Koepke, J. I.; Terlecky, S. R.; Wawrzak, Z.; Winters, M. A.; Merigan, T. C.; Kovari, L. C., Crystal Structures of a Multidrug-Resistant Human Immunodeficiency Virus Type 1 Protease Reveal an Expanded Active-Site Cavity. *J. Virol.* **2004**, 78, (6), 3123-3132.

30. Martin, P.; Vickrey, J. F.; Proteasa, G.; Jimenez, Y. L.; Wawrzak, Z.; Winters, M. A.; Merigan, T. C.; Kovari, L. C., "Wide-Open" 1.3 Å Structure of a Multidrug-Resistant HIV-1 Protease as a Drug Target. *Structure* **2005**, 13, (12), 1887-1895.

31. Taylor, A.; Brown, D. P.; Kadam, S.; Maus, M.; Kohlbrenner, W. E.; Weigl, D.; Turon, M. C.; Katz, L., High-level expression and purification of mature HIV-1 protease in *Escherichia coli* under control of the *araBAD* promoter. *Appl. Microbiol. Biotechnol.* **1992**, 37, (2), 205-210.

32. Toth, M. V.; Marshall, G. R., A simple, continuous fluorometric assay for HIV protease. *Int. J. Pept. Protein Res.* **1990**, 36, (6), 544-550.

33. Yung-Chi, C.; Prusoff, W. H., Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **1973**, 22, (23), 3099-3108.

34. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode. In *Methods Enzymol.*, Carter Jr., C. W., Ed. Academic Press: 1997; Vol. 276, pp 307-326.

35. Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2004**, 60, (3), 432-438.

36. Sheldrick, G. M.; Schneider, T. R., *SHELXL*: High-resolution refinement. In *Methods Enzymol.*, Charles, W. C. J.Robert, M. S., Eds. Academic Press: 1997; Vol. 277, pp 319-343.

37. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2004**, 60, (12 Part 1), 2126-2132.

3. Structure-Guided Design of C_2-Symmetric HIV-1 Protease Inhibitors Based on a Pyrrolidine Scaffold *

3.1 Introduction

Infection with the *human immunodeficiency virus* (HIV) inevitably leads to the development of the *acquired immune deficiency syndrome* (AIDS). The WHO estimates that currently 40 million people are infected worldwide and that this number will increase continuously.[1] In the last 20 years, chemotherapy against HIV infection has been tackled by modern drug discovery and development. Several stages in the viral replication have been targeted to reduce the viral load thus delaying the progression to AIDS. However, an entire remedy of the infection or a vaccination is still unaccomplished goal.[2] One of the most successful approaches in antiviral therapy up to date is the *highly active antiretroviral therapy* (HAART) which combines inhibitors of the viral reverse transcriptase and the viral protease. This viral protease, the HIV protease, belongs to the class of aspartic proteases and cleaves precursor polypeptides into functional viral proteins, which are essential for the infectiousness of the virus particles. The HIV protease is a C_2-symmetric homodimer with each monomer consisting of 99 residues. Two glycine-rich flaps are located above the large active site, which binds six to seven amino acids of the nine different cleavage sites within the *gag* and *pol* polypeptides. Because inhibition of this enzyme leads to immature virions, HIV protease has become a target of several marketed drugs. Up to now, nine inhibitors against the HIV protease have been approved by the FDA,[3-6] but meanwhile the virus has shown a very high degree of adaptability. The high mutation rate of the virus is caused by the error-prone viral reverse transcriptase and its fast replication rate. This leads, especially under the additional selection pressure of HAART, to the development of resistant virus variants.[7]

* Taken from original publication, Andreas Blum, Jark Böttcher, Andreas Heine, Gerhard Klebe, Wibke E. Diederich. Structure-Guided Design of C_2-Symmetric HIV-1 Protease Inhibitors Based on a Pyrrolidine Scaffold. *J. Med. Chem.* **2008**, 51, (7), 2078-2087.

Because a similar mode of action of all approved inhibitors has resulted in pronounced cross-resistance, the need of a steadfast and continuous search for new inhibitors is evident. An increased structural diversity of inhibitor scaffolds could be a possible strategy to at least diminish the accelerated development of multi-drug resistant variants.

Most of the approved inhibitors are transition state mimetics addressing the catalytic dyad (D25A/D25B) via a hydroxyl group. As a different approach, cyclic amines have been identified as non-peptidic inhibitors for the aspartic proteases Renin,[8, 9] β-Secretase,[10] and Plasmepsins.[11, 12] Recently, a pyrrolidine-based inhibitor for HIV protease has been developed.[13, 14] The X-ray structure of the protein inhibitor complex revealed the endocyclic amino function binding to the catalytic dyad. Electrostatic calculations suggest that the amino functionality should be protonated and both aspartic acid residues of the catalytic dyad therefore deprotonated, thus leading to strong hydrogen bonds as well as electrostatic interactions of the resulting cyclic ammonium function to the catalytic dyad.[15] The sulfonamide side-chain of the inhibitor addresses the flap region by one of its oxygen atoms through hydrogen bonding to the backbone NH of I50A (Figure 1). A superposition of this structure with other crystal structures of HIV protease complexes indicated that the distance between the sulfonyl group and the endocyclic nitrogen is too large for appropriate flap-interactions. The carboxamide part of inhibitor **1** showed no directed polar interactions to the enzyme and was therefore not considered for further inhibitor design.

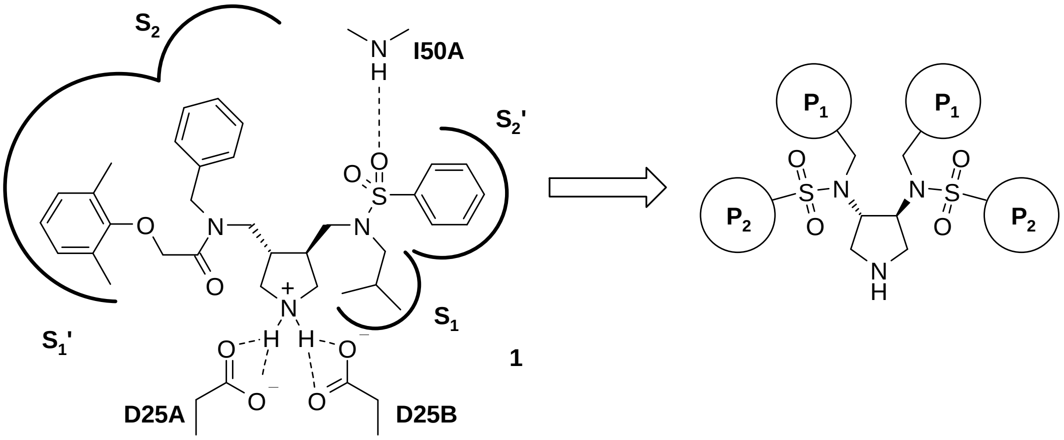
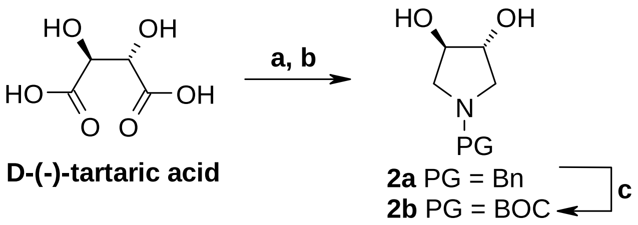


Figure 1: Design of C_2-symmetric inhibitors starting from the cocrystal structure of **1**.

Guided by the described polar interactions of inhibitor **1** to the enzyme, a basic pharmacophore model was derived: As core structure, the 3*S*,4*S*-disubstituted pyrrolidine, addressing the catalytic dyad via its secondary amino group, was retained. However, as indicated by the X-ray structure of **1**, the exocyclic methylene groups were removed in order to improve the interaction pattern of the sulfonyl groups with the flap region of the enzyme. To exploit the C_2-symmetry of the enzyme, exclusively symmetric inhibitors were synthesized. To mimic the subpocket occupation of **1**, arylsulfonamides were introduced to occupy the S_2- and S_2’-subpockets. In addition, a hydrophobic moiety was attached to each of the sulfonamide nitrogen atoms designed to address the S_1- and S_1’-subsites of the enzyme (Figure 1).

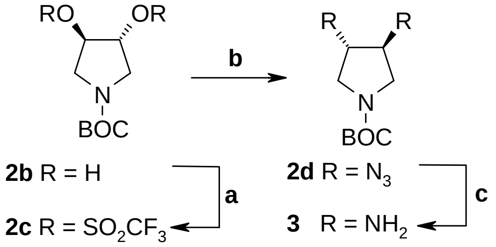
3.2 Chemistry

3,4-Difunctionalized pyrrolidines are accessible in enantiopure form commencing with commercially available tartaric acids.[16, 17] D-(-)-Tartaric acid was condensed with benzylamine to furnish the corresponding cyclic imide, which was further reduced with $LiAlH_4$ to yield the benzyl-protected pyrrolidine-3,4-diol **2a**[18] (Scheme 1).



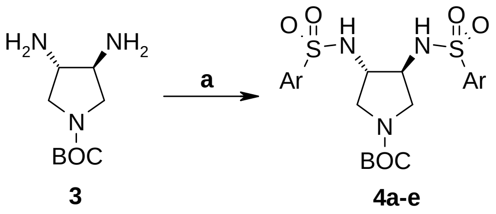
Scheme 1: (a) $BnNH_2$, xylene, Dean-Stark (81%); (b) LAH, THF, reflux (69%); (c) H_2, Pd/C, BOC_2O, MeOH (74%).

The synthesis of the corresponding benzyl protected pyrrolidine-3,4-diamine, accessible via a Mitsunobu reaction of diol **2a** with HN_3 and subsequent reduction of the resulting diazide, has already been described.[19] However, following this synthetic route, the removal of the benzyl-protecting group by catalytic hydrogenation in the last step of the synthesis to give rise to the final inhibitors remained unsuccessful, even after applying higher temperature and hydrogen pressure as well as different hydrogen sources. Furthermore, we intended to avoid the usage of the very toxic hydrazoic acid in the Mitsunobu reaction. Therefore, a change to the easily removable BOC-protecting group was necessary. This exchange was carried out at the diol stage by hydrogenation of **2a** in presence of BOC_2O,[20] yielding the BOC-protected 3,4-diol **2b** (Scheme 1). Although an activation of the 3,4-pyrrolidine diol for the purpose of nucleophilic displacement reactions by conversion into its corresponding disulfonates has been described,[21, 22] in our hands, the respective mesylate did not undergo substitution with NaN_3 in DMSO even at 120°C. Nevertheless, utilization of the related, but significantly more reactive bis-triflate **2c** followed by its nucleophilic substitution with NaN_3 in DMPU at RT now yielded the 3,4-diazide **2d**, which was subsequently reduced to the corresponding pyrrolidine diamine **3** by catalytic hydrogenation (Scheme 2) in excellent overall yield.

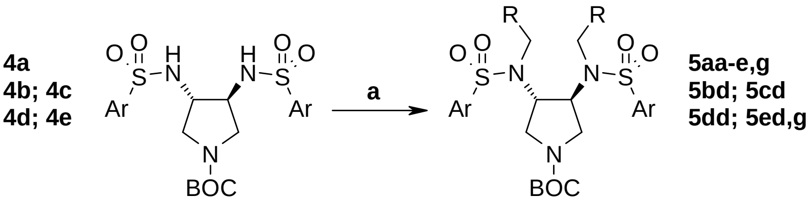


Scheme 2: (a) Tf_2O, pyridine, CH_2Cl_2 -78°C→-10°C; (b) NaN_3, DMPU, RT; (c) H_2, Pd/C, hexane/EtOAc (86% over three steps).

Diamine **3** was then condensed with suitable sulfonyl chlorides to the corresponding sulfonamides **4** (Scheme 3), which were further alkylated with different allyl and benzyl bromides in the presence of K_2CO_3 in acetonitrile (Scheme 4) thus yielding the BOC-protected inhibitor precursors of type **5**.

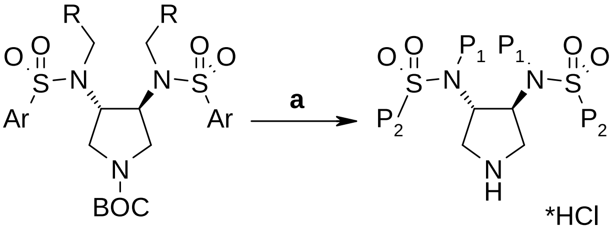


Scheme 3: (a) $ArSO_2Cl$, Et_3N, CH_2Cl_2: **4a** Ar: Ph (86%); **4b** Ar: *o*-Me-Ph (87%); **4c** Ar: *o*-Cl-Ph (94%); **4d** Ar: *p*-NO_2-Ph (69%); **4e** Ar: *p*-CN-Ph (71%).



Scheme 4: (a) RCH_2Br, K_2CO_3, CH_3CN, reflux: **4a** Ar: Ph (**5aa** R: $CH{=}CH_2$ 72%; **5ab** R: $CH(CH_3){=}CH_2$ 73%; **5ac** R: $CH{=}C(CH_3)_2$ 68%; **5ad** R: Ph 75%; **5ae** R: *p*-Br-Ph 76%; **5af** R: *p*-I-Ph 76%; **5ag** R: *p*-CF_3-Ph 49%); **4b** Ar: *o*-Me-Ph (**5bd** R: Ph 78%); **4c** Ar: *o*-Cl-Ph (**5cd** R: Ph 65%); **4d** Ar: *p*-NO_2-Ph (**5dd** R: Ph 96%); **4e** Ar: *p*-CN-Ph (**5ed** R: Ph 67%; **5eg** R: *p*-CF_3-Ph 45%).

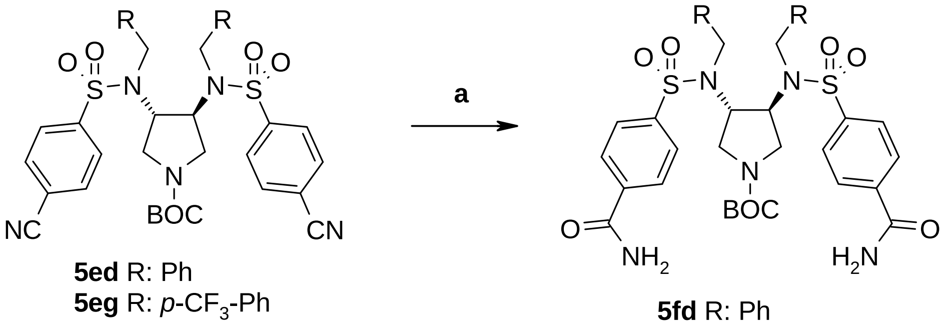
The final deprotection was carried out under non-aqueous acidic conditions using 2 M HCl in Et_2O furnishing the inhibitors **6** as their hydrochlorides in high yields (Scheme 5, yields are given in Table 1). In compounds with double lettering codes, the first letter refers to the sulfonyl-substituents (P_2/P_2’), the second letter to the *N*-alkyl moieties (P_1/P_1’).



5 **6**

Scheme 5: (a) 2 M HCl in Et_2O, RT. Details are given in Table 1.

The carboxamido-substituted compounds **5fd**, and **5fg** are accessible from the corresponding cyano-substituted derivatives **5ed**, and **5eg** by mild hydrolysis applying 30% H_2O_2 in DMSO (Scheme 6).[23]

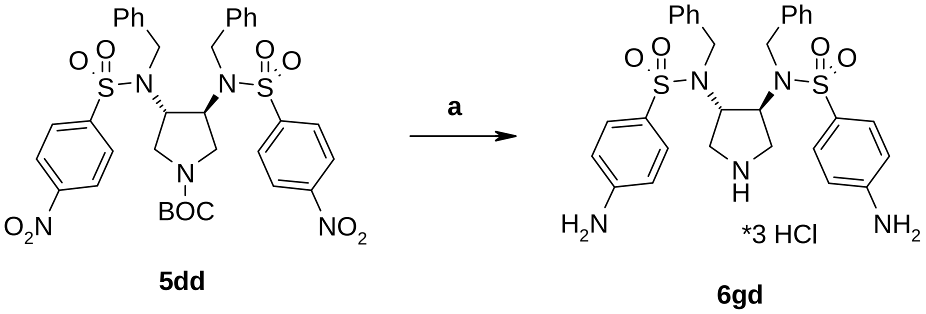


5ed R: Ph
5eg R: *p*-CF_3-Ph

5fd R: Ph
5fg R: *p*-CF_3-Ph

Scheme 6: (a) 30% aq H_2O_2, K_2CO_3, DMSO, 0°C→RT (**5fd** R: Ph 92%; **5fg** R: *p*-CF_3-Ph 71%).

The amino-substituted inhibitor **6gd** was obtained from the corresponding nitro-substituted precursor **5dd** by reduction with $SnCl_2$ in EtOAc at elevated temperature.[24] Under these reaction conditions, the BOC-protecting group was removed simultaneously (Scheme 7).



Scheme 7: (a) $SnCl_2 \cdot 2\,H_2O$, EtOAc, 2h reflux, then 2 M HCl in Et_2O 79%.

The synthesis of inhibitors of type **6** was achieved in 7 or 8 steps, respectively, starting from D(-)-tartaric acid in an overall yield ranging from 7 to 21% with an average yield of more than 70% for each step.

3.3 Results and Discussion

To gain a first insight into the structure-activity-relationship (SAR), the benzene sulfonamide group in **4a** was alkylated with three differently alkyl-substituted alkenyl moieties and a benzyl group using allyl bromide (to **5aa**), 2-methyl-allyl bromide (to **5ab**), 3,3-dimethyl-allyl bromide (to **5ac**), and benzyl bromide (to **5ad**). After acidic deprotection of the BOC-group, inhibitors **6aa-6ad** were obtained. All four bis-substituted sulfonamides showed affinity in the micromolar range against HIV protease, from which those inhibitors with the largest moieties (**6ac** and **6ad**) exhibited the highest affinity (Table 1).

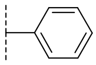
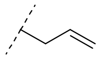
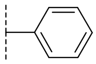
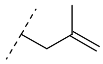
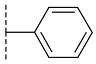
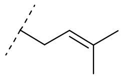
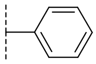
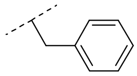
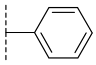
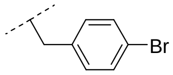
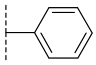
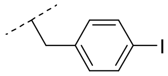
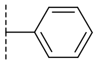
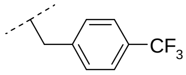
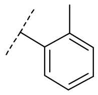
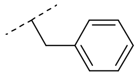
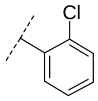
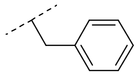
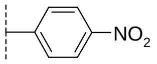
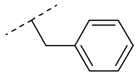
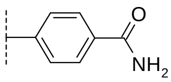
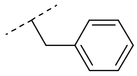
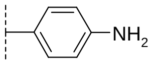
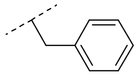
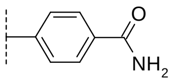
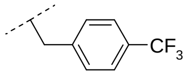
	P_2/P_2’	P_1/P_1’	K_i [μM]	Yield (from **5**)
6aa			12.3	76 (**5aa**)
6ab			74.7	71 (**5ab**)
6ac			1.57	78 (**5ac**)
6ad			2.15	92 (**5ad**)
6ae			0.46	72 (**5ae**)
6af			0.39	81 (**5af**)
6ag			0.80	71 (**5ag**)
6bd			0.67	81 (**5bd**)
6cd			0.77	78 (**5cd**)
6dd			1.72	79 (**5dd**)
6fd			0.26	81 (**5fd**)
6gd			0.27	79 (**5dd**)
6fg			0.07	80 (**5fg**)

Table 1: K_i-values of the inhibitors **6** towards the HIV protease and yield of the terminal deprotection step.

The X-ray structure of **6ad** in complex with the protease was determined with a resolution of 1.55 Å and surprisingly revealed an asymmetric binding mode (Figure 2). Similar to the complex structure of pyrrolidine-based inhibitor **1**, the protonated cyclic amino nitrogen is found at its pivotal position forming a hydrogen bond network to the catalytic aspartic acid residues D25A (2.6 Å, 3.0 Å) and D25B (2.9 Å, 3.0 Å).

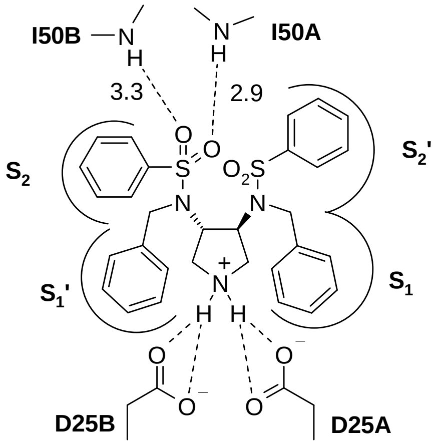


Figure 2: Schematic representation of the binding mode of **6ad** observed in the crystal structure in complex with the HIV protease; bond lengths to I50A and I50B of the flap-region are given in Å.

Instead of a structural water molecule usually present in peptidomimetic or substrate-like complexes, one sulfonyl group establishes two hydrogen bonds to the backbone-NHs of I50A (2.9 Å) and I50B (3.3 Å), each with one sulfonyl oxygen atom. The second sulfonyl group, however, does not form any polar interactions (Figure 3).

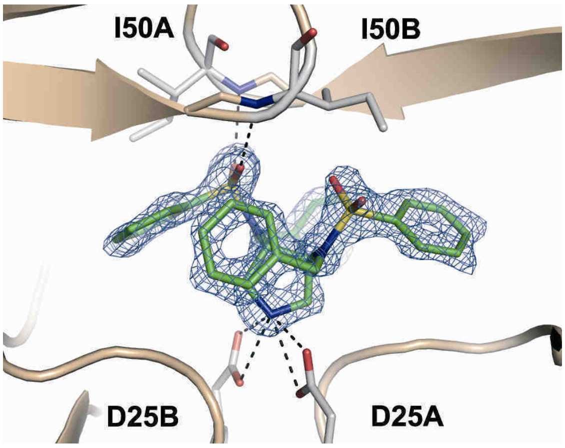


Figure 3: Crystal structure of **6ad** (green, color-coded by atom type) in complex with HIV protease. The protein backbone trace is schematically illustrated in wheat, the catalytic D25A and D25B as well as I50A and I50B of the flaps are displayed in grey color-coded by atom type. The F_o-F_c density for the ligand is displayed at a σ level of 3.0 as blue mesh.

The benzyl substituents occupy the S_1/S_1'-subpockets establishing Van der Waals (VdW)-contacts with L23A, I50B, V82A, I84A (S_1); R8B, L23B, I50A, P81B, V82B, I84B (S_1'). The phenyl moieties reside in the S_2/S_2'-subsites forming VdW-contacts with A28B, V32B, I50A, I84B (S_2); A28A, D30A, V32A, I47A, I50B (S_2'). Additional VdW-contacts are found to G27, G48, and G49 of each subunit. The deviation from a C_2-symmetrical occupation of the subpockets can be analyzed by superpositioning the observed geometry of the inhibitor with its geometry resulting from a rotation around the protein's C_2-axis. The average distance of the corresponding ring atoms is 1.7 Å in the S_1- and 2.1 Å in the S_2-pocket.

This crystal structure enabled us to rationally design a second series of most likely more potent inhibitors. Consequently, **6ad** was selected as lead structure for further optimization. An in-depth analysis of the crystal structure revealed three very promising symmetric substitution patterns for further lead optimization (Figure 4):

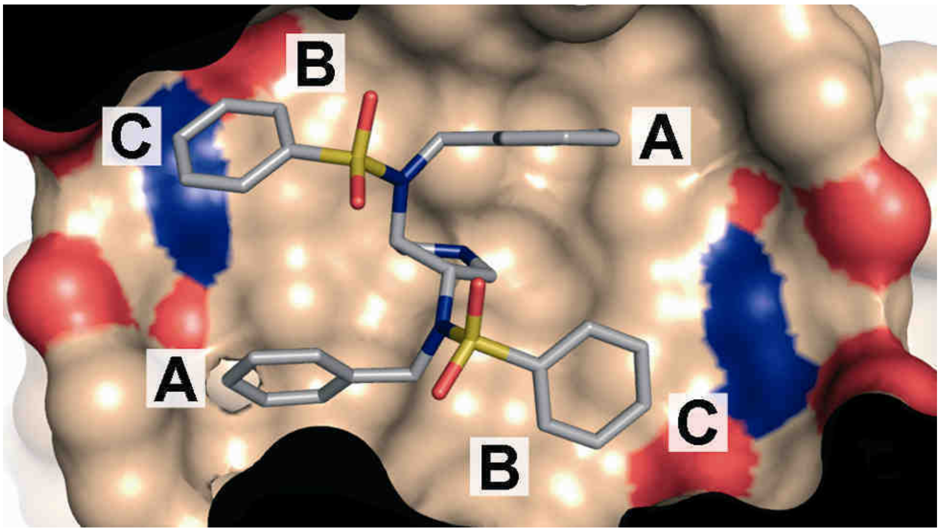


Figure 4: Overview of the optimization strategies of **6ad** by attachment of additional substituents; protein shown in wheat surface representation, D29 and D30 surface patches are shown in color-coded by atom type; (A) hydrophobic P_1/P_1'-elongation; (B) small hydrophobic P_2/P_2'-*ortho*-substitution; (C) polar P_2/P_2'-*para*-substitution to address D29 and D30.

(A) Elongation of the P_1/P_1’-benzyl moieties with hydrophobic substituents in *para*-position to address unoccupied space in the S_1/S_1’-pocket.

(B) *ortho*-substitution at the P_2/P_2’-phenyl ring systems to improve the shape match between the binding pocket surface and the ligand. For this purpose small hydrophobic substituents were selected in order to fill remaining space flanked by V32B and I84B in the S_2-pocket as well as V32A and I84A in the S_2’-pocket.

(C) *para*-substitution at the P_2/P_2’-phenyl moieties with substituents capable of forming hydrogen bonds to D29A and D30A in the S_2’-pocket and D29B and D30B in the S_2-pocket.

All three strategies were pursued, and the resulting inhibitors **6ae-6ag** and **6bd-6gd** were subsequently analyzed by kinetic measurements with respect to their affinity against the target enzyme. The cocrystal structures of selected representatives, at least one of each modification type, in complex with HIV protease were determined. This approach allows the detailed analysis of additional interactions formed by the introduced substituents and provides an extensive understanding of the principles accounting for the SAR. Subsequently, this knowledge should facilitate the selection of the most promising combination of substituents thus leading to an even further improved inhibitor.

(A) Elongation of the P_1-benzyl ring in *para*-position with hydrophobic residues was easily achieved by alkylation of benzene sulfonamide **4a** with appropriately substituted benzyl bromides, yielding after deprotection the corresponding bromo- (**6ae**), iodo- (**6af**), and trifluoromethyl- (**6ag**) substituted inhibitors. These inhibitors indeed showed improved affinity by a factor of five compared to **6ad** for the halide substituted ones (**6ae** 0.46 µM; **6af** 0.39 µM). In case of the trifluoromethyl-substituted inhibitor **6ag**, a factor of three (0.80 µM) is achieved. The crystal structure of **6af** in complex with HIV protease was determined with a resolution of 1.41 Å and revealed a binding mode resembling that of the unsubstituted pyrrolidine **6ad**. A similar hydrogen-bond network of the core structure can be observed (Table 2).

	6ad	**6af**	**6cd**	**6fd**	**6gd**	**6fg**
D25A $O_{\delta 1}$	3.0	2.9	3.0	3.0	3.0	2.9
D25A $O_{\delta 2}$	2.6	2.7	2.5	2.7	2.6	2.7
D25B $O_{\delta 1}$	3.0	2.8	2.8	2.8	2.8	2.8
D25B $O_{\delta 2}$	2.9	3.1	2.9	2.8	2.9	3.0
I50A N	2.9	3.0	3.1	3.0	3.1	2.9
I50B N	3.3	3.1	3.6	3.2	3.7	3.1

Table 2: Conserved hydrogen bonds of the pyrrolidine-nitrogen atom to the catalytic dyad (D25A/D25B) and the sulfonyl oxygen atoms to the flap (I50A/I50B) observed in the determined cocrystal structures; values given in Å.

The root-mean-square deviation (rmsd) between the C_α atoms of the complexes of **6ad** and **6af** constitutes 0.15 Å, resembling the overall similar binding mode. This similarity is also reflected by an rmsd of 0.45 Å from the respective **6ad** substructure in **6af**. The gain in affinity is presumably due to additional VdW interactions in the S_1-subsite with P81A, G48B, and G49B and in the S_1'-subsite with R8B. Obviously, the position and orientation of the P_2/P_2'-moiety is not affected by the additional substituent.

(B) The extension of the P_2/P_2'-phenyl ring in *ortho*-position with small hydrophobic substituents was successfully accomplished by synthesis of the *ortho*-methyl and *ortho*-chloro substituted benzene sulfonamides **4b** and **4c**, respectively, which were subsequently alkylated with benzyl bromide yielding after deprotection the methyl- (**6bd**) and chloro- (**6cd**) substituted inhibitors. Compared to the lead structure **6ad**, a threefold increase in affinity was observed (**6bd** 0.67 µM; **6cd** 0.77 µM). The 2.30 Å-resolved crystal structure of the chloro-derivative **6cd** revealed a similar binding mode compared to that of **6ad**. The protein structure remains unaffected (C_α rmsd to **6ad** is 0.14 Å), however, the binding mode of **6cd** slightly deviates from the expected conformation (rmsd to **6ad** substructure is 0.67 Å). The observed polar interactions are in agreement with the structure of **6ad** (Table 2), furthermore the *ortho*-chloro-substitution is as intended in additional VdW-contacts of the chlorine atom to V32B and I84B in the S_2'-pocket. Moreover, the *ortho*-substituent induces an unanticipated change in the orientation of the phenyl moiety in the S_2-pocket. There, the chloro-substituent points in the opposite direction toward the S_1-pocket now interacting with G48B and G49B, hence retaining the orientation of **6ad**.

(C) For the decoration of the P_2/P_2'-phenyl moieties in *para*-position with polar groups, the 4-nitrophenyl sulfonamide **4d** as well as the 4-cyanophenyl sulfonamide **4e** were synthesized and subsequently alkylated with benzyl bromide. The alkylated nitro compound **5dd** was on the one hand directly deprotected to yield inhibitor **6dd**. On the other hand, the nitro group in **5dd** was first reduced to the corresponding amino function concurrently removing the protecting group now giving rise to the amino-substituted inhibitor **6gd**. The alkylated cyano compound **5ed** was converted into the carboxamide **5fd** and subsequently deprotected yielding inhibitor **6fd**. Compared with the unsubstituted pyrrolidine **6ad**, the nitro-substituted inhibitor **6dd** only showed a comparable affinity (**6dd** 1.72 µM), whereas those inhibitors with hydrogen bond donor substituents revealed an eightfold increase in affinity (**6fd** 0.26 µM; **6gd** 0.27 µM). To elucidate the hydrogen bond network established by the polar substituents, both inhibitors were crystallized in complex with HIV protease and the resulting crystal structures were determined (resolution: **6fd** 1.50 Å; **6gd** 1.78 Å). Both crystal structures show a high similarity to that of **6ad** with respect to the C_α atoms of the protein (C_α rmsd **6fd** 0.25 Å; **6gd** 0.12 Å) as well as the binding-mode of the ligand (rmsd of **6ad** substructure **6fd** 0.54 Å; **6gd** 0.36 Å). The polar interactions of the central scaffold are similar to those of **6ad** (Table 2). Due to the asymmetric binding mode, the interaction patterns in the S_2-and S_2'-subpockets of each ligand differ from each other (Figure 5, Figure 6): The *para*-amino derivative **6gd** establishes two hydrogen bonds in the S_2-pocket, one directly to the backbone carbonyl of D30B (3.2 Å) and one to the sidechain of D30B bridged by a water molecule (2.8 Å, 2.7 Å). In the S_2'-pocket, also two hydrogen bonds are formed, however, now to the side-chains of D30A (2.7 Å) and D29A again mediated by an interstitial water molecule (3.2 Å, 2.9 Å).

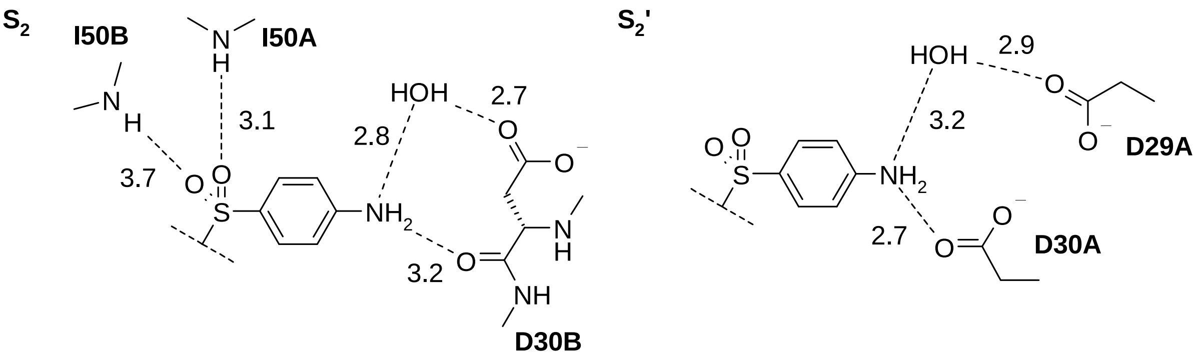


Figure 5: Schematic representation of the hydrogen bond network observed in the crystal structure of **6gd**; bond lengths are given in Å.

The same number of hydrogen bonds can be observed for the *para*-carboxamido substituted inhibitor **6fd**: In the S_2-pocket the backbone NH of D30B is addressed via the carboxamide oxygen (2.8 Å), whereas the amide-nitrogen establishes a hydrogen bond to the side-chain of D30B (2.9 Å). In the S_2'-pocket a water molecule mediates the interaction between the ligand amido nitrogen and the D29A side-chain (3.0 Å, 2.7 Å). The carbonyl function of the ligand accepts a hydrogen bond from the backbone NH of D30A (3.0 Å). The position of the P_1/P_1'-residues in both ligand structures is not affected by these additional interactions in the S_2- and S_2'-pockets.

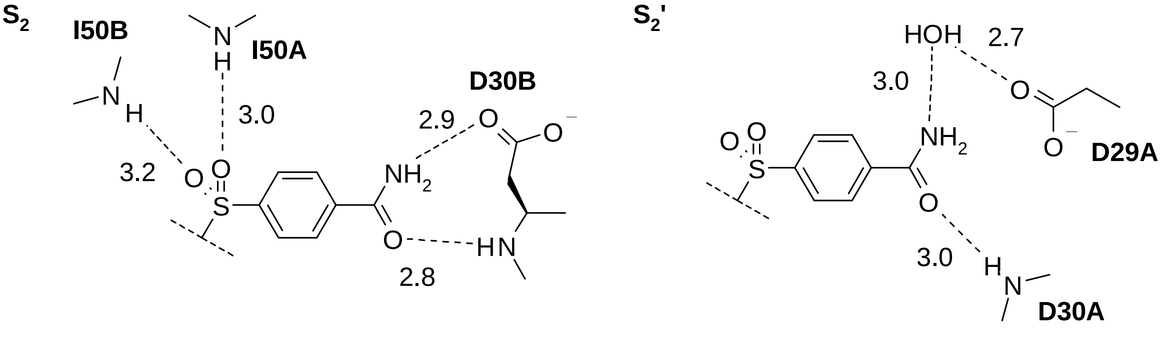


Figure 6: Schematic representation of the hydrogen bond network observed in the crystal structure of **6fd**; bond lengths are given in Å.

All crystal structures analyzed at this stage of the project are very similar; the rmsd deviation between the C_α atoms of any two complexes is less than 0.28 Å, and all inhibitors also show a very high degree of similarity in their binding modes (Table 3).

	6ad	**6af**	**6cd**	**6fd**	**6gd**	**6fg**	**6ad**	**6af**	**6cd**	**6fd**	**6gd**	**6fg**
	Protein C_α						**6ad** substructure					
6ad	-	0.15	0.14	0.25	0.12	0.19	-	0.45	0.67	0.54	0.36	0.69
6af	0.15	-	0.17	0.23	0.19	0.18	0.45	-	0.56	0.66	0.57	0.61
6cd	0.14	0.17	-	0.28	0.15	0.20	0.67	0.56	-	0.85	0.72	0.81
6fd	0.25	0.23	0.28	-	0.23	0.23	0.54	0.66	0.85	-	0.28	0.32
6gd	0.12	0.19	0.15	0.23	-	0.20	0.36	0.57	0.72	0.28	-	0.48
6fg	0.19	0.18	0.20	0.23	0.20	-	0.69	0.61	0.81	0.32	0.48	-

Table 3: Calculated rmsd after alignment of the protein C_α-atoms; for the ligands the respective **6ad** substructure within each ligand is compared after C_α-alignment of the protein structure; values given in Å.

The hydrogen-bond network of the lead structure **6ad** is conserved in all structures (Table 2). The rmsd between the respective substructure and **6ad** is less than 0.69 Å. For the comparison of any two ligands, the rmsd is less than 0.85 Å. The largest differences are observed between the *ortho*-chloro-substituted ligand (**6cd**) and the *para*-amino (**6gd**) or the *para*-carboxamido-substituted ligand (**6fd**) with 0.72 Å and 0.85 Å respectively. This reflects the different binding mode of **6cd** (Table 3). This different binding mode makes an *ortho*-substitution at the P_2-residue less attractive for further combination with other modifications (Figure 7), so only the two remaining substitution strategies were combined. As substituents for the combined inhibitor, a trifluoromethyl group as P_1/P_1'- and a carboxamido moiety as P_2/P_2'-substituent were selected.

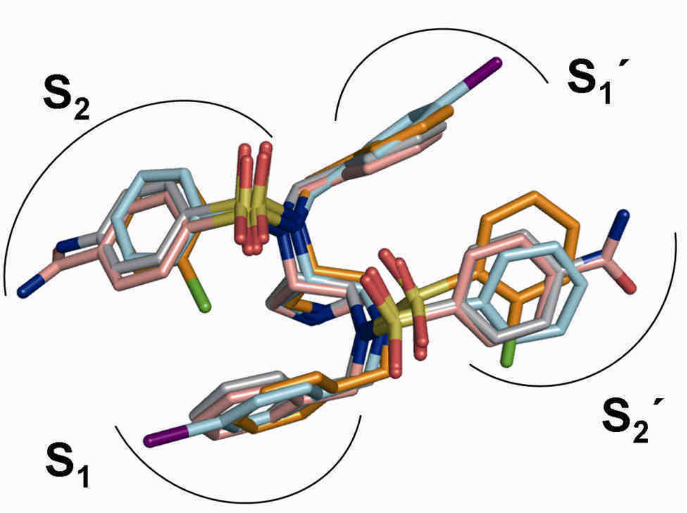


Figure 7: Superposition of the ligand conformations observed in the corresponding crystal structures aligned by C_{α}-fit of the proteins; ligands color-coded by atom type **6af** (cyan), **6cd** (orange), **6fd** (salmon), **6gd** (grey).

The combined inhibitor **6fg** was prepared in similar manner as already described for carboxamide inhibitor **6fd** starting from sulfonamide **4e**, followed by alkylation to **5eg**, and subsequent hydrolysis of the nitriles to the carboxamide **5fg.** After acidic, non-aqueous deprotection, the inhibitor **6fg** was tested for its affinity against HIV protease. With 74 nM the trifluoromethyl and carboxamido-substituted inhibitor **6fg** exhibits the highest affinity within this series. To validate our combination strategy pursued within this project, the crystal structure of **6fg** in complex with HIV protease was determined with a resolution of 1.48 Å (Figure 8).

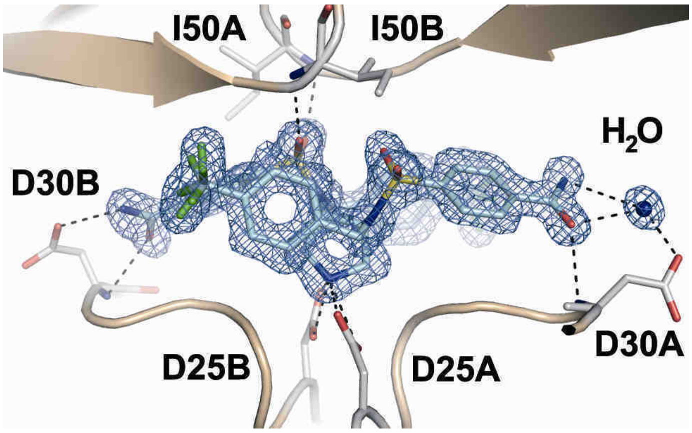


Figure 8: Crystal structure of **6fg** (cyan, color-coded by atom type) in complex with HIV protease. The protein backbone trace is schematically illustrated in wheat, the catalytic D25A/D25B, D30A/D30B, and I50A/I50B of the flaps are displayed in grey color-coded by atom type. The F_o-F_c density for the ligand is displayed at a σ level of 3.0 as blue mesh.

No major differences could be observed comparing the protein structure to the formerly determined complexes (C_α rmsd to **6af** 0.18 Å; to **6fd** 0.23 Å). Our successful combination strategy is also represented in the conserved ligand conformation (rmsd of **6ad** substructure to **6af** 0.61 Å; **6fd** 0.32 Å). The polar interaction pattern of the carboxamido-substituted inhibitor **6fd** is retained. Compared to **6af**, additional VdW-contacts of the trifluoromethyl group to R8A in the S_1-subsite are observed. In the structure of **6fg**, both CF_3 groups show disorder of a rigid rotor. They were refined as double conformations.

3.4 Summary and Conclusion

C_2-symmetric 3,4-disubstituted pyrrolidines have been developed as a new class of HIV protease inhibitors. Starting from the initial lead **6ad**, which showed affinity in the low micromolar range, the activity of this new class of HIV protease inhibitors could be significantly optimized by means of rational structure-based design up to the two digit nanomolar range for the final inhibitor **6fg**. Based on our developed synthetic strategy, the synthesis of the enantiopure key intermediate **3** commencing with D-(-)-tartaric acid is straightforward and high-yielding. Condensation of **3** with appropriately chosen sulfonyl chlorides renders the corresponding sulfonamides. Alkylation of which, followed by further functional group transformations and deprotection gave rise to the desired inhibitors. These are obtained via a 7 or 8 step synthesis with an overall yield ranging from 7 to 21%. Following the outlined synthetic route, the synthesis of a plethora of putative protease inhibitors is readily feasible.

Initial SAR-studies as well as the crystal structure determination resulted in selection of **6ad** as lead compound for further structural optimization. The analysis of the cocrystal structure of **6ad** revealed three possible strategies for optimization via symmetric introduction of substituents to the original P_1/P_1'- and P_2/P_2'-phenyl moieties. From each class of possible modifications, at least one representative was synthesized and thoroughly analyzed by crystal structure determination of the protein-ligand complex. The observed interactions of the core structure are highly conserved throughout this series of inhibitors. These structures provided deeper insights into the protein-ligand interactions and the underlying principles of the SAR. Taking this information into account, the most promising combination of P_1/P_1'- and P_2/P_2'-moieties was selected and the resulting inhibitor **6fg** indeed showed the expected improvement in affinity with a K_i = 74 nM. The cocrystal structure of this inhibitor confirmed the successful application of our optimization strategy. The complete structural characterization of crucial intermediates prevents misdirection by only taking the affinity of the compounds into account. This project clearly shows that the rational design of inhibitors based on the successful cooperation between synthetic medicinal chemistry and structural biology can lead to a highly efficient optimization of lead structures.

3.5 Experimental Section

Kinetic Assay: Inhibition data for HIV protease were determined as follows: IC_{50} values were taken from plots of v_i/v_0 versus inhibitor concentration, in which v_i is the velocity in presence, and v_0 the velocity in the absence of an inhibitor. The fluorogenic substrate Abz-Thr-Ile-Nle-(*p*-NO_2-Phe)-Gln-Arg-NH_2 was purchased from Bachem. Recombinant HIV protease was expressed from *Escherichia coli* and purified as previously described.[25] Enzymatic assays were performed in 172 µL assay buffer (100 mM MES, 300 mM KCl, 5 mM EDTA, 1 mg/mL BSA, pH 5.5) by the addition of substrate dissolved in 4 µL DMSO, distinct inhibitor concentrations dissolved in 4 µL DMSO and 20 µL HIV-1 protease in assay buffer to a final volume of 200 µL (final DMSO concentration 4%). The hydrolysis of the substrate was recorded as the increase in fluorescence intensity (excitation wavelength 337 nm, emission wavelength 410 nm) over a time period of 10 min during which the signal increased linearly with time.[26] The kinetic constants for HIV protease ($K_m = 14.6\ \mu M$) were determined by the method of Lineweaver and Burk with a HIV protease concentration of 2.8 nM with varied substrate concentrations. The active-site concentration was quantified by titrating three different HIV protease concentrations with the strong binding inhibitor Saquinavir ($K_i = 0.3$ nM). K_i values were calculated from the following equation: $K_i = [IC_{50} - (E_t/2)][1 + (S/K_m)]^{-1}$

Crystallization of HIV protease inhibitor complexes: HIV protease inhibitor complexes were crystallized at 18°C in 0.1 M BisTris, pH 6.5, 2.0-3.5 M NaCl and a protein concentration of 7 mg/mL in the space group $P2_12_12$ (crystal data, Table 4 and 5). The crystals were obtained by cocrystallization of the enzyme with inhibitor concentrations ranging from 20 to 100 fold the K_i value. Crystals were further optimized using streak-seeding techniques. For cryoprotection the crystals were briefly soaked in mother liquor containing 25% glycerol.

Data collection, phasing and refinement: The data sets were collected at the synchrotron BESSY II in Berlin/Germany on PSF beamline 14.2 (PDB ID: 2PQZ **6ad**, 2QNP **6af**, 2PWC

6gd, 2QNN **6fg**), the EMBL/DESY in Hamburg/Germany on beamline X13, (pdb ID: 2PWR **6fd**) both equipped with a MAR-CCD detector and (pdb ID: 2QNQ **6cd**) on a Rigaku R-AXIS IV image plate detector using Cu K_α radiation from an in-house Rigaku rotating anode. Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000.[27] The structures were determined by the molecular replacement method with Phaser,[28] one monomer of the 1.5 Å structure of the HIV-1 protease in complex with a pyrrolidine based inhibitor (PDB ID: 1XL2) and consecutively a monomer of the determined structure (PDB ID: 2PQZ **6ad**) was used as the search model. The structure (PDB ID: 2QNP **6af**) was determined by single isomorphous replacement with anomalous scattering method (SIRAS) using the HKL2MAP graphics user interface,[29] the dataset of the unsubstituted pyrrolidine was used as native data. The positions of the iodine atoms were identified using SHELXD[30] and the phase problem was solved with SHELXE.[31] The resulting phases were applied and after density modification with DM,[32] ARPwARP[33] as implemented in CCP4[34] was used for automated model building.

Refinement was continued with CNS[35] and SHELXL-97,[36] for each refinement step at least 10 cycles of conjugate minimization were performed, with restraints on bond distances, angles and B-values. Intermittent cycles of model building were done with the program COOT.[37] The coordinates have been deposited in the PDB (http://www.rcsb.org/pdb/) with access codes: **6ad** 2PQZ; **6af** 2QNP; **6cd** 2QNQ; **6fd** 2PWR; **6gd** 2PWC; **6fg** 2QNN.

	6ad 2PQZ	**6af** 2QNP	**6cd** 2QNQ
Resolution (Å)	25-1.55	25-1.41	30-2.30
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
cell dimensions (Å)	a = 57.53 b = 86.00 c = 46.72	a = 57.36 b = 85.90 c = 46.65	a = 57.57 b = 85.90 c = 46.54
highest resolution shell (Å)	1.58-1.55	1.43-1.41	2.35-2.30
no. of measured reflections	153817	168404	57399
no. of independent reflections	32912	44939	10478
completeness (%)[a]	96.3 [87.3]	99.5 [94.1]	97.1 [93.9]
I/σ[a]	13.4 [3.0]	21.9 [2.83]	16.1 [5.27]
R_{sym} (%)[a]	8.6 [33.5]	5.0 [33.4]	11.0 [25.2]
refined residues	198	198	198
refined ligand atoms	39	41	41
refined water molecules	187	183	70
refined chlorides	3	2	3
refined glycerols	-	-	-
resolution in refinement (Å)	10-1.55	10-1.41	25-2.30
R_{cryst} (%) $F > 4\ \sigma\ F_{o;}\ F_o$	16.4	17.6	20.5
R_{free} (%) $F > 4\ \sigma\ F_{o;}\ F_o$	20.4	21.3	23.9
mean B-factor ($Å^2$)			
(peptide chain A; B)	15.7; 14.2	14.9; 13.5	23.3;20.5
main-chain ($Å^2$)	12.8; 11.5	12.5; 10.9	22.6; 20.1
side-chains ($Å^2$)	18.8; 17.2	17.7; 15.4	24.2; 21.2
ligand ($Å^2$)	24.0	12.7[b]	38.0
water ($Å^2$)	28.1	24.8	23.4
Cl ions ($Å^2$)	17.1	16.9	22.1
Glycerol ($Å^2$)	-	-	-
Ramachandran plot			
most favorate geometry (%)	96.8	95.6	97.5
additionally allowed (%)	3.2	4.4	2.5
generously allowed (%)	-	-	-
disallowed (%)	-	-	-

Table 4: X-ray data processing and refinement for the HIV protease complexes of derivatives **6ad, 6af** and **6cd**. (a) Values in brackets refer to the highest resolution shell. (b) In case of **6af** the ligand occupancy was refined to 63% with respect to the protein.

	6fd 2PWR	**6gd** 2PWC	**6fg** 2QNN
Resolution (Å)	25-1.50	25-1.78	25-1.48
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
cell dimensions (Å)	a = 56.77	a = 57.28	a = 57.38
	b = 84.80	b = 85.78	b = 86.18
	c = 46.06	c = 46.45	c = 46.29
highest resolution shell (Å)	1.53-1.5	1.81-1.78	1.51-1.48
no. of measured reflections	141076	87548	137979
no. of independent reflections	33531	22211	38945
completeness (%)[a]	92.1 [93.7]	98.2 [99.5]	99.5 [96.4]
I/σ[a]	34.2 [6.6]	13.5 [2.9]	23.5 [3.71]
R_{sym} (%)[a]	3.9 [23.0]	9.7 [46.3]	4.8 [28.7]
refined residues	198	198	198
refined ligand atoms	45	41	53
refined water molecules	182	145	134
refined chlorides	3	3	2
refined glycerols	1	1	2
resolution in refinement (Å)	10-1.50	10-1.78	10-1.48
R_{cryst} (%) $F > 4 \sigma F_o$; F_o	16.8	16.3	17.0
R_{free} (%) $F > 4 \sigma F_o$; F_o	20.2	21.1	19.3
mean B-factor (Å^2)			
(peptide chain A; B)	14.4; 13.0	17.1; 15.5	16.3; 14.4
main-chain (Å^2)	12.3; 13.0	14.7; 13.2	13.8; 12.2
side-chains (Å^2)	16.9; 15.4	19.8; 17.9	19.1; 16.9
ligand (Å^2)	13.5	21.2	14.1
water (Å^2)	24.7	26.7	26.2
Cl ions (Å^2)	16.7	17.1	18.4
Glycerol (Å^2)	22.9	40.3	30.5
Ramachandran plot			
most favorate geometry (%)	96.2	96.8	96.8
additionally allowed (%)	3.8	3.2	3.2
generously allowed (%)	-	-	-
disallowed (%)	-	-	-

Table 5: X-ray data processing and refinement for the HIV protease complexes of derivatives **6fd** **6gd** and **6fg**. (a) Values in brackets refer to the highest resolution shell.

3.6 References

1. UNAIDS: *2006 Report on the global AIDS epidemic.*
2. Pomerantz, R. J.; Horn, D. L., Twenty years of therapy for HIV-1 infection. *Nat. Med.* **2003**, 9, (7), 867-873.
3. Wlodawer, A.; Erickson, J. W., Structure-Based Inhibitors of HIV-1 Protease. *Annu. Rev. Biochem.* **1993**, 62, (1), 543-585.
4. Rodríguez-Barrios, F.; Gago, F., HIV Protease Inhibition: Limited Recent Progress and Advances in Understanding Current Pitfalls. *Curr. Top. Med. Chem.* **2004**, 4, 991-1007.
5. Chrusciel, R. A.; Strohbach, J. W., Non-Peptidic HIV Protease Inhibitors. *Curr. Top. Med. Chem.* **2004**, 4, (10), 1097-1114.
6. Randolph, J. T.; DeGoey, D. A., Peptidomimetic Inhibitors of HIV Protease. *Curr. Top. Med. Chem.* **2004**, 4, (10), 1079-1095.
7. D’Aquila, R. T.; Schapiro, J. M.; Brun-Vézinet, F.; Clotet, B.; Conway, B.; Demeter, L. M.; Grant, R. M.; Johnson, V. A.; Kuritzkes, D. R.; Loveday, C.; Shafer, R. W.; Richman, D. D., Drug Resistance Mutations in HIV-1. *Top. HIV Med.* **2002**, 10, (5), 21-25.
8. Güller, R.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Hirth, G.; Jenny, C.; Kansy, M.; Montavon, F.; Müller, M.; Oefner, C.; Stadler, H.; Vieira, E.; Wilhelm, M.; Wostl, W.; Märki, H. P., Piperidine-renin inhibitors compounds with improved physicochemical properties. *Bioorg. Med. Chem. Lett.* **1999**, 9, (10), 1403-1408.
9. Vieira, E.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Güller, R.; Hirth, G.; Märki, H. P.; Müller, M.; Oefner, C.; Scalone, M.; Stadler, H.; Wilhelm, M.; Wostl, W., Substituted piperidines - highly potent renin inhibitors due to induced fit adaptation of the active site. *Bioorg. Med. Chem. Lett.* **1999**, 9, (10), 1397-1402.
10. John, V.; Beck, J. P.; Bienkowski, M. J.; Sinha, S.; Heinrikson, R. L., Human β-Secretase (BACE) and BACE Inhibitors. *J. Med. Chem.* **2003**, 46, (22), 4625-4630.
11. Prade, L.; Jones, A. F.; Boss, C.; Richard-Bildstein, S.; Meyer, S.; Binkert, C.; Bur, D., X-ray Structure of Plasmepsin II Complexed with a Potent Achiral Inhibitor. *J. Biol. Chem.* **2005**, 280, (25), 23837-23843.
12. Hof, F.; Schütz, A.; Fäh, C.; Meyer, S.; Bur, D.; Liu, J.; Goldberg, D. E.; Diederich, F., Starving the Malaria Parasite: Inhibitors Active against the Aspartic Proteases Plasmepsins I, II, and IV. *Angew. Chem., Int. Ed.* **2006**, 45, (13), 2138-2141.

13. Specker, E.; Böttcher, J.; Lilie, H.; Heine, A.; Schoop, A.; Müller, G.; Griebenow, N.; Klebe, G., An Old Target Revisited: Two New Privileged Skeletons and an Unexpected Binding Mode For HIV-Protease Inhibitors. *Angew. Chem., Int. Ed.* **2005**, 44, (20), 3140-3144.

14. Specker, E.; Böttcher, J.; Brass, S.; Heine, A.; Lilie, H.; Schoop, A.; Müller, G.; Griebenow, N.; Klebe, G., Unexpected Novel Binding Mode of Pyrrolidine-Based Aspartyl Protease Inhibitors: Design, Synthesis and Crystal Structure in Complex with HIV Protease. *ChemMedChem* **2006**, 1, (1), 106-117.

15. Czodrowski, P.; Sotriffer, C. A.; Klebe, G., Atypical Protonation States in the Active Site of HIV-1 Protease: A Computational Study. *J. Chem. Inf. Model.* **2007**, 47, (4), 1590-1598.

16. Nagel, U., Asymmetric Hydrogenation of α-(Acetylamino)cinnamic Acid with a Novel Rhodium Complex; the Design of an Optimal Ligand. *Angew. Chem., Int. Ed.* **1984**, 23, (6), 435-436.

17. Nagel, U.; Kinzel, E.; Andrade, J.; Prescher, G., Enantioselektive Katalyse, 4. Synthese *N*-substituierter (*R*,*R*)-3,4-Bis(diphenylphosphino)-pyrrolidine und Anwendung ihrer Rhodiumkomplexe zur asymmetrischen Hydrierung von α-(Acylamino)acrylsäure-Derivaten. *Chem. Ber.* **1986**, 119, (11), 3326-3343.

18. Rocha Gonsalves, A. M. d'. A.; Serra, M. E. S.; Murtinho, D.; Silva, V. F.; Matos Beja, A.; Paixão, J. A.; Ramos Silva, M.; Alte da Veiga, L., Pyrrolidine-based amino alcohols: novel ligands for the enantioselective alkylation of benzaldehyde. *Journal of Molecular Catalysis A: Chemical* **2003**, 195, (1-2), 1-9.

19. Skarżewski, J.; Gupta, A., Synthesis of C_2 symmetric primary vicinal diamines. Double stereospecific Mitsunobu reaction on the heterocyclic diols derived from tartaric acid. *Tetrahedron: Asymmetry* **1997**, 8, (11), 1861-1867.

20. Rodríguez Sarmiento, R. M.; Wirz, B.; Iding, H., Chemoenzymatic preparation of non-racemic *N*-Boc-pyrrolidine-3,4-dicarboxylic acid 3-ethyl esters and their 4-hydroxymethyl derivatives. *Tetrahedron: Asymmetry* **2003**, 14, (11), 1547-1551.

21. Kuppert, D.; Sander, J.; Roth, C.; Wörle, M.; Weyhermüller, T.; Reiss, G. J.; Schilde, U.; Müller, I.; Hegetschweiler, K., The Coordination Chemistry of *cis*-3,4-Diaminopyrrolidine and Related Polyamines. *Eur. J. Inorg. Chem.* **2001**, (10), 2525-2542.

22. Song, C. E.; Yang, J. W.; Roh, E. J.; Lee, S.-g.; Ahn, J. H.; Han, H., Heterogeneous Pd-Catalyzed Asymmetric Allylic Substitution Using Resin-Supported Trost-Type Bisphosphane Ligands. *Angew. Chem., Int. Ed.* **2002**, 41, (20), 3852-3854.
23. Katritzky, A. R.; Pilarski, B.; Urogdi, L., Efficient Conversion of Nitriles to Amides with Basic Hydrogen Peroxide in Dimethyl Sulfoxide. *Synthesis* **1989**, (12), 949-950.
24. Kettler, K.; Sakowski, J.; Wiesner, J.; Ortmann, R.; Jomaa, H.; Schlitzer, M., Novel lead structures for antimalarial farnesyltransferase inhibitors. *Pharmazie* **2005**, 60, (5), 323-327.
25. Taylor, A.; Brown, D. P.; Kadam, S.; Maus, M.; Kohlbrenner, W. E.; Weigl, D.; Turon, M. C.; Katz, L., High-level expression and purification of mature HIV-1 protease in *Escherichia coli* under control of the *araBAD* promoter. *Appl. Microbiol. Biotechnol.* **1992**, 37, (2), 205-210.
26. Toth, M. V.; Marshall, G. R., A simple, continuous fluorometric assay for HIV protease. *Int. J. Pept. Protein Res.* **1990**, 36, (6), 544-550.
27. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode. In *Methods Enzymol.*, Carter Jr., C. W., Ed. Academic Press: 1997; Vol. 276, pp 307-326.
28. Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Cryst. D* **2004**, 60, (3), 432-438.
29. Pape, T.; Schneider, T. R., *HKL2MAP*: a graphical user interface for macromolecular phasing with *SHELX* programs. *J. Appl. Cryst.* **2004**, 37, (5), 843-844.
30. Usón, I.; Sheldrick, G. M., Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **1999**, 9, (5), 643-648.
31. Sheldrick, G. M., Macromolecular phasing with *SHELXE*. *Zeitschrift für Kristallographie* **2002**, 217, (12), 644-650.
32. Cowtan, K. D.; Zhang, K. Y. J., Density modification for macromolecular phase improvement. *Prog. Biophys. Mol. Biol.* **1999**, 72, (3), 245-270.
33. Lamzin, V. S.; Wilson, K. S., Automated Refinement of Protein Models. *Acta Cryst. D* **1993**, 49, (1), 129-147.
34. The *CCP*4 Suite: Programs for Protein Crystallography. *Acta Cryst. D* **1994**, 50, (5), 760-763.
35. Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L., *Crystallography & NMR System*: A New

Software Suite for Macromolecular Structure Determination. *Acta Cryst. D* **1998**, 54, (5), 905-921.

36. Sheldrick, G. M.; Schneider, T. R., *SHELXL*: High-resolution refinement. In *Methods Enzymol.*, Charles, W. C. J.; Robert, M. S., Eds. Academic Press: 1997; Vol. 277, pp 319-343.

37. Emsley, P.; Cowtan, K., *Coot*: model-building tools for molecular graphics. *Acta Cryst. D* **2004**, 60, (12 Part 1), 2126-2132.

4. Structural and Kinetic Analysis of Pyrrolidine-based Inhibitors of the Drug Resistant Ile84Val Mutant of HIV-1 Protease *

4.1 Introduction

Human immunodeficiency virus (HIV) protease is a well established target for the development of antiviral therapeutics.[1] The protease processes the *gag* and *pol* encoded polyproteins into functional enzymes which are essential for the viral replication.[2] Inhibition of the viral protease leads to immature non-infectious virions.[3] Currently, nine HIV protease inhibitors (PIs) are approved by the FDA.[4] The *highly active antiretroviral therapy* (HAART), which combines protease and reverse transcriptase inhibitors, not only significantly prolongs but also improves the quality of the patient's life.[5, 6] However, an entire remedy of the infection still remains an unaccomplished goal. Moreover, although most of the inhibitors initially exhibit a strong inhibitory effect, this efficacy decreases over time due to the continuously new formation of drug resistant virus variants.

The relatively low fidelity of the viral reverse transcriptase and the fast replication rate are the main driving factors for the high mutation rate of the HIV. It has been estimated that up to 10^5 point mutations occur daily within the viral population of an infected individual.[7] Due to high mutation rates, there are several non-identical genomes termed viral quasispecies, hence making it difficult to define one wild-type protease as reference point.[8] The following studies were performed using a clone of the BH10 isolate.

---

* Taken from original publication, Jark Böttcher, Andreas Blum, Andreas Heine, Wibke E. Diederich, Gerhard Klebe. Structural and Kinetic Analysis of Pyrrolidine-based Inhibitors of the Drug Resistant Ile84Val Mutant of HIV-1 Protease. *J. Mol. Biol* **2008**, (383),2, 347-57

Mutations in the binding pocket as well as in distal sites of the HIV protease affect inhibitor and substrate binding by altering the number and strength of subsite interactions. Since the applied inhibitors effectively inhibit wild-type proteases and limit wild-type duplication, they also provide a selective advantage to those HIV variants with reduced susceptibility. The occurrence of multi-drug resistant HIV variants diminishing the efficacy of all protease inhibitors in clinical use hampers an effective antiviral treatment of patients infected by these variants. Under selective drug pressure, mutations have been proven to occur nearly in one-half of the HIV-1 protease amino acid positions.[9] The mutations can generally be classified into major and minor ones.[10] Major mutations are initially selected by drug treatment and mainly occur at residues forming interactions upon drug binding. Minor mutations consecutively improve the replicative fitness of virus variants carrying major mutations.

The high adaptability of the virus demands a continuous and persistent search for new inhibitors. Novel inhibitors should exhibit a different mutation profile compared to that of the marketed drugs to circumvent the development of cross-resistance. Most of the currently approved inhibitors are transition state mimics targeting the catalytic aspartates 25A and 25B via a secondary hydroxyl group. As an alternative skeleton, cyclic amines have been proposed as novel anchoring groups.[11] Very recently, we have described the design and synthesis of C_2-symmetric 3,4-disubstituted pyrrolidines as a new class of HIV-1 protease inhibitors. The cocrystal structures of six derivatives were determined in complex with wild-type HIV-1 protease (compound (PDB ID): **1 (**2PQZ); **3** (2QNP); **6** (2QNQ); **8** (2PWC); **9** (2PWR); **10** (2QNN) Table 2) and further utilized within a structure-guided optimization process.[12] These crystal structures revealed a conserved binding mode for all investigated derivatives, schematically represented in Figure 1: The endocyclic amino function addresses the catalytic dyad and a unique flap interaction pattern is observed. Only one sulfonyl group of the C_2-symmetric inhibitor is involved in two hydrogen bonds to the backbone NHs of Ile50A and Ile50B each formed by one of the sulfonyl oxygen atoms. The second sulfonyl group remains uncoordinated and establishes no polar contacts. This leads to an overall slightly asymmetric binding mode of the C_2-symmetric inhibitors.

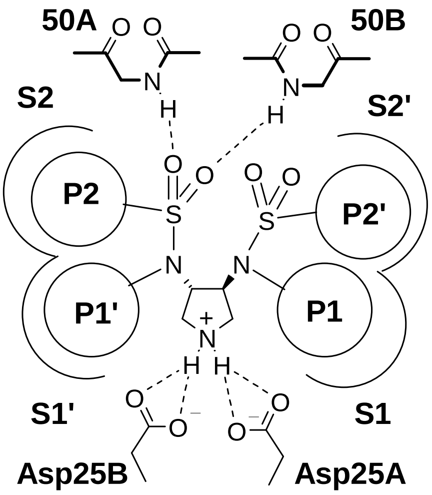


Figure 1: Schematic representation of the conserved binding mode of the pyrrolidine-3,4-bis-*N*-benzyl-sulfonamides. Hydrogen bonds are indicted by broken lines and main-chain bonds are denotated by bold lines.

However, each symmetry-related subpocket is occupied by the corresponding symmetry-related substituent of the inhibitor. The cocrystal structure of the most potent derivative (compound **10**)[12] is shown together with the cocrystal structures of all currently approved HIV protease inhibitors in Figure 2 (inhibitor (PDB ID): Ritonavir (1HXW),[13] Atazanavir (2AQU),[14] Darunavir (1T3R),[15] Amprenavir (1HPV),[16] Indinavir (1HSG),[17] Nelfinavir (1OHR),[18] Saquinavir (1HXB),[19] Lopinavir (2O4S)[20] and Tipranavir (2O4P)[20]).

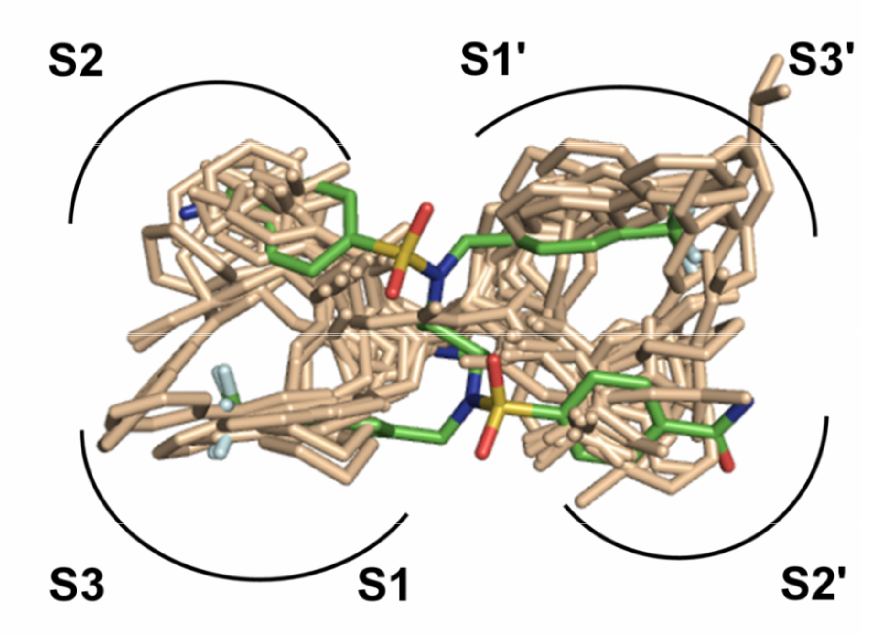


Figure 2: Superposition of the final lead structure (compound **10)** with the cocrystal structures of all approved HIV protease inhibitors (PIs). Compound **10** in green, color-coded by atom type, whereas the ligand geometries of the approved PIs are displayed in wheat.

This comparison reveals a completely different mode in targeting the protease's subsites by this new class of inhibitors: Whereas the approved inhibitors bind to the subsites successively from S3 to S3', the pyrrolidine-based inhibitors occupy the S1-S3/S2' and the S1'-S3'/S2 pockets via the two substituents at the pyrrolidine ring, respectively. This leads to a unique binding mode in the active site, particularly at the borders of the S1/S2' and the S1'/S2 pockets, composed by the amino acids Ile50 and Ile84 of chain A and B. Both residues are referred to as major mutations. Whereas the Ile84Val mutation in HIV-1 protease is associated with resistance to all approved inhibitors, the Ile50Val mutation is mainly linked to the resistance against Amprenavir.[21] Due to their importance and the remarkable differences in binding of the pyrrolidine-based inhibitors, particularly in the contact area next to Ile50 and Ile84, we selected these two mutations for further investigations. Both mutations have already been studied extensively by kinetic measurements,[22, 23] structural biology,[24, 25] isothermal titration calorimetry,[26, 27] and computational methods.[28] These studies attribute the reduced affinity with respect to the Ile84Val mutant primarily to the loss of Van der Waals contacts with the P1/P1' moieties of the inhibitors. The sensitivity of Amprenavir towards the Ile50Val mutation is attributed to strong interactions of the C_δ of Ile50A with the P2'-residue of the inhibitor.[26]

Encouraged by the deviating and up to now unique mode in which the C_2-symmetric pyrrolidine-based inhibitors bind to the protease's subsites, we investigated their potential against these drug-resistant mutations. The two active site mutants Ile50Val and Ile84Val were generated by site directed mutagenesis. Followed by kinetic studies two representatives **8** and **9** were crystallized in complex with the wild-type and the two corresponding mutant proteases. Throughout the paper the wild-type protease (BH10 isolate) is referred to as PR_{WT} and the mutant proteases PR_{I50V} and PR_{I84V}, respectively.

4.2 Results

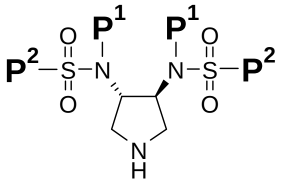
4.2.1 Kinetic characterization

In order to estimate the consequences of the point mutations to putative natural substrates, the kinetic parameters of PR_{WT}, PR_{I50V} and PR_{I84V} were determined in a fluorescence-based assay (Table 1) using four different commercially available substrates. The four fluorescence substrates were purchased from Bachem (substrates A and B) and Sigma Aldrich (substrates C and D). PR_{I50V} and PR_{I84V} exhibited a significantly reduced K_m to all studied substrates compared to PR_{WT}. For PR_{I50V}, the affinity towards the substrates is reduced tenfold whereas for PR_{I84V}, an affinity reduction by a factor of five was determined. The reduced catalytic efficiency (k_{cat}/K_m) is driven by this loss in affinity of the mutant proteases towards these substrates.

Anthranilyl-HIV protease substrate	**WT**		**PR_{I50V} (/WT)**		**PR_{I84V} (/WT)**	
	K_m [µM]	k_{cat}/K_m [µM^{-1}s^{-1}]	K_m [µM]	k_{cat}/K_m [µM^{-1}s^{-1}]	K_m [µM]	k_{cat}/K_m [µM^{-1}s^{-1}]
A	14.6 ±0.9	0.41 ±0.07	139 ±9 **(10)**	0.013 ±0.003 **(0.03)**	70.6 ±1.3 **(5)**	0.057 ±0.003 **(0.03)**
B	22.7 ±2.6	0.36 ±0.13	158 ±24 **(7)**	0.048 ±0.023 **(0.1)**	72.0 ±7.2 **(3)**	0.090 ±0.026 **(0.1)**
C	9.7 ±0.6	1.87 ±0.34	117 ±5 **(12)**	0.21 ±0.025 **(0.1)**	51.3 ±2.6 **(5)**	0.46 ±0.056 **(0.1)**
D	5.2 ±0.4	1.66 ±0.36	54.9 ±1.9 **(11)**	0.21 ± 0.026 **(0.1)**	27.7 ±2.1 **(5)**	0.41 ±0.08 **(0.1)**

Table 1: Kinetic characterization of PR_{WT}, PR_{I50V} and PR_{I84V} with fluorogenic substrates. Values in parentheses refer to the relative change compared to the wild-type protease.

The affinities of the pyrrolidine-based inhibitors were determined towards the protease variants using substrate A, and the results are summarized in Table 2. In case of the Ile50Val mutation, on average, a decrease in affinity of the pyrrolidine-based inhibitors by a factor of five is observed. This overall decrease in affinity seems therefore to occur as a general phenomenon of this class of compounds and not a result of a special substitution pattern. In contrast, the pyrrolidine-based inhibitors exhibit an equal or even improved affinity towards the Ile84Val mutant: Pyrrolidine **1** exhibits a twofold increase in affinity. Compounds with only one additional hydrophobic substituent at the P1/P1’-aromatic rings (**2**, **3** and **4**) at least retain their affinity, whereas all other derivatives carrying additional polar or hydrophobic substituents at the P2/P2’ moieties exhibit significantly improved potency. Particularly for the amide-substituted derivatives (**9** and **10**), an up to sevenfold increase in affinity towards this drug-resistant point mutant is observed.



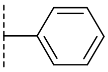
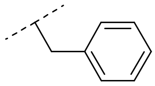
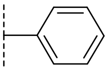
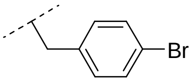
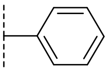
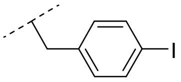
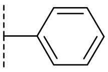
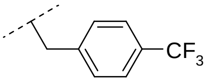
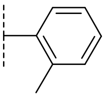
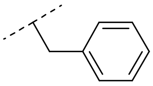
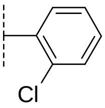
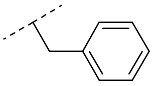
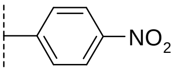
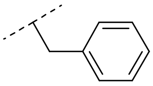
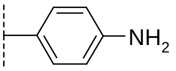
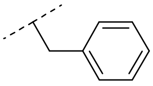
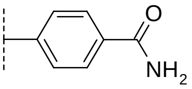
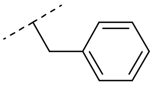
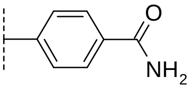
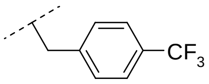
	$\mathbf{P^2}$	$\mathbf{P^1}$	$\mathbf{PR_{WT}}$ K_i [μM]	$\mathbf{PR_{I50V}}$ K_i [μM]	$\mathbf{PR_{I84V}}$ K_i [μM]	K_i I50V/WT	K_i I84V/WT
1			2.2	10.9	1.1	5	0.5
2		 Br	0.46	2.6	0.55	6	1.2
3		 I	0.39	1.4	0.33	4	0.8
4		 CF_3	0.80	3.0	0.50	4	0.6
5			0.67	2.1	0.46	3	0.7
6	 Cl		0.77	2.6	0.48	3	0.6
7	 NO_2		1.7	8.0	0.92	5	0.5
8	 NH_2		0.27	1.0	0.128	4	0.5
9	 O, NH_2		0.26	1.4	0.036	5	0.1
10	 O, NH_2	 CF_3	0.07	0.26	0.012	4	0.2

Table 2: K_i-values of the inhibitors towards the wild-type and mutant HIV proteases.

To investigate the contrasting behaviour of the inhibitors towards the single point mutants PR_{I50V} and PR_{I84V}, we determined the cocrystal structures of two derivatives in complex with the mutant proteases. Due to the largest relative difference in affinity towards the mutants, along with good potency, compounds **8** and **9** were selected, exhibiting deviating P2/P2’ substituents.

4.2.2 Structural analysis

The amino (**8**) and carboxamido (**9**) derivatives were crystallized in complex with the PR_{I50V} and PR_{I84V} variants. The crystallographic data and refinement statistics of the four structures are listed in Table 3. All cocrystal structures, discussed within this study, crystallized with identical packing, in the space group $P2_12_12$ with very similar lattice constants. Therefore, an influence of crystal packing on the observed differences can be neglected. In the space group $P2_12_12$, the asymmetric unit contains the HIV-1 protease with two symmetry-independent chains labelled as 1A to 99A and 1B to 99B, respectively. The collected datasets have comparable quality with resolutions between 1.58 Å and 1.92 Å. In all cases, the ligand geometries were clearly visible in the F_o-F_c omit-map at a sigma level of 3.0 (Figure 3) and could be refined as one single conformer. All structures were determined by molecular replacement and consecutively refined following the same protocol.

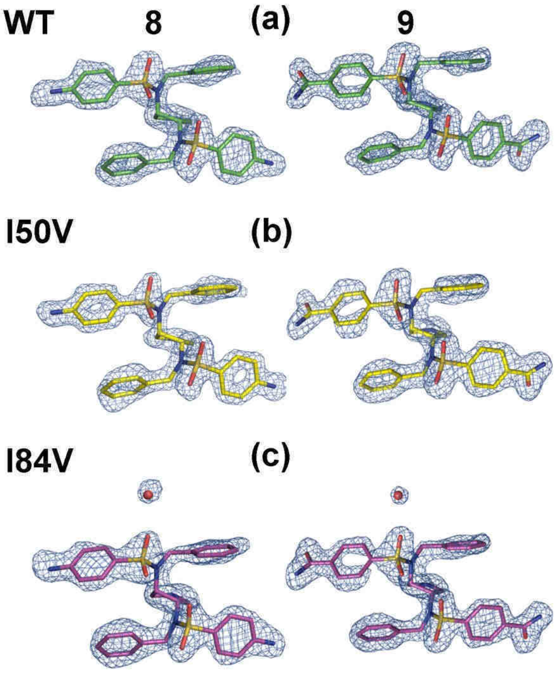


Figure 3: Ligand geometries of **8** and **9** color-coded by atom type in the cocrystal structures with PR_{WT}: (a) green; PR_{I50V}: (b) yellow; PR_{I84V}: (c) magenta; The F_o-F_c omit map for the ligand and the water in case of (c) is displayed at a σ level of 3.0 as blue mesh.

	8 **PR_{I50V}**	**8** **PR_{I84V}**	**9** **PR_{I50V}**	**9** **PR_{I84V}**
resolution (Å)	25-1.58	25-1.92	30-1.80	30-1.81
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
cell dimensions (Å)	a = 57.9	a = 57.3	a = 57.2	a = 57.4
	b = 86.0	b = 85.7	b = 85.5	b = 85.7
	c = 46.3	c = 46.6	c = 46.7	c = 46.6
highest resolution shell (Å)	1.61-1.58	1.95 -1.92	1.83 -1.80	1.84 -1.81
no. of measured reflections	141102	76478	89944	75858
no. of independent reflections	31985	17926	20882	21217
completeness (%)[a]	99.1 [90.6]	98.6 [97.0]	95.5 [100]	97.8 [100]
I/σ[a]	13.5 [2.5]	15.8 [3.2]	18.8 [3.1]	16.2 [2.6]
R_{sym} (%)[a]	9.0 [37.5]	9.1 [46.3]	7.3 [49.1]	7.0 [42.2]
resolution in refinement (Å)	10-1.58	10-1.92	10-1.80	10-1.81
R_{cryst} ($F > 4\sigma F_{o;} F_o$) (%)	17.0 ; 19.8	17.8 ; 20.1	17.7 ; 19.3	18.1 ; 20.4
R_{free} ($F > 4\sigma F_{o;} F_o$) (%)	20.7 ; 24.5	23.0 ; 26.4	23.3 ; 25.3	23.8 ; 26.0
mean B-factor ($Å^2$) (peptide chain A; B)	15.2 ; 13.5	27.3 ; 24.9	31.9 ; 28.4	23.3 ; 21.1
main-chain ($Å^2$)	12.3 ; 11.2	24.8 ; 22.5	28.3 ; 25.8	20.7 ; 19.0
side-chains ($Å^2$)	18.2 ; 15.9	30.2 ; 27.5	35.9 ; 31.2	26.4 ; 23.5
ligand ($Å^2$)	20.3	27.1	25.3	20.7
water molecules ($Å^2$)	24.8	31.0	35.4	27.9
Ramachandran plot				
most favorate geometry (%)	96.8	95.6	97.5	94.9
additionally allowed (%)	3.2	4.4	2.5	5.1
rmsd bonds (Å)	0.007	0.005	0.006	0.006
rmsd angles (°)	2.0	1.9	2.0	1.9

Table 3: Crystallographic data for the complexes of the amino derivative **8** and the carboxamido derivative **9** Values in brackets refer to the highest resolution shell.

4.2.2.1 Crystal structures of 8

The crystal structures of the amino derivative **8** in complex with the HIV protease variants were determined with a resolution of 1.58 Å (PR_{I50V}) and 1.92 Å (PR_{I84V}), respectively. Both complexes exhibit a high level of similarity to the corresponding wild-type complex (determined at a resolution of 1.78 Å), studied previously.[12] This similarity is reflected by the root-mean-square deviation (rmsd) between the C_α atoms of 0.12 Å for PR_{I50V} and 0.18 Å for PR_{I84V} compared to the PR_{WT} complex. Using the C_α-based alignment, the ligand geometries also show a high level of similarity with an rmsd of 0.19 Å for the PR_{I50V} and 0.28 Å for the PR_{I84V} in comparison to the PR_{WT} ligand coordinates (Figure 4).

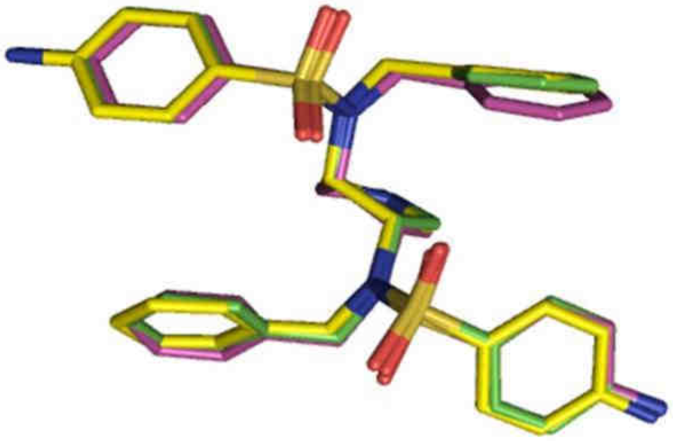


Figure 4: Ligand geometries of the C_α superposition of the cocrystal structures of **8** in complex with PR_{WT} (green), PR_{I50V} (yellow), PR_{I84V} (magenta), color-coded by atom type.

The overall binding mode is conserved and schematically represented in Figure 5, the distances of polar contacts are given in Table 4.

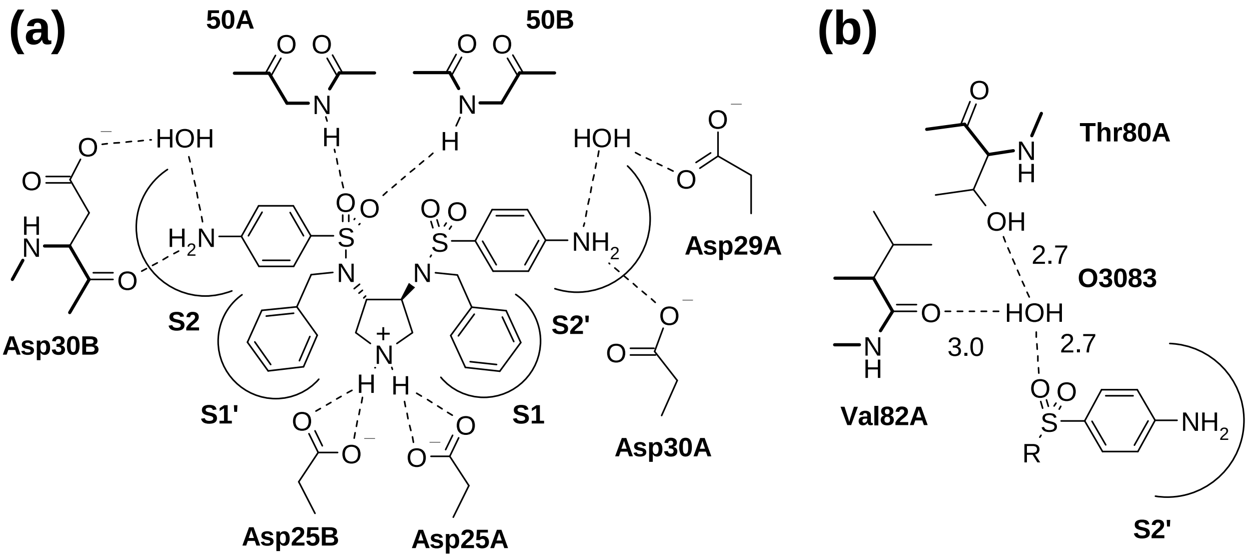


Figure 5: (a) Schematic representation of the binding modes of **8** observed in the crystal structures in complex with PR$_{WT}$, PR$_{I50V}$ and PR$_{I84V}$. (b) Additional hydrogen bond network of water O3083 in complex with PR$_{I84V}$. Hydrogen bonds are indicted by dashed and main-chain bonds using bold lines.

	PR$_{WT}$	**PR$_{I50V}$**	**PR$_{I84V}$**
D25 O_δ	2.6-3.0	2.7-3.0	2.7-3.0
I/V50A N	3.1	3.0	3.2
I/V50B N	3.7	3.8	3.7
D29A O_δ/H_2O	3.2/2.9	3.3/2.8	-
D30A O_δ	2.7	2.7	2.7
D30B O_δ/H_2O	2.8/2.7	2.9/2.7	2.8/2.8
D30B O	3.2	3.1	3.1

Table 4: Polar contacts of the ligand **8** observed in the cocrystal structures with the different HIV-protease variants. Values are given in Å.

The pivotal nitrogen of the pyrrolidine core structure interacts with the catalytic dyad. In case of the wild-type complex, the nitrogen atom is in close contact with the carboxylic oxygen atoms of Asp25A and Asp25B. This important interaction is conserved for the mutant proteases. The flap interactions formed via hydrogen bonds of the sulfonyl oxygens to the main-chain NH of Ile50A and Ile50B or Val50A and Val50B are comparable in all complexes. The inhibitor's benzyl moieties occupy the S1 and S1' pockets whereas the *p*-amino-phenyl sulfonamide groups bind in the S2 and S2' pockets. The amino substituents form a hydrogen-bond network, which, due to the asymmetric binding mode, differs in S2 and S2'. The observed polar interactions in S2 are highly conserved: The amino group hydrogen bonds to the main-chain carbonyl oxygen of Asp30B and, mediated via a water molecule, to the corresponding side-chain carboxyl oxygen atom. In the S2' pocket, differences are noticed: The water-mediated interaction to the carboxyl function of Asp29A is only observed in PR_{WT} and PR_{I50V}, whereas the H-bond to the side-chain of Asp30A is conserved in all complexes. In PR_{WT} and PR_{I50V}, the C_δ atoms of Ile84A and Ile84B are in close Van der Waals contact to the benzylic methylene groups of the ligand (3.5/3.3 Å PR_{WT} and 3.6/3.4 Å PR_{I50V}), compared to C_γ of the corresponding Val84A and Val84B in PR_{I84V} (3.8/4.0 Å). The burial of the ligand surface varies slightly between 95% and 96%, exhibiting the highest value in the PR_{WT} complex.

The most pronounced differences are observed in the PR_{I84V} complex: The Ile84Val exchange at the border of the S1/S2' pockets enlarges the binding site, and the newly created additional space is filled by a water molecule (O3083, Figure 5b). This water mediates polar interactions between one oxygen atom of the sulfonyl group of **8** and to the main-chain carbonyl of Val82A and the side-chain of Thr80A. The binding pockets of the three structures were consecutively analyzed with a water probe using the program GRID to evaluate whether the position of the additional water is energetically favourable. In Figure 6c, the GRID interaction fields for the PR_{I84V} binding pocket are displayed.

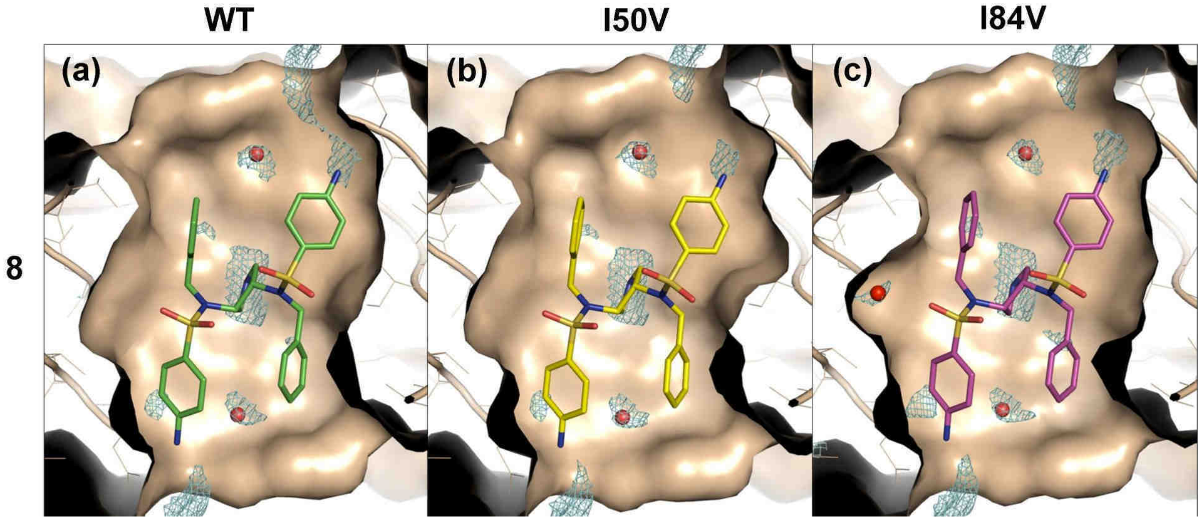


Figure 6: The calculated GRID interaction fields for the binding pockets of PR_{WT} (a) PR_{I50V} (b) and (PR_{I84V} (c) in complex with **8** are displayed as blue mesh at a contour level of 8 kcal/mol. The protein surface is schematically represented in wheat, the corresponding ligand is color-coded by atom type in green, yellow and magenta for PR_{WT}, PR_{I50V} and PR_{I84V}, respectively. Experimental observed water molecules are shown as red spheres.

A contour is observed at the position of the additionally enclosed water molecule, together with an area next to the catalytic Asp25A and Asp25B and at the far end of S2 and S2' and near Gly27A and Gly27B, which are part of the peptide recognition motif. The positions of the experimentally observed water molecules are in good agreement with the calculated fields. In case of the PR_{WT} (Figure 6a) and the PR_{I50V} (Figure 6b) complexes, similar observations are made except for the GRID interaction field near Thr80A and Val82A.

4.2.2.2 Crystal structures of 9

The crystal structures of the carboxamido derivative **9** were determined with a resolution of 1.80 Å (PR_{I50V}) and 1.81 Å (PR_{I84V}) (Table 3). As for **8**, the overall binding of **9** is conserved compared to the wild-type counterpart (1.50 Å), described previously.[12] Compared to the PR_{WT} structure, the C_α rmsd of PR_{I50V} is 0.23 Å and of PR_{I84V} is 0.24 Å. The corresponding rmsd for the ligand atoms using this alignment is 0.19 Å for PR_{I50V} and 0.28 Å for PR_{I84V} (Figure 7).

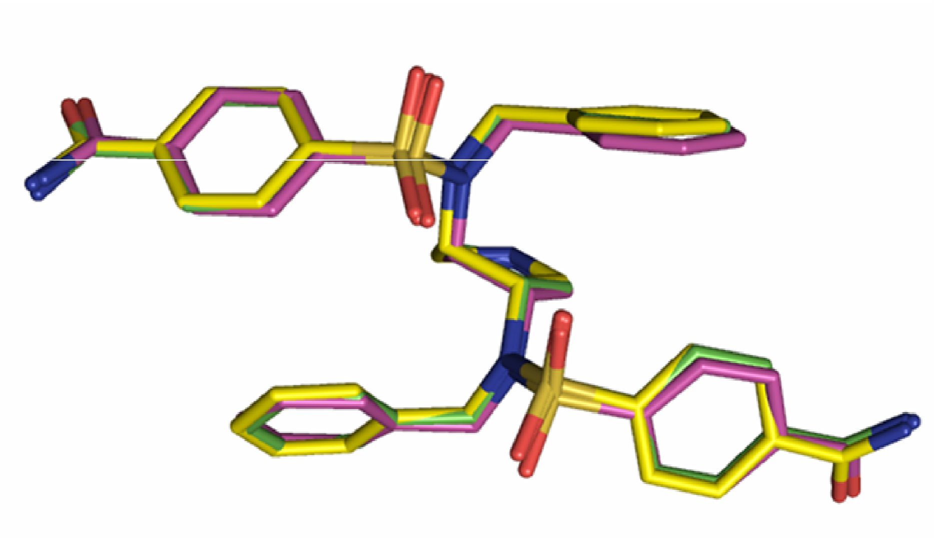


Figure 7: Ligand geometries of the C_α superposition of the cocrystal structures of **9** in complex with PR_{WT} (green), PR_{I50V} (yellow), PR_{I84V} (magenta), color-coded by atom type.

The same subsite occupancy is observed in all complexes and is schematically represented in Figure 8a.

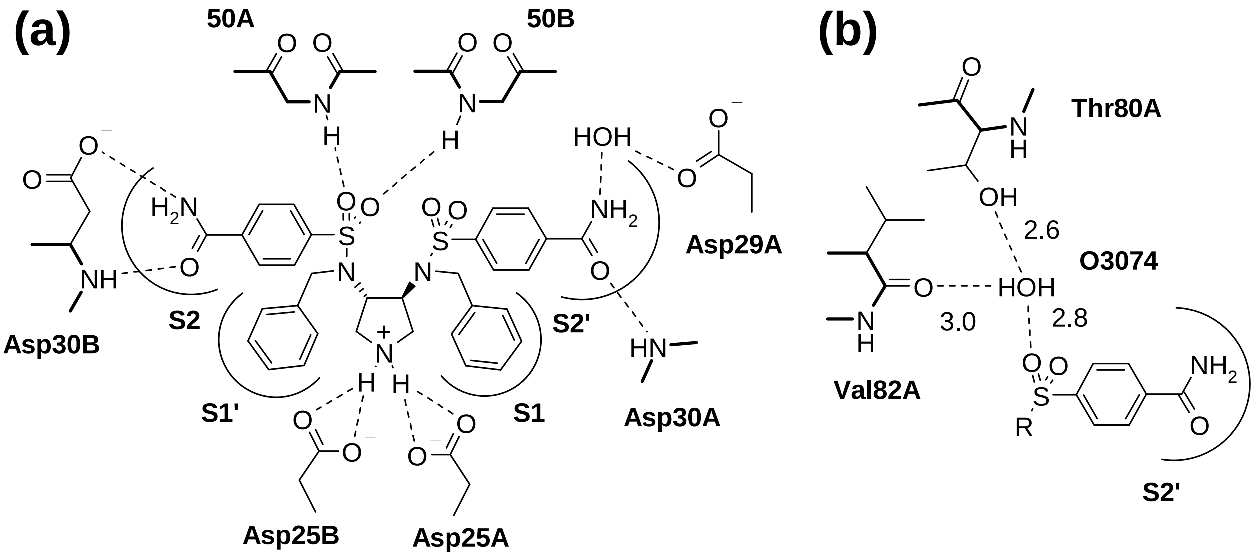

(a)
50A
50B
Asp30B
S2
S1'
Asp25B
Asp25A
S1
S2'
Asp29A
Asp30A
HOH
(b)
Thr80A
O3074
Val82A
2.6
3.0
2.8
S2'
R

Figure 8: (a) Schematic representation of the binding modes of **9** observed in the crystal structures in complex with PR_{WT}, PR_{I50V} and PR_{I84V}. (b) Additional hydrogen bond network of water O3074 in complex with PR_{I84V}. Hydrogen bonds are indicted by dashed and main-chain bonds using bold lines.

The hydrogen bond network established by the inhibitor is conserved, and the polar contact distances are given in Table 5.

	PR_{WT}	**PR_{I50V}**	**PR_{I84V}**
D25 O_{δ}	2.7-3.0	2.7-3.0	2.7-2.9
I/V50A N	3.0	3.0	3.2
I/V50B N	3.2	3.5	3.6
D29A O_{δ}/H_2O	3.0/2.7	3.4/2.5	3.1/2.4
D30A N	3.0	3.0	3.0
D30B O_{δ}	2.9	2.9	3.1
D30B N	2.8	3.0	2.9

Table 5: Polar contacts of the ligand **9** observed in the cocrystal structures with the different HIV-protease variants. Values are given in Å.

The pivotal nitrogen is in close polar contact to the carboxylate oxygens of Asp25A and Asp25B, while the flap only interacts with the oxygen atoms of one sulfonyl group forming hydrogen bonds to the main-chain NHs of amino acid residues 50A and 50B. In S2, direct interactions of the carboxamido oxygen to the main-chain NH of Asp30B and of the carboxamido nitrogen to the corresponding Asp30B side-chain are observed in all complexes. The water-mediated interaction to the side-chain of Asp29A and the direct interaction to the backbone NH of Asp30A in S2’ exhibit a high level of similarity. The C_{δ} of Ile84A and Ile84B are in close contact to the benzylic methylene groups of the ligand (3.3/3.3 Å PR_{WT} and 3.5/3.6 Å PR_{I50V}). In PR_{I84V}, the C_{γ} of Val84A and Val84B only show Van der Waals contacts of 3.8 and 4.1 Å. The burial of the ligand surface within the binding pocket sums up to 95% in case of PR_{I50V} and to 96% in case of PR_{WT} and PR_{I84V}.

An intercalation of a water molecule (O3074), is observed in PR_{I84V} and results in additional polar protein-ligand interactions. The water molecule establishes hydrogen bonds to backbone carbonyl of Val82A, to side-chain of Thr80A, and to one ligand sulfonyl oxygen atom (Figure 8b). The binding pockets were analyzed with the program GRID. The resulting interaction fields of a water probe for PR_{I84V} are displayed in Figure 9c.

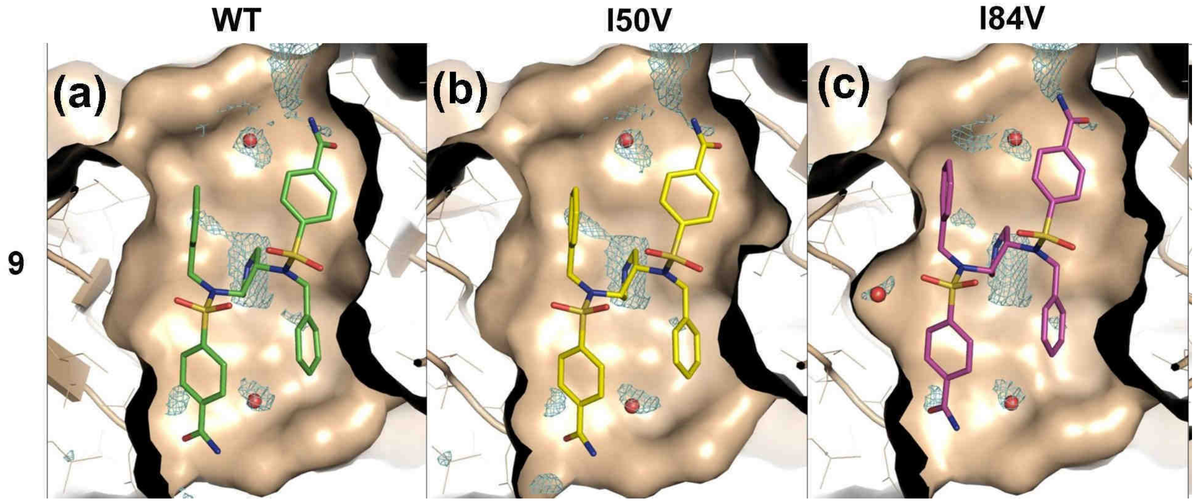


Figure 9: The calculated GRID interaction fields for the binding pockets of PR_{WT} (a) PR_{I50V} (b) and (PR_{I84V} (c) in complex with **9** are displayed as blue mesh at a contour level of 8 kcal/mol. The protein surface is schematically represented in wheat, the corresponding ligand is color-coded by atom type in green, yellow and magenta for PR_{WT}, PR_{I50V} and PR_{I84V}, respectively. Experimental observed water molecules are shown as red spheres.

Energetically favourable water contacts are predicted in the region surrounding the catalytic Asp25A and Asp25B, at the far end of the S2 and S2' pockets and next to Gly27A and Gly27B. The position of the additionally enclosed water molecule is identified as preferred position for a water molecule only in case of PR_{I84V} (Figure 9a-c). The positions of the experimentally observed waters are in good agreement with the calculated interaction fields.

4.3 Discussion

The intention of this study was to elucidate the potential of a new class of HIV protease inhibitors against crucial point mutations. The unique mode in addressing the enzyme's subpockets observed for the C_2-symmetric pyrrolidines suggests deviating susceptibility towards amino acid exchanges at the S1/S2' and S1'/S2 pocket interfaces. The studied variants Ile50Val and Ile84Val are major mutations, both expanding the protease's binding pocket. Kinetic characterization reveals significant affinity decrease towards all four applied fluorogenic substrates with respect to the mutants. Similar observations are made for the inhibitors in case of the Ile50Val mutation, however, they were less pronounced as compared to the substrates. In contrast, the inhibitors retain or even increase affinity towards the Ile84Val mutant. Obviously, removal of the methylene groups drew a different response in terms of binding affinity.

Both inhibitors **8** and **9** are almost completely buried in all complexes due to the tunnel-shaped binding cavity of the protease. The mutations at the pivotal positions do not influence the burial of the ligands. With respect to the Ile50Val complexes high similarity with the wild-type counterparts is observed. The side-chains of Ile50B form presumably unfavourable close contacts to the sulfonyl oxygen atoms in PR_{WT} and PR_{I50V}. In the wild-type protease, C_{δ} of Ile50A is in Van der Waals contact to carbon atoms of the phenyl moiety of **8** and **9**. In contrast, in PR_{I50V}, the isoleucine exchange results in a loss of these favourable interactions, thus explaining the drop in inhibitor affinity. In case of Amprenavir, which bears a similar P2’ moiety, the susceptibility towards the Ile50Val mutation, is also attributed to these interactions.[26]

Larger differences are observed in case of the PR_{I84V} complexes. Ligand and protein conformation, along with the formed polar interactions, is highly conserved. In PR_{I84V}, C_{δ} of Ile50B forms Van der Waals contacts to the phenyl moiety of the inhibitors. This contact is not observed in the wild-type. Regarding the PR_{WT} complexes, the C_{δ} atoms of Ile84A and Ile84B have unfavourable short Van der Waals contacts with ligand atoms. Through exchange of these residues to valine, these unfavourable contacts are relaxed. The overall hydrophobic surface present in PR_{WT} and PR_{I50V} changes significantly, leading to a stronger exposure of the polar surfaces of Thr80 and Val82. In the wild-type complexes, the hydroxyl group of Thr80 is generally involved in a hydrogen bond to the main-chain carbonyl of Val82. As a result of the created additional space in PR_{I84V}, a water molecule is picked up and participates in polar interactions, forming a hydrogen-bond network connecting Thr80A and Val82A with one sulfonyl oxygen atom of the ligand. In PR_{WT} complexes, this oxygen atom is not involved in any polar interactions. At the symmetry-related B-domain, intercalation of a water molecule is not observed in the mutant, likely due to the shorter distance of the sulfonyl oxygen atom to the polar groups. Considering the results of a GRID analysis, the intercalation of a water molecule seems to be highly favourable in the A-domain. The additional hydrogen bonds to Thr80A and Val82A and the loss of short Van der Waals contacts to the residue in position 84A induce a slightly different loop conformation of the A-chain residues 79-84 of PR_{I84V}. This is reflected by an increased rmsd of this region to PR_{WT}: 0.56 Å in case of **8** and

0.54 Å in case of **9**. This reorganization leads to a better mutual surface match of the ligand in the S1 pocket.

In summary these contributions add up to an improved affinity of this class of ligands towards the Ile84Val mutant. The amount of each individual contribution is difficult to assign and will depend on slight differences in the shape and accordingly accommodated binding mode. Clearly, the individual contributions are not additive as can be seen when compounds **8, 9** and **10** are compared. Inhibitors **8** versus **9** and **10** are equipped with different polar substituents (NH_2 versus $CONH_2$). They induce slightly different arrangements in the corresponding S2/S2'-pockets. Via the rigid central scaffold, this discrepancy is translated into small orientational differences of the ligands in the S1/S1'-pockets. They result in shorter Van der Waals contacts to Ile84 for **9** compared to **8**. Mutating this residue to the smaller valine results in a relief of steric strain and provokes the observed increase in affinity discrimination between PR_{I84V} and PR_{WT}.

4.4 Summary and Conclusion

The investigated pyrrolidine-based inhibitors provide a novel strategy to target the subpockets of HIV protease. This leads to a different occupation of the binding pocket, especially next to the border between the S1/S2' and S1'/S2-pockets. The properties of the new inhibitor class have been examined towards the active-site mutations Ile50Val and Ile84Val. Kinetic data indicate a contrary behaviour: Reduced affinity towards PR_{I50V} and improved potency towards PR_{I84V}. The determination of cocrystal structures of two representative inhibitors (**8** and **9**) with the wild-type and the two variants provides the structural basis to explain this observation. The reduced affinity towards PR_{I50V} could be rationalized in terms of reduced Van der Waals contacts, an observation already described in literature for Amprenavir. In case of PR_{I84V}, multiple contributions are responsible: Amplified Van der Waals contacts, enhanced polar interactions mediated by an intercalating water molecule that fills the gap created by the mutational exchange, and relief of steric strain likely given in the wild-type complexes. The Ile84Val mutation is discussed in literature mainly as an expansion of the active-site volume, lowering beneficial Van der Waals contacts to a bound ligand. Here, we can trace the exchange in more detail and make changes in the exposed physicochemical properties responsible for the modified interaction profile. The investigated pyrrolidines are very rigid ligands. They cannot optimally adapt their shape to the PR_{WT} binding pocket. Surface complementarity of the PR_{I84V} binding pocket is more appropriate to accommodate an inhibitor compared to the wild-type, exhibiting in best case, a K_i = 12 nM for inhibitor **10**.

The message to be learned from this study concerns the change in physicochemical properties of the binding pocket upon Ile84Val mutation that can be exploited for inhibitor design. Taking advantage of these features might avoid the susceptibility towards this mutation, which has proven to be problematic for all presently approved inhibitors. Furthermore, this study unambiguously shows that large affinity changes, e.g. a factor of 40 in case of **9,** might result from very small deviations in the protein-ligand complexes. It also points to the limitations in modeling such differences. Only the detailed analysis of multiple complexes in terms of structure-activity relationships provides deeper insights into the various contributions responsible for protein-ligand interactions.

4.5 Experimental Section

Inhibitor preparation: The Inhibitors were synthesized enantioselectively from D-(-)-tartaric acid as recently reported.[12] Stock solutions of inhibitors for kinetic measurements and cocrystallization were obtained by dissolving them in DMSO.

Mutant preparation, purification and crystallization: Mutants were prepared using the polymerase chain reaction with the Quick Change Site-directed Mutagenesis Kit (Stratagene).[29] The pET11a plasmid containing the HIV-1 protease gene was kindly provided by Professor Helena Danielson, University of Uppsala. Synthetic oligonucleotides containing the desired mutation were used as mutagenic primers, 5'- CCAAAAATGATA-GGGGGAGTTGGAGGTTTTATCAAAGTAAG-3' and 5'- CTTACTTTGATAAAACCTC-CAACTCCCCCTATCATTTTTGG-3' for I50V; 5'- GGACCTACACCTGTCAACGTAAT-TGGAAGAAATCTGTTG-3' and 5'- CAACAGATTTCTTCCAATTACGTTGACAGGTGT-AGGTCC-3' for I84V. All constructs were verified by DNA sequencing. The PR_{WT}, PR_{I50V},

PR_{I84V} HIV protease variants were consecutively expressed from *Escherichia coli* and purified as previously described.[30]

The HIV protease inhibitor complexes were crystallized at 18°C by the sitting drop vapour diffusion method using a 1:1 ratio. Crystals were obtained by cocrystallization of the enzyme with inhibitor concentrations ranging from 50- to 100-fold the K_i value and a final DMSO concentration of 5%. 1 µL of the well buffer (0.1 M BisTris, 2.5-3.0 M NaCl, pH 6.5) was mixed with 1 µL protein solution (50 mM NaAc, 1 mM EDTA, 1 mM DTT, pH 6.5) with a HIV protease concentration of 7-10 mg/mL. Crystals were obtained within 1-2 days and exhibited a rectangular shape. Crystals were further optimized using streak-seeding techniques. For cryoprotection the crystals were briefly soaked in mother liquor containing 25% glycerol.

Kinetic Assay: Enzymatic assays were performed in 172 µL assay buffer (100mM MES, 300 mM KCl, 5 mM EDTA, 1 mg/mL BSA, pH 5.5) by the addition of substrate dissolved in 4 µL DMSO, distinct inhibitor concentrations dissolved in 4 µL DMSO and 20 µL HIV-1 protease in assay buffer to a final volume of 200 µL (final DMSO concentration 4%). The fluorogenic anthranilyl-HIV protease substrates **A** (Abz-Thr-Ile-Nle-(*p*-NO_2-Phe)-Gln-Arg-NH_2) and **B** (Abz-Lys-Ala-Arg-Val-Nle-(*p*-NO_2-Phe)-Phe-Glu-Ala-Nle-NH_2) were purchased from Bachem and **C** (Abz-Ala-Arg-Val-Nle-(*p*-NO_2-Phe)-Phe-Glu-Ala-Nle-NH_2) and **D** (Abz-Arg-Val-Nle-(*p*-NO_2-Phe)-Phe-Glu-Ala-Nle-NH_2) from Sigma Aldrich. The kinetic constants for PR_{WT}, PR_{I50V} and PR_{I84V} were determined by the method of Lineweaver and Burk. The total enzyme concentration was quantified by titrating with the strong binding inhibitor Saquinavir at a 200-fold higher enzyme concentration. Inhibition data for PR_{WT}, PR_{I50V} and PR_{I84V} were determined as follows: IC_{50} values were taken from plots of v_i/v_0 versus inhibitor concentration, in which v_i is the velocity in presence, and v_0 the velocity in the absence of an inhibitor. The hydrolysis of the substrate **A** was recorded as the increase in fluorescence intensity (excitation wavelength 337 nm, emission wavelength 410 nm) over a time period of 10 min during which the signal increased linearly with time.[31] K_i values were calculated from the following equation: $K_i = [IC_{50} - (E_t/2)][1 + (S/K_m)]^{-1}$

Data collection, phasing and refinement: The data sets were collected at the synchrotron BESSY II in Berlin/Germany on PSF beamline 14.2 (**8** PR_{I50V}, PDB ID: 3R43), equipped with a MAR-CCD detector and on a Rigaku R-AXIS IV image plate detector using Cu K_α radiation from an in-house Rigaku rotating anode (**8** PR_{I84V}, **9** PR_{I50V}, **9** PR_{I84V}, PDB ID:

2R3W, 2R3T, 2R38). Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000.[32] The structures were determined by the molecular replacement method with Phaser,[33] one monomer of the 1.50 Å HIV-1 protease in complex with a pyrrolidine based inhibitor (PDB ID: 2PQZ) was used as the search model. The structure refinement was continued with SHELXL-97,[34] for each refinement step at least 10 cycles of conjugate minimization were performed, with restraints on bond distances, angles and B-values. Intermittent cycles of model building were done with the program COOT.[35] The coordinates have been deposited in the PDB (http://www.rcsb.org/pdb/) with access codes: (**8** PR_{I50V} 2R43, **8** PR_{I84V} 2R3W, **9** PR_{I50V} 2R3T, **9** PR_{I84V} 2R38).

GRID Analysis and surface calculations: The calculations of the molecular interaction fields were performed with GREATER, a graphical user interface for the program GRID, version 22.[36] The interaction box was defined to enclose the entire active sites. The grid spacing was = 0.33 Å (NPLA = 3) and the maximum positive cut-off energy was set to 8 kcal/mol and visualized with Pymol. The buried surface area was calculated with the program MS developed by Connolly.[37] The program calculates the molecular surface from atom coordinates; the probe radius was set to 1.4 Å.

4.6 References

1. Wlodawer, A. & Erickson, J. W. (1993). Structure-Based Inhibitors of HIV-1 Protease. *Annu. Rev. Biochem.* **62**, 543-585.
2. Kramer, R. A., Schaber, M. D., Skalka, A. M., Ganguly, K., Wong-Staal, F. & Reddy, E. P. (1986). HTLV-III gag protein is processed in yeast cells by the virus pol-protease. *Science* **231**, 1580-1584.
3. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A. F., Scolnick, E. M. & Sigal, I. S. (1988). Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Nat. Acad. Sci. U.S.A.* **85**, 4686-4690.
4. Flexner, C. (2007). HIV drug development: the next 25 years. *Nat. Rev. Drug Discov.* **6**, 959-966.
5. Bartlett, J. A., DeMasi, R., Quinn, J., Moxham, C. & Rousseau, F. (2001). Overview of the effectiveness of triple combination therapy in antiretroviral-naive HIV-1 infected adults. *AIDS* **15**, 1369-1377.

6. Gulick, R. M., Mellors, J. W., Havlir, D., Eron, J. J., Meibohm, A., Condra, J. H., Valentine, F. T., McMahon, D., Gonzalez, C., Jonas, L., Emini, E. A., Chodakewitz, J. A., Isaacs, R. & Richman, D. D. (2000). 3-Year Suppression of HIV Viremia with Indinavir, Zidovudine, and Lamivudine. *Ann. Intern. Med.* **133**, 35-39.
7. Coffin, J. M. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* **267**, 483-489.
8. Mendoza, C. d. & Vincent, S. (2004). Resistance to HIV Protease Inhibitors: Mechanisms and Clinical Consequences. *Curr. Drug Metab.* **5**, 321-328.
9. Ohtaka, H., Muzammil, S., Schön, A., Velazquez-Campoy, A., Vega, S. & Freire, E. (2004). Thermodynamic rules for the design of high affinity HIV-1 protease inhibitors with adaptability to mutations and high selectivity towards unwanted targets. *Int. J. Biochem. Cell Biol.* **36**, 1787-1799.
10. D’Aquila, R. T., Schapiro, J. M., Brun-Vézinet, F., Clotet, B., Conway, B., Demeter, L. M., Grant, R. M., Johnson, V. A., Kuritzkes, D. R., Loveday, C., Shafer, R. W. & Richman, D. D. (2002). Drug Resistance Mutations in HIV-1. *Top. HIV Med.* **10**, 21-25.
11. Specker, E., Böttcher, J., Lilie, H., Heine, A., Schoop, A., Müller, G., Griebenow, N. & Klebe, G. (2005). An Old Target Revisited: Two New Privileged Skeletons and an Unexpected Binding Mode For HIV-Protease Inhibitors. *Angew. Chem. Int. Ed.* **44**, 3140-3144.
12. Blum, A., Böttcher, J., Heine, A., Klebe, G. & Diederich, W. E. (2008). Structure guided Design of C_2-symmetric HIV-1 Protease Inhibitors based on a Pyrrolidine Scaffold. *J. Med. Chem.* **51**, 2078-2087
13. Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., Green, B. E., Fino, L., Park, C. H., Kong, X., Wideburg, N. E., Saldivar, A., Ruiz, L., Kati, W. M., Sham, H. L., Robins, T., Stewart, K. D., Hsu, A., Plattner, J. J., Leonard, J. M. & Norbeck, D. W. (1995). ABT-538 is a Potent Inhibitor of Human Immunodeficiency Virus Protease and has High Oral Bioavailability in Humans. *Proc. Nat. Acad. Sci. U.S.A.* **92**, 2484-2488.
14. Clemente, J. C., Coman, R. M., Thiaville, M. M., Janka, L. K., Jeung, J. A., Nukoolkarn, S., Govindasamy, L., Agbandje-McKenna, M., McKenna, R., Leelamanit, W., Goodenow, M. M. & Dunn, B. M. (2006). Analysis of HIV-1 CRF_01 A/E Protease Inhibitor Resistance: Structural Determinants for Maintaining Sensitivity and Developing Resistance to Atazanavir. *Biochemistry* **45**, 5468-5477.

15. Surleraux, D. L. N. G., Tahri, A., Verschueren, W. G., Pille, G. M. E., de Kock, H. A., Jonckers, T. H. M., Peeters, A., De Meyer, S., Azijn, H., Pauwels, R., de Bethune, M., King, N. M., Prabu-Jeyabalan, M., Schiffer, C. A. & Wigerinck, P. B. T. P. (2005). Discovery and Selection of TMC114, a Next Generation HIV-1 Protease Inhibitor. *J. Med. Chem.* **48**, 1813-1822.
16. Kim, E. E., Baker, C. T., Dwyer, M. D., Murcko, M. A., Rao, B. G., Tung, R. D. & Navia, M. A. (1995). Crystal Structure of HIV-1 Protease in Complex with VX-478, a Potent and Orally Bioavailable Inhibitor of the Enzyme. *J. Am. Chem. Soc.* **117**, 1181-1182.
17. Chen, Z., Li, Y., Chen, E., Hall, D. L., Darke, P. L., Culberson, C., Shafer, J. A. & Kuo, L. C. (1994). Crystal Structure at 1.9-Å Resolution of Human Immunodeficiency Virus (HIV) II Protease Complexed with L-735,524, an Orally Bioavailable Inhibitor of the HIV Proteases. *J. Biol. Chem.* **269**, 26344-26348.
18. Kaldor, S. W., Kalish, V. J., Davies, J. F., Shetty, B. V., Fritz, J. E., Appelt, K., Burgess, J. A., Campanale, K. M., Chirgadze, N. Y., Clawson, D. K., Dressman, B. A., Hatch, S. D., Khalil, D. A., Kosa, M. B., Lubbehusen, P. P., Muesing, M. A., Patick, A. K., Reich, S. H., Su, K. S. & Tatlock, J. H. (1997). Viracept (Nelfinavir Mesylate, AG1343): A Potent, Orally Bioavailable Inhibitor of HIV-1 Protease. *J. Med. Chem.* **40**, 3979-3985.
19. Krohn, A., Redshaw, S., Ritchie, J. C., Graves, B. J. & Hatada, M. H. (1991). Novel Binding Mode of Highly Potent HIV-Proteinase Inhibitors Incorporating the (*R*)-Hydroxyethylamine Isostere. *J. Med. Chem.* **34**, 3340-3342.
20. Muzammil, S., Armstrong, A. A., Kang, L. W., Jakalian, A., Bonneau, P. R., Schmelmer, V., Amzel, L. M. & Freire, E. (2007). Unique Thermodynamic Response of Tipranavir to Human Immunodeficiency Virus Type 1 Protease Drug Resistance Mutations. *J. Virol.* **81**, 5144-5154.
21. Tisdale, M., Myers, R., Randall, S., Maguire, M., Ait-Khaled, M., Elston, R. & Snowden, W. (2000). Resistance to the HIV Protease Inhibitor Amprenavir In Vitro and in Clinical Studies: A Review. *Clin. Drug Invest.* **20**, 267-285.
22. Pazhanisamy, S., Stuver, C. M., Cullinan, A. B., Margolin, N., Rao, B. G. & Livingston, D. J. (1996). Kinetic Characterization of Human Immunodeficiency Virus Type-1 Protease-resistant Variants. *J. Biol. Chem.* **271**, 17979-17985.

23. Klabe, R. M., Bacheler, L. T., Ala, P. J., Erickson-Viitanen, S. & Meek, J. L. (1998). Resistance to HIV Protease Inhibitors: A Comparison of Enzyme Inhibition and Antiviral Potency. *Biochemistry* **37**, 8735-8742.
24. Tie, Y., Kovalevsky, A. Y., Boross, P., Wang, Y., Ghosh, A. K., Tozser, J., Harrison, R. W. & Weber, I. T. (2007). Atomic Resolution Crystal Structures of HIV-1 Protease and Mutants V82A and I84V with Saquinavir. *Proteins* **67**, 232–242.
25. Wang, Y., Tie, Y., Boross, P. I., Tozser, J., Ghosh, A. K., Harrison, R. W. & Weber, I. T. (2007). Potent New Antiviral Compound Shows Similar Inhibition and Structural Interactions with Drug Resistant Mutants and Wild Type HIV-1 Protease. *J. Med. Chem.* **50**, 4509-4515.
26. Ohtaka, H., Velázquez-Campoy, A., Xie, D. & Freire, E. (2002). Overcoming drug resistance in HIV-1 chemotherapy: The binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease. *Protein Sci.* **11**, 1908-1916.
27. Todd, M. J., Luque, I., Velázquez-Campoy, A. & Freire, E. (2000). Thermodynamic Basis of Resistance to HIV-1 Protease Inhibition: Calorimetric Analysis of the V82F/I84V Active Site Resistant Mutant. *Biochemistry* **39**, 11876-11883.
28. Wang, W. & Kollman, P. A. (2001). Computational study of protein specificity: The molecular basis of HIV-1 protease drug resistance. *Proc. Nat. Acad. Sci. U.S.A.* **98**, 14937-14942.
29. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.
30. Taylor, A., Brown, D. P., Kadam, S., Maus, M., Kohlbrenner, W. E., Weigl, D., Turon, M. C. & Katz, L. (1992). High-level expression and purification of mature HIV-1 protease in *Escherichia coli* under control of the *araBAD* promoter. *Appl. Microbiol. Biotechnol.* **37**, 205-210.
31. Toth, M. V. & Marshall, G. R. (1990). A simple, continuous fluorometric assay for HIV protease. *Int. J. Pept. Protein Res.* **36**, 544-550.
32. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods Enzymol.* (Carter, Jr. C.W. & Sweet R. M.), Vol. 276, pp. 307-326. Academic Press.
33. Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **60**, 432-438.

34. Sheldrick, G. M. & Schneider, T. R. (1997). *SHELXL*: High-resolution refinement. In *Methods Enzymol.* (Carter, Jr. C.W. & Sweet R. M.), Vol. 277, pp. 319-343. Academic Press.

35. Emsley, P. & Cowtan, K. (2004). *Coot*: model-building tools for molecular graphics. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **60**, 2126-2132.

36. Goodford, P. J. (1985). A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **28**, 849-857.

37. Connolly, M. L. (1983). Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**, 548-558.

5. Two Solutions for the same Problem: Multiple Binding Modes of Pyrrolidine-based HIV-1 Protease Inhibitors *

5.1 Introduction

The first successful application of structure-based drug design (SBDD) was reported in 1976 by *Beddell et al.* who utilized the three dimensional structure of hemoglobin for ligand development.[1] Since then the number of protein structures publicly available in the protein data bank (PDB) has increased exponentially resulting in more than 50000 entries nowadays.[2] Due to the steadily growing amount of available protein structures and methodological advances in structural analysis, SBDD has meanwhile become an integral part of drug discovery projects.[3]

The analysis of a protein-ligand complex gives detailed information about the protein-ligand interactions. Once an initial lead has been identified and the target protein is accessible for structural analysis it can further be optimized using iterative cycles comprising synthesis, biological evaluation, computational methods and structural analysis.[4] The structural analysis using X-ray crystallography usually suggests the existence of one single well-defined state, anticipated as a unique binding mode, which is then utilized to propose the most promising modifications of the initial scaffold to further optimize a given lead. Numerous successful projects of SBDD have been published which led at best to the development of an approved drug, as for example in case of Tipranavir.[5]

* Taken from a manuscript in preparation, Andreas Blum, Jark Böttcher, Stefanie Dörr, Andreas Heine, Gerhard Klebe, Wibke E. Diederich. Two solutions for the same problem: multiple binding modes of Pyrrolidine-based HIV-1 Protease Inhibitors.

However, the iterative process can face a lot of obstacles as well as surprises and numerous examples of a non-unique behavior within a series of similar ligands have been described of which only the three following should be mentioned as representative examples: A reversed binding mode was detected by X-ray crystallography in due course SBDD process of non-peptidic HIV protease inhibitors. In the reported case an alkylation of the inhibitor's amide functionaleties resulted in a completely different occupation of the individual specificity sub-pockets.[6] The small energy deviations of such opposing binding modes have been investigated in detail, e.g. by the pH-dependency of the binding mode observed in case of trypsin inhibitor complexes.[7] An extraordinary observation was made in case of Protein Kinase C: the inhibitor binds in different orientations to the two monomers present in the asymmetric unit.[8] Observations like these mentioned above suggest that the occurrence of several binding orientations within a series of similar compounds or even for one certain ligand is not a rare case. Such phenomena can be indicated e.g. from a complex SAR, which cannot be explained using the broadly accepted hypothesis that similar ligands bind in a similar fashion.[9]

Recently we reported our design and synthesis of C_2-symmetric 3,4-disubstituted pyrrolidines as a new class of HIV-1 protease inhibitors.[10] Our structure guided optimization was based on the initially observed cocrystal structure of the *N*-benzyl substituted inhibitor **1** (PDB code: 2PQZ, Figure 1 and Table 1).

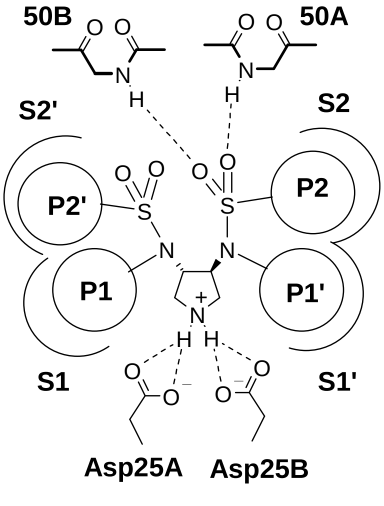


Figure 1: Schematic representation of the conserved binding mode of the pyrrolidine-3,4-bis-*N*-benzyl-sulfonamides. Hydrogen bonds are indicted by dashed and main-chain bonds using bold lines.

In due course of the project additional cocrystal structures of the further developed inhibitors were determined to elucidate their interactions explaining a fully consistent SAR.

The conserved binding mode observed in the five crystal structures is depicted in Figure 1. All complexes between the topologically C_2-symmetric inhibitors and the sequentially C_2-symmetric protein adopt an unsymmetrical binding mode with respect to the ligand. However, each sub-pocket related by an imposed C_2-symmetry is occupied by substituents at the core skeleton of the inhibitor that mutually correspond by topological symmetry. The benzyl moieties occupy the S1- and S1'- subpockets whereas the sulfonamide side-chains of the inhibitor address the S2- and S2'-subpockets. The *N*-benzyl substituted derivatives were additionally analyzed with respect to their susceptibility towards the active site mutants Ile50Val (PR_{I50V}) and Ile84Val (PR_{I84V}), and revealed promising properties towards the Ile84Val mutation located at the borders of the S1/S2'- and S2/S1'-subpockets, respectively. Particularly, compounds bearing polar substituents in the S2/S2'-pockets showed significantly improved affinity towards the mutation of isoleucine by valine (e.g. compound **2**).[11] This observation was further analyzed by X-ray crystallography of the corresponding complexes of two derivatives together with the protease variants and could finally be explained by amplified Van der Waals (VdW) contacts, enhanced polar interactions mediated by an interstitial water molecule, and relief of steric strain likely given in the wild-type complexes.

Interestingly enough, the determination of the binding affinities towards the PR_{I84V} and PR_{I50V} mutants of inhibitors bearing smaller *N*-alkyl substituents (**3**-**5**) revealed a different selectivity profile (Table 1). In order to further elucidate this peculiar difference in the SAR, a new small series of inhibitors was designed. Inhibitor **5,** comprising the *N*-dimethylallyl substituents, which showed the best potency of the series, was picked as a reference. Its properties were further varied by decorating the sulfonamide benzyl group with *para*-substituents of different physicochemical properties, such as polar carboxylate and amide groups and non-polar bromine and cyano groups. Based on the previously observed binding-modes, this moiety was thought to occupy the S2- and S2'-pockets.

5.2 Results and Discussion

5.2.1 Chemistry

The inhibitor scaffold is easily accessible in enantiomerically pure form from *D*-(-)-tartaric acid and inhibitors **1-5** have been synthesized as previously described.[10] 3,4-Diaminopyrrolidine **10** reacts with *para*-bromo-phenylsulfonyl chloride. It renders the corresponding sulfonamide which is then alkylated using 3,3-dimethylallyl bromide thus leading to the key intermediate **11** (Figure 2).

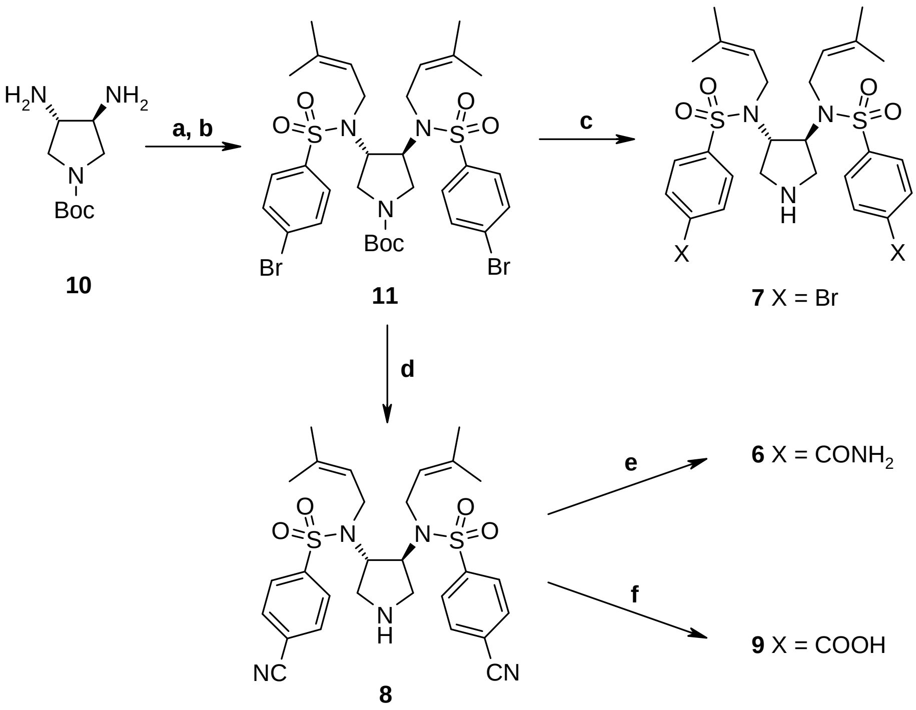


Figure 2: Preparation of the inhibitors, the BOC-protected 3,4-diamino-pyrrolidine was prepared from D-(-)-tartaric acid as previously described.[10] **a**) *p*-Br-$PhSO_2Cl$, NEt_3, CH_2Cl_2: **b**) 3,3-dimethylallyl bromide, K_2CO_3, CH_3CN, 0° to RT.: 92%; **c**) CH_2Cl_2:TFA 1:1: 74%; **d**) $Zn(CN)_2$, $Pd(PPh_3)_4$, DMF, µW: 51%; **e**) 30% H_2O_2, K_2CO_3, DMSO: 38%; f) NaOH, H_2O, µW: 46%.

Deprotection of **11** under acidic non-aqueous conditions renders **7**. Substitution of the *para*-bromo substituent using $Zn(CN)_2$ as cyanide source using Pd-catalysis under microwave irradiation required prolonged reaction times and high temperatures, which resulted in a concomitant cleavage of the BOC-protecting group and thus directly furnished inhibitor **8**.[12] Intermediate **8** was further utilized for the preparation of the inhibitors **6** and **9**. Saponification of the nitrile functionality in **8** under basic conditions and microwave yielded the corresponding acid **9**. Mild hydrolysis with 30% hydrogen peroxide in DMSO under catalysis of K_2CO_3 gave rise to **6**.[13] Via the microwave-assisted approach the differently substituted pyrrolidines **6-9** are easily accessible in only 1-2 steps utilizing **11** as key intermediate in moderate overall yields. After purification with RP-MPLC, the inhibitors were obtained as triflouroacetate salts and tested for their affinity against the different HIV proteases.

5.2.2 Biological evaluation

Determination of the affinities towards the protease variants gave rise to a structure-activity relationship (SAR) which surprisingly showed no clear preference for one type of substituents (Table 1).

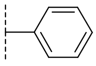
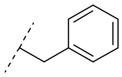
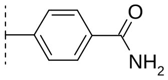
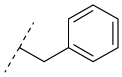
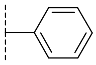
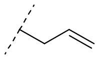
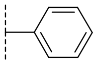
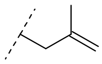
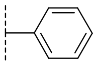
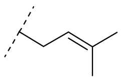
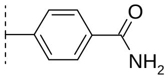
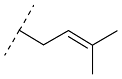
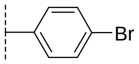
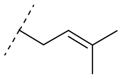
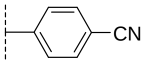
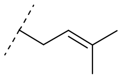
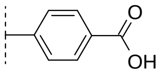
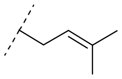
			PR$_{WT}$	**PR$_{I50V}$**	**PR$_{I84V}$**
.	P2/P2’	P1/P1’	K_i [µM]	K_i [µM]	K_i [µM]
1			2.2	11	1.1
2			0.26	1.4	0.036
3			12	n.i.*	84
4			75	340	53
5			1.6	10	5.8
6			0.48	13	1.9
7			0.69	2.0	0.91
8			2.3	7.3	4.9
9			3.3	11	2.9

Table 1: K_i-values of the inhibitors towards the HIV protease. * n.i. = Ki > 500µM

Whereas **6** bearing a polar amide functionality and **7** bearing a hydrophobic bromide substituent gain similarly in affinity towards PR_{WT} and PR_{I84V}, substitution with a carboxylate functionality **9** or the corresponding nitrile moiety **8** did not lead to an improvement in any of the cases. In comparison to PR_{WT} and PR_{I84V} binding, all compounds within this series reveal a reduced affinity towards PR_{I50V}, similar to the previously studied series of *N*-benzyl substituted derivatives. In contrast to the *N*-benzyl substituted inhibitors no significant improvement in affinity towards PR_{I84V} is observed for all the inhibitors comprising 3,3-dimethylallyl substituents. The substitution with polar substituents in *para*-position of the sulfonyl-aryl moieties improves the binding affinities towards this mutation in case of the *N*-benzyl derivatives. However, in case of the dimethylallyl substituted inhibitors, particularly for the most potent derivative **6,** no change in the relative affinity difference can be observed, when compared to the unsubstuitueted counterpart **5**. Compared with the corresponding benzyl derivative **2** which showed a seven-fold improvement with respect to PR_{I84V}, the now four fold decrease can hardly be explained regarding only the steric demand of the different *N*-alkyl moieties. To rationalize these observations in structural terms and in order to compare with the known binding mode of **2**, we crystallized compound **6** in complex with the wildtype HIV-protease and consecutively determined its crystal structure.

5.2.3 Structural Analysis

Crystals of the PR_{WT} in complex with **6** were obtained by cocrystallization of the enzyme with an inhibitor concentration of 100 µM and final DMSO concentration of 10% following our standard protocol. Crystals grew within a week and showed deviating appearance, cubic shaped at a precipitant concentration of 3.5 M NaCl and needle-like shaped at a concentration of 3 M NaCl. Whereas the cubic crystals exhibited the space group $P2_12_12_1$ the needle-shaped crystal form corresponded to the space group $P6_122$. Both structures could be determined with a resolution of 1.65 Å and ligand geometries were clearly visible in the F_o-F_c omit-map at a sigma level of 2 (Figure 3a and 3b). However, two completely different binding modes were observed and will be described separately: The geometry in space group $P2_12_12_1$ is referred to as orthorhombic binding mode, whereas the binding mode in $P6_122$ is named hexagonal binding mode.

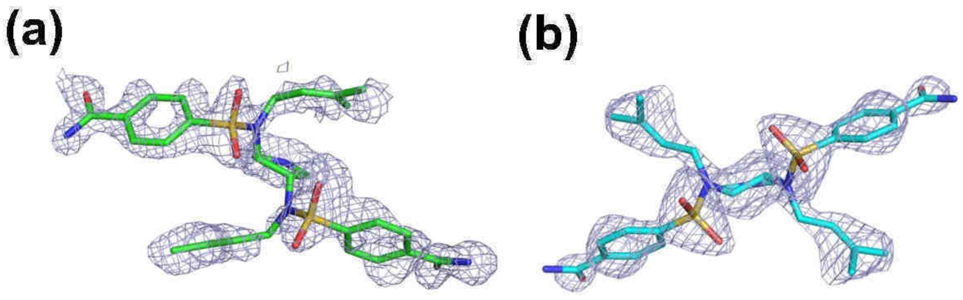


Figure 3: The F_o-F_c densities for the ligands is displayed at a σ level of 2.0 as blue mesh. (a) space group $P2_12_12_1$, (green, color-coded by atom type) (b) space group $P6_122$ (blue, color-coded by atom type).

	$P2_12_12_1$	$P6_122$
resolution (Å)	25-1.65	25-1.65
space group	$P2_12_12_1$	$P6_122$
cell dimensions (Å)	a = 52.0	a = 62.5
	b = 57.6	b = 62.5
	c = 61.4	c = 82.3
highest resolution shell (Å)	1.68-1.65	1.68-1.65
no. of measured reflections	99069	80672
no. of independent reflections	22526	11774
completeness (%)[a]	98.8 [89.4]	97.8 [96.2]
I/σ[a]	26.1 [2.7]	23.2 [4.3]
R_{sym} (%)[a]	5.2 [35.0]	8.0 [36.8]
resolution in refinement (Å)	10-1.65	10-1.65
R_{cryst} ($F > 4\sigma F_o$; F_o)	18.3; 20.2	22.4; 23.4
R_{free}($F > 4\sigma F_o$; F_o)	23.9; 26.2	27.4; 29.7
mean B-factor ($Å^2$)	18.3; 17.3	30.5
(peptide chain A; B)		
main-chain ($Å^2$)	15.6; 15.5	27.5
side-chains ($Å^2$)	21.2; 19.4	33.9
ligand ($Å^2$)	32.4	38.2
Cl ($Å^2$)	19.2	-
Water ($Å^2$)	26.1	33.5
Ramachandran plot		
most favored geometry (%)	96.2	94.9
additionally allowed (%)	3.8	5.1

Table 4: Crystallographic data. (a) Values in brackets refer to the highest resolution shell.

5.2.3.1 Binding mode in the orthorhombic form

In the orthorhombic space group inhibitor **6** exhibits a binding mode closely related to the one previously observed for all our pyrrolidine-3,4-bis-*N*-benzyl-sulfonamides. The endocyclic amino functionality addresses the catalytic dyad and forms a hydrogen bond network with the aspartic acid residues 25A (2.8 Å / 2.8 Å) and 25B (2.8 Å / 2.9 Å). Flap interactions are established via hydrogen bonding of only one of the inhibitor's sulfonyl oxygen atoms to the main-chain NH of Ile50A (2.6 Å). The second sulfonyl oxygen atom of this sulfonamide group remains uninvolved in any polar interaction (Figure 4).

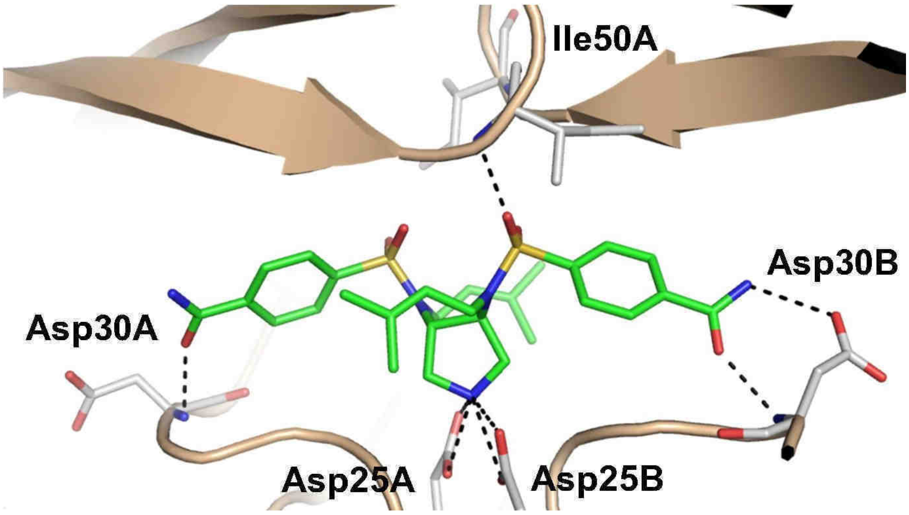


Figure 4: Crystal structure of **6** (green, color-coded by atom type) in complex with HIV protease ($P2_12_12_1$). The protein backbone trace is schematically illustrated in wheat, the catalytic Asp25A and Asp25B, Ile50A, Ile50B, Asp30A and Asp30B are displayed in grey color-coded by atom type. Hydrogen bonds are indicated by dashed lines.

The opposing sulfonyl group of **6** does not form any polar interactions. Also the subsite occupancy is comparable to the previously determined complexes. Whereas the dimethylallyl substituents occupy the S_1/S_1'-subpockets establishing contacts to Gly27A, Gly48A, Gly49A, Leu23B, Pro81B, Val82B, Ile84B (S1) and Leu23A, Gly27B, Ile84A (S1'), the S_2/S_2'-pockets are addressed by the phenyl moieties of **6** forming numerous vdW contacts with Ala28A, Val32A, Ile47A, Gly48A (S2); Ile50A, Ala28B, Gly48B, Ile84B (S2'). The amide functionalities form hydrogen bonds to the main-chain NH of Asp30A (2.6 Å) and Asp30B (2.7 Å), and additionally in the S2' pocket to the corresponding Asp30B side-chain (2.8 Å).

The C_α superposition of this structure with the crystal structure of the corresponding *N*-benzyl substituted inhibitor (PDB ID: 2PWR) reveals that the overall binding mode is only slightly affected by the exchange of the *N*-benzyl to *N*-dimethylallyl moiety, which is reflected by a root mean square deviation (rmsd) between the C_α atoms of the complexes of 0.53 and an rmsd of 0.60 Å for the identical ligand atoms of **2** and **6**, using this alignment. The ligand is deeply buried in the binding pocket (90%).

5.2.3.2 Binding mode in the hexagonal form

The binding mode observed in the hexagonal space group deviates dramatically from the previously discribed binding mode which is in agreement to all examples yet observed for our pyrrolidine-based HIV protease inhibitors. In the hexagonal case, the complex of the C_2-symmetric protease and the C_2-symmetric inhibitor **6** fully agrees to C_2-symmetry (Figure 5).

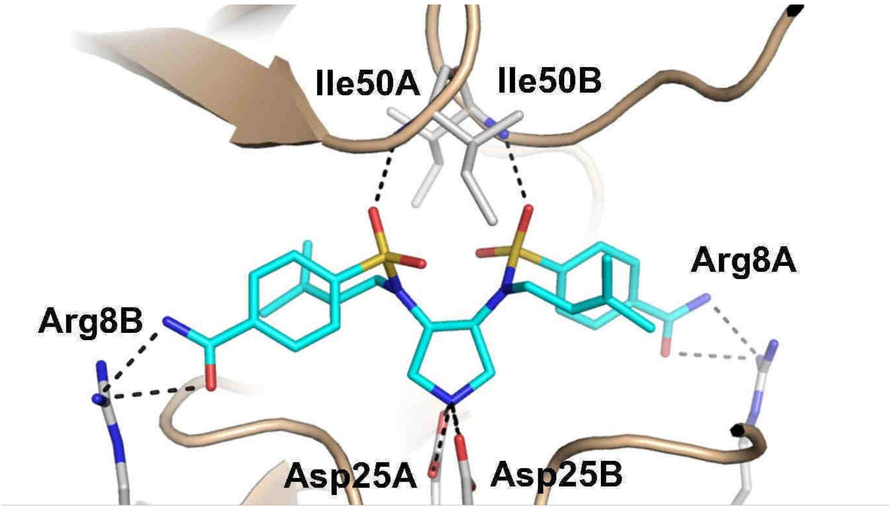


Figure 5: Crystal structure of **6** (light blue, color-coded by atom type) in complex with HIV protease ($P6_122$). The protein backbone trace is schematically illustrated in wheat, the catalytic Asp25A and Asp25B, Ile50A, Ile50B, Arg8A and Arg8B are displayed in grey color-coded by atom type. Hydrogen bonds are indicated by dashed lines.

In this space group, the asymmetric unit comprises one monomer of HIV-1 protease. The functional dimer is generated via a twofold crystallographic symmetry. For purpose of comparison amino acids of the protein chain A and their symmetry equivalents will be referred to as 1A to 99A and 1B to 99B, respectively. According to the imposed symmetry any ligand interactions to the A and B chain are identical. Similar to the orthorhombic case the catalytic aspartates Asp25A and Asp25B are addressed by the endocyclic pyrrolidine nitrogen of **6** (2.8 Å / 2.9 Å). However, this is the only conserved feature shared by the two binding modes. In contrast to the orthorhombic mode each sulfonyl group of **6** forms a direct hydrogen bond to Ile50A and Ile50B, respectively, with one sulfonyl oxygen atom of each sulfonamide group (3.0 Å). The *para*-carboxamido-benzenesulfonamide moieties occupy the S1 and S1'-pockets forming vdW interactions to Leu23, Gly27, Gly49, Pro81, Val82, Ile84 and polar interactions to Arg8 with the amide substituent (3.0 Å / 3.2 Å), of chain A and B, respectively. The S2 and S2'-pockets are now occupied by the dimethylallyl moieties establishing vdW contacts to Ala28, Val32, Ile47, and Ile50 of chain A and B. Comparison of the complex with the corresponding complex of the *N*-benzyl substituted inhibitor using a C_α superposition reveals high similarity of the protein structures reflected by a root mean square deviation of 0.60 Å, but high dissimilarity for the identical ligand atoms (rmsd = 5.4 Å) using this alignment. Similarly to the orthorhombic case, the ligand is deeply buried in the binding pocket (86%).

5.3 Discussion

Investigation of the inhibitory potency of a series of pyrrolidine-based inhibitors against mutant-HIV protease variants resulted in a remarkably different selectivity profile of the *N*-allyl substituted derivatives compared to the *N*-benzyl substituted initial lead structure. To elucidate the reasons for this opposing behavior especially against the Ile84Val mutant, a small series of inhibitors was synthesized based on the most potent derivative **5** of the initial series.

However, the non-uniform SAR obtained within this series, comprising substituents with different physicochemical properties in *para*-position of the inhibitor's sulfonamide aryl groups, did not match the activity pattern being expected from the binding mode of the *N*-benzyl derivatives. For a detailed analysis of this phenomenon and for comparison with the corresponding *N*-benzyl derivative **2**, **6** was selected for crystal structure determination in complex with HIV protease. To our surprise protein crystals were obtained in two crystal forms at very similar crystallization conditions. Both exhibited a space group deviating from the former series, in which 10 complexes of *N*-benzyl substituted inhibitors revealed identical crystal packing in $P2_12_12_1$.

The analysis of the complex in space group $P2_12_12_1$ revealed a binding mode closely resembling the one observed for all examples of the *N*-benzyl series. Due to the similarity of the surface complementarities observed for this orthorhombic binding mode with those previously determined for **2**, particular in the contact area to Ile84 (Figure 6a and 6c) a convincing explanation for the deviating selectivity profile is difficult to provide.

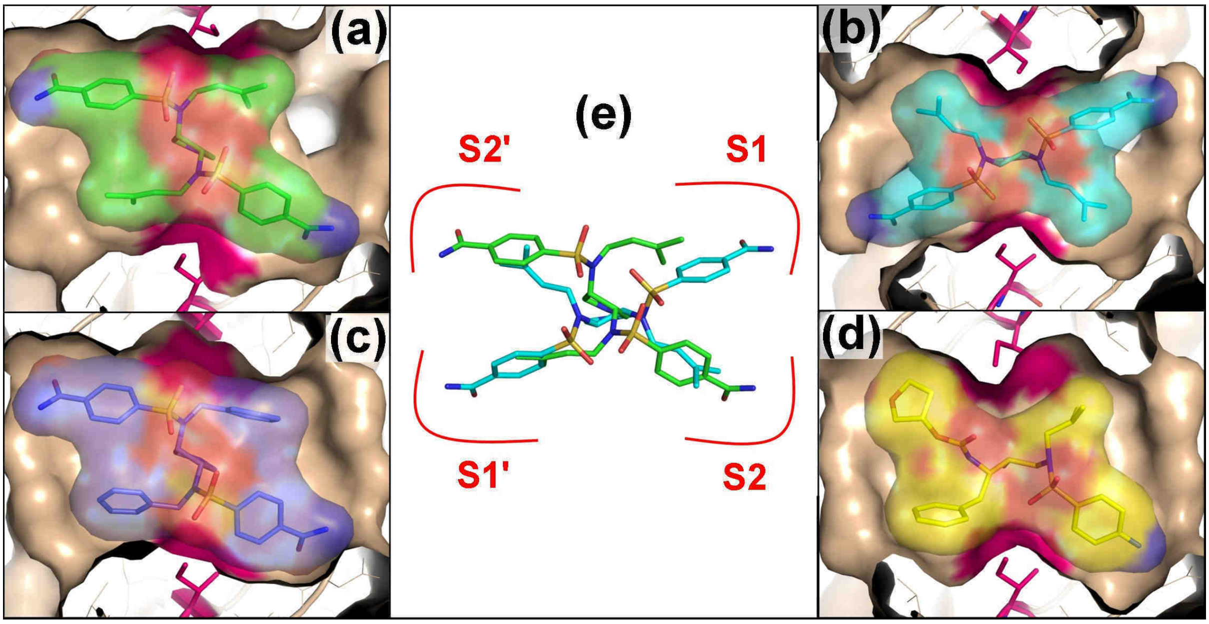


Figure 5: Crystallographically observed binding modes in HIV protease in equal orientation and size (a – d) Surface representation of the ligand (color-coded by atom type) and the corresponding protein structure (wheat), view from the top with the flap region clipped off. Ile84 is highlighted in hot pink (a) orthorhombic binding mode of **6** (b) hexagonal binding mode of **6** (c) binding mode of **2** (PDB ID: 2PWR) (d) binding mode of Amprenavir (pdb ID: 1HPV) (e) Superposition of the ligand orientations observed in the corresponding orthorhombic (green) and hexagonal (blue) crystal structures of **5** aligned by a C_α-fit of the protein coordinates**.**

However, the structural evidence found in the second hexagonal crystal form suggests that a different binding mode can be adopted. This binding mode reveals an inhibitor orientation much closer to that found for all approved peptidomimetic inhibitors, which are all known to exhibit susceptibility towards the I84V substitution (Figure 6b and 6d). The commonly accepted assumption that similar ligands bind in a similar fashion can definitely be ruled out in this case at least considering the crystalline state. Regarding the two different binding modes, it is not evident which orientation will be preferred under biological conditions. The nature of the substituent and the different binding-pocket shape of the mutant enzymes might induce different preferences, resulting in an overall non-uniform SAR. Whether the adopted space groups impose special symmetry constrains onto the deviating binding modes or whether the preferred binding mode reinforces the crystal packing cannot be resolved at this point. A comparison of protein-protein contacts due to crystal packing as well as the additional observation of two chloride ions present in the $P2_12_12_1$ complex, which exhibit at least a distance of 10Å to the closest ligand atom, however, do not satisfactorily explain why the binding mode observed in the orthorhombic case should not be possible in the hexagonal $P6_122$ space group and vice versa. Both binding modes seem to possess a very close by energy content, however, they obviously resolve spontaneously by crystallization.

5.4 Summary and Conclusion

Recently, we developed *N*-benzyl substituted HIV protease inhibitors and investigated in detail their behavior toward the active site mutants PR_{I50V} and PR_{I84V}. The observation of an overall affinity increase towards PR_{I84V} was, amongst others, attributed to relief of steric strain present in the wildtype complexes. The series **3-9** presented in this contribution was intended to elucidate whether smaller hydrophobic residues might be better suited to address the S1 and S1’ pockets appropriately. However, an inconsistent SAR demanded the determination of a cocrystal structure of at least one derivative bearing such modifications. Crystal structures of the most potent derivative **6** were determined based on two different crystal forms, found serendipitely. Remarkably they could be grown under very similar crystallization conditions exhibiting only a concentration difference in the precipitating NaCl concentration. Both crystal forms exhibited comparable diffraction quality and in each case a well defined however, distinct inhibitor orientation could be identified. To our surprise, the complexes of identical overall composition exhibit different protein-inhibitor interactions. Only the interactions to the catalytic dyad are similar along with a comparable burial of the ligands. The observation of two different binding modes for **6**, suggests that also other members of this series could occur with more than one binding mode thus explaining the non-consistent SAR within the series. Are the results obtained by this study a peculiar rarely occuring phenomenon or of fundamental importance also for other structure-based drug discovery projects? There might be some statistical evidence for the widely accepted oppinion that similar ligands bind in similar fashion, however, this is by no means a “truism”! The question remains how often a similar behavior as found in this study remains unnoticed in a congeneric ligand series. Thoroughly and exhaustively performed structural biology has to be integral part of drug discovery projects and has to escort the entire optimization process. The same holds for the biological evaluation, which has to be interpreted in very careful manner to identify possible changes of binding orientations during lead optimization. In fact, the observation of a different binding mode for members of a series might even provide unexpected opportunities to optimize ligands towards a particular binding profile. Furthermore, such ligands might exhibit remarkable advantage to escape resistance development. Likely both binding modes are close in energy. Mutational changes can affect and influence either of the two modes and change the energetic balance. However, an inhibitor capable to bind in two close-by alternative modes might swap from one to the other in order to escape resistance development. A similar concept has been followed during the development of TMC125-R165335 (etravirine) and TMC120-R147681 (dapivirine), two novel allosteric site inhibitors of the reverse transcriptase with remarkable resistance profile.[14] The authors describe such an advantage to overcome resistance development by the ability to bind the target enzyme in multiple conformations.

5.5 Experimental Section

Kinetic Assay: Enzymatic assays were performed in 172 µL assay buffer (100 mM MES, 300 mM KCl, 5 mM EDTA, 1 mg/mL BSA, pH 5.5) by the addition of substrate dissolved in 4 µL DMSO, distinct inhibitor concentrations dissolved in 4 µL DMSO and 20 µL HIV-1 protease in assay buffer to a final volume of 200 µL (final DMSO concentration 4%). The fluorogenic anthranilyl-HIV protease substrate (Abz-Thr-Ile-Nle-(*p*-NO_2-Phe)-Gln-Arg-NH_2) was purchased from Bachem. The hydrolysis of the fluorogenic was recorded as the increase in fluorescence intensity (excitation wavelength 337 nm, emission wavelength 410 nm).[15] The kinetic parameters of PR_{WT} (K_m = 14.6 µM), PR_{I50V} (K_m = 139 µM) and PR_{I84V} (K_m = 70 µM) were determined by the method of Lineweaver and Burk. IC_{50} values were generated by nonlinear regression analysis from plots of v_i/v_0 versus inhibitor concentration, in which v_i is the velocity in presence, and v_0 the velocity in the absence of an inhibitor. K_i values were calculated from the following equation: $K_i = IC_{50}/[1 + (S/K_m)]$

Crystallization of HIV-1 protease: Crystals of HIV-1 protease in complex with **6** were obtained by cocrystallization using the sitting drop vapor diffusion method. The well buffer (0.1 M BisTris, pH 6.5, 3-3.5 M NaCl) was mixed with 1µl protein solution (50 mM NaAc, pH 6.5, 1 mM EDTA, 1 mM DTT) with a HIV protease concentration of 7mg/mL. Crystals exhibited space group $P2_12_12_1$ at a precipitant concentration of 3.5 M NaCl and $P6_122$ at 3 M NaCl. For cryo-protection both crystal forms were briefly soaked in mother liquor containing 25% glycerol.

Data collection, phasing and refinement: The data sets were collected at the synchrotron BESSY II in Berlin/Germany on PSF beamline 14.2. Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000.[16] The structures were determined by the molecular replacement method with Phaser,[17] one monomer of the 1.5 Å structure of the HIV-1 protease in complex with a pyrrolidine-based inhibitor (PDB ID: 2PQZ) was used as the search model. Refinement was continued with SHELXL-97,[18] for each refinement step at least 10 cycles of conjugate gradient minimization were performed, with restraints on bond distances, angles and B-values. Intermittent cycles of model building were done with the program COOT.[19] The coordinates have been deposited in the PDB (http://www.rcsb.org/pdb/) with access codes: 3CKT (orthorhombic), 2ZGA (hexagonal).

5.6 References

1. Beddell, C. R.; Goodford, P. J.; Norrington, F. E.; Wilkinson, S.; Wootton, R., Compounds designed to fit a site of known structure in human haemoglobin. *Br. J. Pharmacol.* **1976**, 57, (2), 201–209.
2. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M., The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, 112, (3), 535-542.
3. Hubbard, R. E., 3D structure and the drug-discovery process. *Mol. BioSyst.* **2005**, 1, (5-6), 391-406.
4. Anderson, A. C., The Process of Structure-Based Drug Design. *Chem. Biol.* **2003**, 10, (9), 787-797.

5. Suvit Thaisrivongs, J. W. S., Structure-based discovery of tipranavir disodium (PNU-140690E): A potent, orally bioavailable, nonpeptidic HIV protease inhibitor. *J. Pept. Sci.* **1999**, 51, (1), 51-58.

6. Reich, S. H.; Melnick, M.; Davies Jf, II; Appelt, K.; Lewis, K. K.; Fuhry, M. A.; Pino, M.; Trippe, A. J.; Nguyen, D.; Dawson, H.; Wu, B.; Musick, L.; Kosa, M.; Kahil, D.; Webber, S.; Gehlhaar, D. K.; Andrada, D.; Shetty, B., Protein Structure-Based Design of Potent Orally Bioavailable, Nonpeptide Inhibitors of Human Immunodeficiency Virus Protease. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, 92, (8), 3298-3302.

7. Stubbs, M. T.; Reyda, S.; Dullweber, F.; Möller, M.; Klebe, G.; Dorsch, D.; Mederski, W. W.; Wurziger, H., pH-Dependent Binding Modes Observed in Trypsin Crystals: Lessons for Structure-Based Drug Design. *ChemBioChem* **2002**, 3, (2-3), 246-249.

8. Gassel, M.; Breitenlechner, C. B.; Konig, N.; Huber, R.; Engh, R. A.; Bossemeyer, D., The Protein Kinase C Inhibitor Bisindolyl Maleimide 2 Binds with Reversed Orientations to Different Conformations of Protein Kinase A. *J. Biol. Chem.* **2004**, 279, (22), 23679-23690.

9. Bostrom, J.; Hogner, A.; Schmitt, S., Do Structurally Similar Ligands Bind in a Similar Fashion? *J. Med. Chem.* **2006**, 49, (23), 6716-6725.

10. Blum, A.; Böttcher, J.; Heine, A.; Klebe, G.; Diederich, W. E., Structure-Guided Design of C2-Symmetric HIV-1 Protease Inhibitors Based on a Pyrrolidine Scaffold. *J. Med. Chem.* **2008**, 51, (7), 2078-2087.

11. Böttcher, J.; Blum, A.; Heine, A.; Diederich, W. E.; Klebe, G., Structural and Kinetic Analysis of Pyrrolidine-based Inhibitors Showing no Susceptibility towards the Crucial Ile84Val Mutation of HIV-1 Protease **to be published**.

12. Alterman, M.; Hallberg, A., Fast Microwave-Assisted Preparation of Aryl and Vinyl Nitriles and the Corresponding Tetrazoles from Organo-halides. *J. Org. Chem.* **2000**, 65, (23), 7984-7989.

13. Katritzky, A.; Pilarski, B.; Urogdi, L., Efficient Conversion of Nitriles to Amides with Basic Hydrogen Peroxide in Dimethyl Sulfoxide. *Synthesis* **1989**, (12), 949-950.

14. Das, K.; Clark, A. D.; Lewi, P. J.; Heeres, J.; deJonge, M. R.; Koymans, L. M. H.; Vinkers, H. M.; Daeyaert, F.; Ludovici, D. W.; Kukla, M. J.; DeCorte, B.; Kavash, R. W.; Ho, C. Y.; Ye, H.; Lichtenstein, M. A.; Andries, K.; Pauwels, R.; deBethune, M. P.; Boyer, P. L.; Clark, P.; Hughes, S. H.; Janssen, P. A. J.; Arnold, E., Roles of Conformational and Positional Adaptability in Structure-Based Design of TMC125-R165335 (Etravirine) and Related Non-nucleoside Reverse Transcriptase Inhibitors

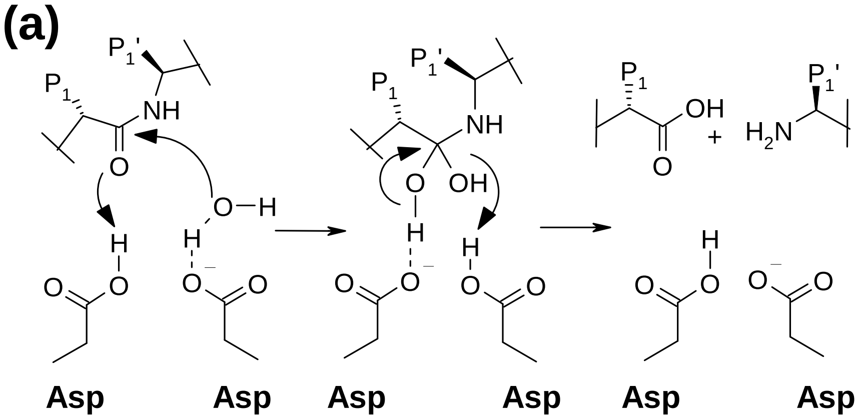
That Are Highly Potent and Effective against Wild-Type and Drug-Resistant HIV-1 Variants. *J. Med. Chem.* **2004**, 47, (10), 2550-2560.

15. Toth, M. V.; Marshall, G. R., A simple, continuous fluorometric assay for HIV protease. *International Journal of Peptide and Protein Research* **1990**, 36, (6), 544-550.
16. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*, Carter Jr., C. W., Ed. Academic Press: 1997; Vol. 276, pp 307-326.
17. Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, 60, (3), 432-438.
18. Sheldrick, G. M.; Schneider, T. R., *SHELXL*: High-resolution refinement. In *Methods in Enzymology*, Charles, W. C. J.Robert, M. S., Eds. Academic Press: 1997; Vol. 277, pp 319-343.
19. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, 60, (12 Part 1), 2126-2132.

6. Achiral Oligoamines as Versatile Tool for the Development of Aspartic Protease Inhibitors *

6.1 Introduction

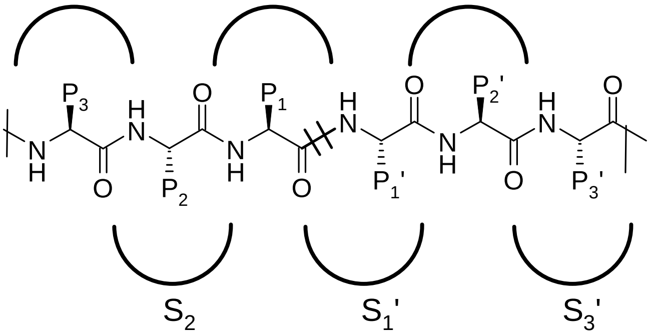
Aspartic proteases belong to the class of endopeptidases and have shown to play an important role in many physiological but also patho-physiological processes. The active site comprises two aspartic acid residues which activate a water molecule hence facilitating the nucleophilic attack at the scissile amide bond. The cleavage of the substrate follows a general catalytic acid-base mechanism in which one of the two aspartates is protonated, the other deprotonated (Scheme 1). [1, 2]



Scheme 1: Schematic representation of the catalytic mechanism of aspartic proteases.

The nucleophilic attack of the activated water molecule leads to a tetrahedral *gem*-diol intermediate which collapses under cleavage of the peptide bond. In general, 6-10 amino acids of the natural polypeptide substrates are recognized by aspartic proteases, thus demanding an extended active site. The standard nomenclature defines the substrate residues as e.g. P3, P2, P1, P1’, P2’, P3’ and the corresponding recognition pockets as e.g. S3, S2, S1, S1’, S2’, S3’ as depicted in Scheme 2.[3]

* Taken from original publication, Andreas Blum, Jark Böttcher, Benedikt Sammet, Torsten Luksch, Andreas Heine, Gerhard Klebe, Wibke E. Diederich. Achiral Oligoamines as Versatile Tool for the Development of Aspartic Protease Inhibitors. *Bioorg. Med. Chem.* **2008**, 16, (18), 8574-86



Scheme 2: Nomenclature of the protease's subsites according to Berger and Schechter,[3] the scissile peptide bond is indicated by crossing lines.

Two classes of aspartic proteases have received pronounced attention as potential drug targets; the pepsin-like[4] (family A1) and the retroviral proteases (family A2).[5-7] The family of the pepsin-like proteases includes e.g. renin (cardiovascular diseases),[8, 9] BACE-1 (Alzheimer's disease),[10, 11] and the plasmepsins (malaria)[12, 13] which all share a common folding motif: The N-terminal and the C-terminal domain, each contributing one catalytic aspartic acid residue, are connected via a β-sheet domain. The active site is formed by the closure of a flexible loop region termed "flap" which covers the active site and forms two hydrogen bonds to the substrate carbonyl groups adjacent to the cleavage site (Figure 1, PDB code: 1PSO).[14]

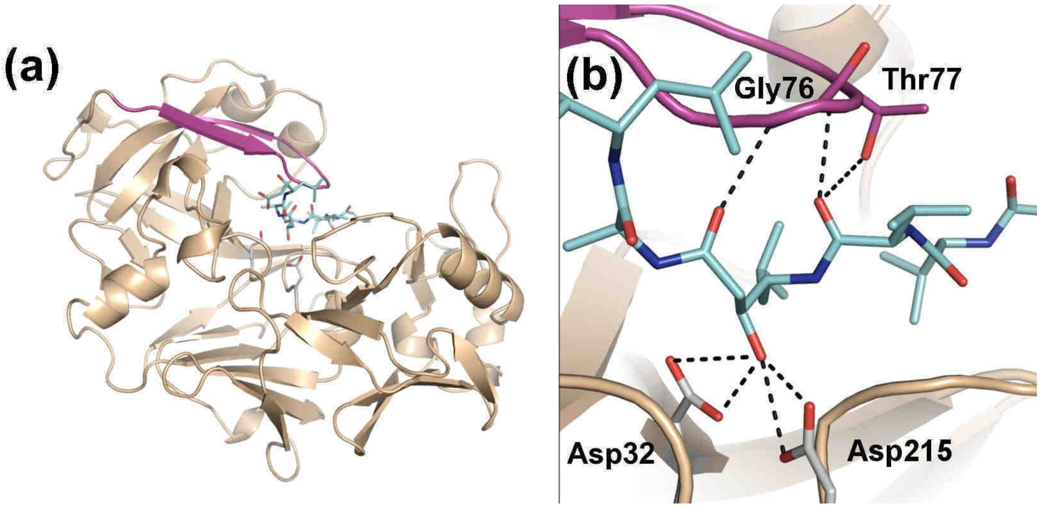


Figure 1: Crystal structures of the peptidomimetic inhibitor Pepstatin A, shown in light blue color-coded by atom type, in complex with pepsin (PDB code 1PSO). The protein is represented as cartoon and flap region establishing interactions to the inhibitor is highlighted in magenta. Hydrogen bonds are indicated by dashed lines. The catalytic aspartates are color-coded by atom type in grey and selected flap amino acid residues are color-coded by atom type in magenta.

In the S1 and S1’ pocket, preferably hydrophobic amino acids are recognized, whereas in the S2, S3, S2’, and S3’ pockets the recognition is non-uniform (Table 1). The retroviral aspartic proteases belong to the A2 family, among which the HIV-1 protease was shown to be a valuable drug target for treatment of HIV-infection.[15-18] The HIV-1 protease is a C_2-symmetric protein consisting of two identical subunits each contributing one catalytic aspartate. In contrast to family A1, the binding pocket is formed by the closure of two flaps. A conserved water molecule mediates the interactions of the peptide carbonyls adjacent to the cleavage site and the flap residues Ile50A and Ile50B (Figure 2, PDB code: 5HVP).[19]

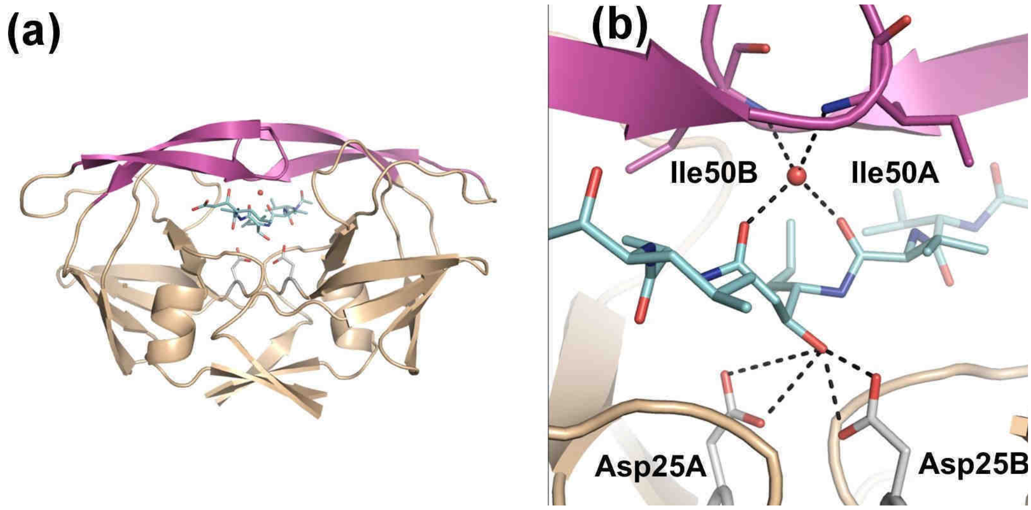


Figure 2: Crystal structure of the peptidomimetic inhibitor Pepstatin A, shown in light blue color-coded by atom type, in complex with the HIV-1 protease (PDB code 5HVP). The proteins are represented as cartoons and flap regions establishing interactions to the inhibitor are highlighted in magenta. Hydrogen bonds are indicated by dashed lines. The catalytic aspartates are color-coded by atom type in grey, selected flap amino acid residues are color-coded by atom type in magenta and the flap water molecule is shown as red sphere.

The HIV protease recognizes a variety of natural substrates with a strong preference for those bearing hydrophobic amino acids in P1 and P1’ and variable amino acids in farther positions (Table 1).[20]

protease	**substrate**	**P3**	**P2**	**P1**	**P1’**	**P2’**	**P3’**
pepsin A	-	-	-	F/L	F	-	-
Plm II	hemoglobin α-chain	R	M	F	L	S	F
Plm IV	hemoglobin α-chain	R	M	F	L	S	F
renin	angiotensinogen	F	H	L	L/V	V/I	H
BACE-1	amyloid-precursor-protein	V	K	M	D	A	Q
HIV-1 protease	gag polyprotein	Q	N	Y	P	I	V
		R	V	L	A	E	A
		T	I	M	M	Q	R
		A	N	F	L	H	K
		G	N	F	L	Q	S
	pol polyprotein	F	N	F	P	Q	I
		L	N	F	P	I	S
		E	T	F	Y	V	D
		K	I	L	F	L	D

Table 1: Cleavage sites in natural substrates of the selected aspartic proteases as annotated in the MEROPS database.[6]

Despite immense efforts in the development of clinically effective drugs targeting aspartic proteases, up to now, only for HIV protease[21] and renin inhibitors have been approved for disease therapy.[22-24] The rational design of inhibitors has mostly been guided by the structure of the natural peptide substrates thus resulting in very potent inhibitors. However, the pharmacokinetic properties of these peptidic compounds bearing secondary hydroxyl groups as transition state isosters are often not optimal, thus hampering their clinical efficacy.[25] The synthesis of these complex and chiral inhibitors often turns out to be very challenging and the optimization of the sub-pocket-addressing moieties is very resource-intensive and time-consuming. Replacement of the secondary hydroxyl group in transition state mimetics by the nearly isosteric amino functionality as well as by cyclic amidines has also been successfully pursued.[26-29] Additionally, reduced amide transition state isoters have been exploited recently as BACE-1 inhibitors.[30, 31] However, for a more efficient development of novel aspartic protease inhibitors, an easily accessible and achiral core structure would be preferable. Utilizing such scaffolds would facilitate the development of promising moieties addressing the protease's sub-pockets. As suitable core structure for this approach we selected secondary amines, which have already been successfully utilized as anchoring groups for the development of aspartic protease inhibitors.

In the late 1990s, researchers at Roche discovered the piperidine moiety as a novel privileged skeleton addressing the catalytic dyad of aspartic proteases with its secondary nitrogen as revealed by X-ray crystallography.[32-34] Based on this discovery, several projects for the development of non-peptidic amine-based aspartic protease inhibitors have been pursued, which led to the development of potent inhibitors for the aspartic proteases plasmepsin (Plm) II and IV,[35-38] BACE-1,[39] and very recently for HIV-1 protease.[40] All inhibitors share in common a cyclic amino functionality addressing the enzyme's catalytic dyad directly or mediated by a water molecule. However, numerous scaffolds have been utilized. Depending on the nature of the scaffold and the target enzyme studied, the specificity pockets are addressed differently. In case of a pyrrolidine-based HIV-1 protease inhibitor, the interaction of the endocyclic, secondary amino functionality to the catalytic aspartic acid residues was studied by X-ray crystallography and *Poisson-Boltzmann* calculations. The latter study suggests the amine being in the protonated and the catalytic dyad in the fully deprotonated state resulting in strong electrostatic interactions.[41]

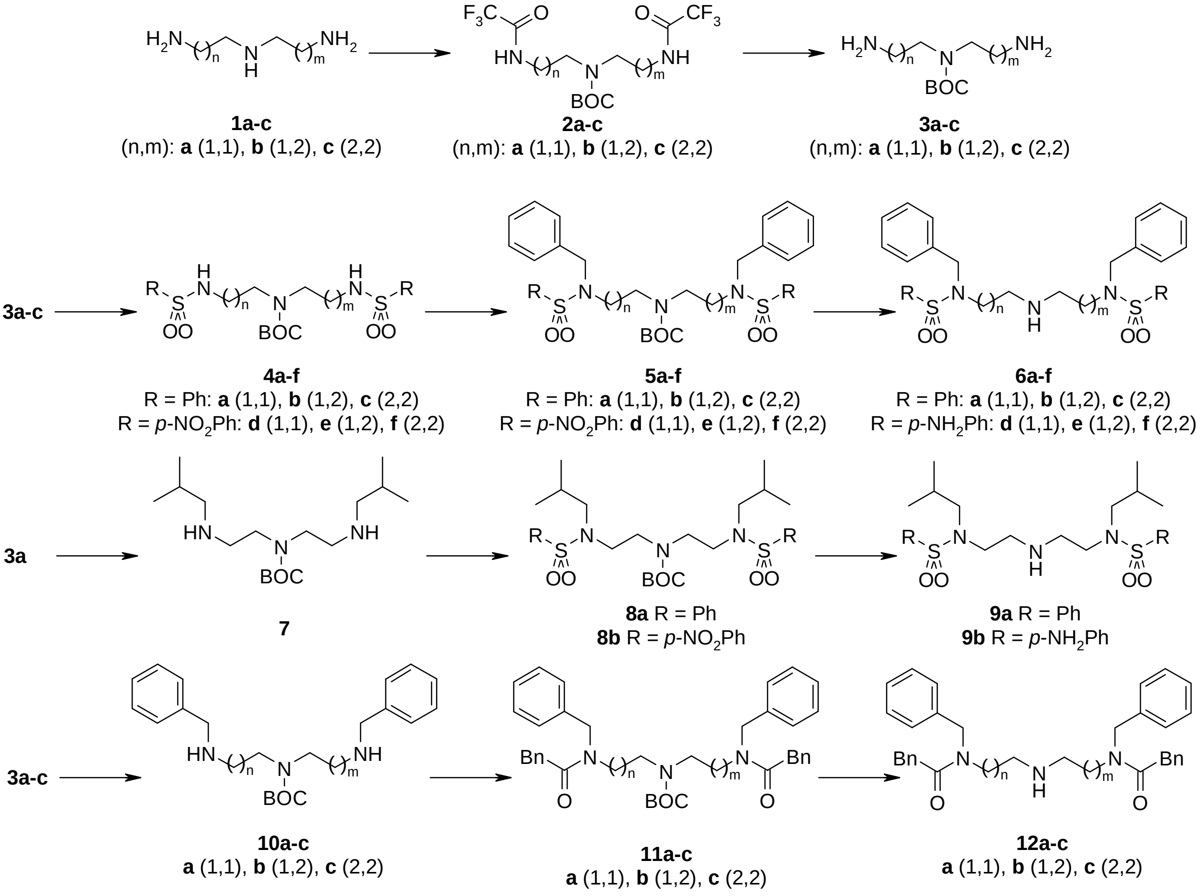
6.2 Results

To further exploit this promising anchoring group for the design of novel, easily accessible inhibitors, linear oligoamines were selected as novel core structure. This scaffold bears a central secondary amino function aimed to address the catalytic dyad. Via the distal amino functionalities, the introduction of appropriate acceptor groups addressing the flap regions of the respective target enzyme is easily amenable. Concomitantly hydrophobic moieties intended to address the S2 and S2' specificity pockets are introduced. Via further alkylation of the distal nitrogen atoms, additional hydrophobic substituents can be implemented to address the S1/S1'-pockets, thereby avoiding the generation of any chiral center. This strategy allows the quick and straightforward generation of achiral inhibitors following a highly flexible and short synthetic route. To optimize the amine-acceptor distances with respect to a certain target enzyme, oligoamines with varying chain length between the central and each of the distal amino functionalities can be employed.

For the design of an initial library, linkers with two and three methylene groups separating the amino groups were selected. As hydrophobic moieties intended to address the S1/S1' sub-pockets, *iso*-butyl and benzyl groups resembling the side chains of leucine and phenylalanine of substrate peptides were introduced. For appropriate flap interactions, two types of acceptor functionalities were investigated: on the one hand carboxamides present in the natural substrate and on the other hand sulfonamides, which allow additional rotational degrees of freedom. The introduction of residues addressing the S2/S2' pockets is easily achieved via further substitution at the lateral nitrogen atoms. In case of the sulfonamides, the hydrophobic phenyl and the more polar *p*-amino phenyl group were chosen. For carboxamides, phenyl acetic acid derivatives were utilized offering additional flexibility between the rigid amide functionality and the phenyl ring.

6.2.1 Chemistry

Oligoamines **3a-c**, BOC-protected at the pivotal, secondary amino function, are accessible from commercially available precursors **1a-c** by transient protection of the terminal amino groups as their trifluoroacetamides **2a-c** (Scheme 3).[42] Condensation with sulfonyl chlorides furnished the sulfonamides **4a-f** which were further alkylated with benzyl bromide giving rise to the protected inhibitors **5a-f**. The phenyl-substituted sulfonamides **6a-c** were obtained via deprotection of **5a-c** with HCl in Et_2O. The *p*-amino phenyl-substituted sulfonamides **6d-f** were synthesized by reduction of the corresponding nitro-compounds **5d-f** with $SnCl_2*2H_2O$ in refluxing 32% HCl under concomitant cleavage of the BOC-protecting group. The *iso*-butyl-substituted inhibitors **9a**,**b** are accessible via reductive amination of **3a** with *iso*-butyraldehyde and $NaBH_4$ as reducing agent yielding **7**, followed by condensation with sulfonyl chlorides rendering **8a**,**b**. The final inhibitors **9a**,**b** were obtained by reduction of the nitro functionality or by acidic cleavage of the BOC-protecting group as described above. Carboxamide inhibitors **12** were synthesized following a similar synthetic sequence: The benzyl-substituted oligoamines **10a-c** were obtained by reductive amination of the amines **3a-c** and benzaldehyde utilizing $Pd/BaSO_4$ as catalyst. Subsequent condensation with phenylacetic acid chloride gave rise to the protected inhibitors **11a-c,** which were deprotected under acidic anhydrous conditions yielding **12a-c**.


1a-c
(n,m): a (1,1), b (1,2), c (2,2)
2a-c
(n,m): a (1,1), b (1,2), c (2,2)
3a-c
(n,m): a (1,1), b (1,2), c (2,2)
3a-c
4a-f
R = Ph: a (1,1), b (1,2), c (2,2)
R = p-NO2Ph: d (1,1), e (1,2), f (2,2)
5a-f
R = Ph: a (1,1), b (1,2), c (2,2)
R = p-NO2Ph: d (1,1), e (1,2), f (2,2)
6a-f
R = Ph: a (1,1), b (1,2), c (2,2)
R = p-NH2Ph: d (1,1), e (1,2), f (2,2)
3a
7
8a R = Ph
8b R = p-NO2Ph
9a R = Ph
9b R = p-NH2Ph
3a-c
10a-c
a (1,1), b (1,2), c (2,2)
11a-c
a (1,1), b (1,2), c (2,2)
12a-c
a (1,1), b (1,2), c (2,2)

Scheme 3: Synthetic sequence for the preparation of the oligoamines.

6.2.2 Kinetic characterization

To elucidate the potential of this new class of inhibitors, the affinity toward selected aspartic proteases was determined in fluorescence-based assays using commercially available substrates. The results are listed in Table 2. The anthranyl-HIV protease substrate was utilized in case of the HIV-1 protease (K_m = 14.6 µM), pepsin (K_m = 13.3 µM), and the plasmepsins II (K_m = 63 µM) and IV (K_m = 28 µM). For renin and BACE-1, suitably labeled oligopeptides derived from their natural substrates with comparable affinity were used (renin substrate K_m = 3.3 µM, BACE-1 substrate K_m = 7.9 µM).

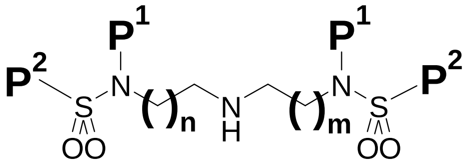
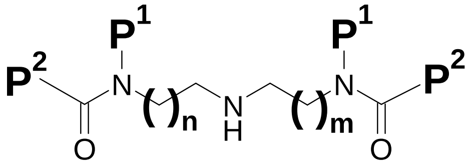
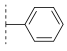
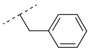
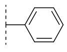
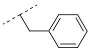
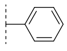
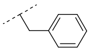
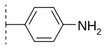
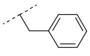
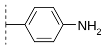
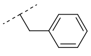
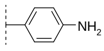
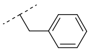
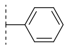
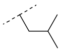
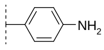
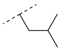
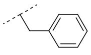
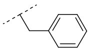
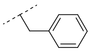
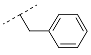
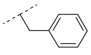
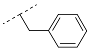
 **6a-f, 9a-b**					 **12a-c**					
	n	**m**	**P^2**	**P^1**	**HIV-1 pr.**	**Plm II**	**Plm IV**	**renin**	**BACE-1**	**pepsin**
6a	1	1			3.8	7.0	36	n.i.	n.i.	5.7
6b	1	2			14	18	17	n.i.	n.i.	15
6c	2	2			n.i.	21	42	n.i.	150	n.i.
6d	1	1			9.6	4.2	7.5	n.i.	n.i.	3.9
6e	1	2			n.i.	22	7.5	n.i.	n.i.	n.i.
6f	2	2			n.i.	2.8	20	n.i.	64	4.0
9a	1	1			10	23	87	n.i.	n.i.	4.5
9b	1	1			0.9	6.8	56	n.i.	n.i.	4.8
12a	1	1			n.i.	91	98	5,7	n.i.	n.i.
12b	1	2			n.i.	63	56	n.i.	n.i.	n.i.
12c	2	2			n.i.	120	170	16	n.i.	n.i.

Table 2: Ki-values of the inhibitors towards selected aspartic proteases in µM (n.i. (IC_{50} > 250 µM)).

Except for BACE-1, for each of the investigated target enzymes an inhibitor with an affinity in the single-digit micromolar range could be identified (Table 2). In case of the HIV-1 protease, a strong preference for inhibitors bearing ethylene linkers and a sulfonyl moiety was observed, with affinities ranging from 10 µM for **9a** up to 0.9 µM for **9b**. Considering the pepsin-like proteases, the tolerance of extended linkers is noteworthy. Particularly the plasmepsins II and IV reveal a remarkable tolerance toward the linker length: inhibitor **6f** exhibits the highest affinity towards Plm II (2.8 µM) and the corresponding analogues **6d** and **6e** toward Plm IV (7.5 µM). An inhibition of renin could only be observed in case of derivatives equipped with carbonyl groups as acceptor functionalities. With compound **12a** bearing the shortest linkers, the highest affinity (5.7 µM) could be achieved. An inhibition of pepsin is only accomplished with sulfonamide derivatives, **6d** being the most active compound (3.9 µM) in the series. However, no preference toward a certain chain length could be observed. In case of BACE-1, the sulfonamides equipped with the longest linkers show moderate inhibition with an affinity of 64 µM for compound **6f**.

6.2.3 Structural analysis

To elucidate whether the oligoamine derivatives indeed exhibit the expected binding mode thus being in agreement with our initial design concept, the crystal structure in complex with one of the target enzymes was determined. Since the crystallization of HIV-1 protease is well established in our lab, this enzyme was selected for crystallization. The inhibitors **6a**, **6d**, and **9b** exhibiting the highest potency were selected for cocrystallization experiments. However, crystals could only be obtained in case of the better soluble compounds **6d** and **9b**. The X-ray structures of these compounds in complex with HIV-1 protease were determined with a resolution of 1.90 and 1.80 Å, respectively. The crystallographic data and refinement statistics are listed in Table 3. Both complexes adopt space group $P2_12_12$ and the inhibitors are clearly visible in the F_o-F_c density at a sigma level of 3. They could be refined as one single conformer (Figure 3).

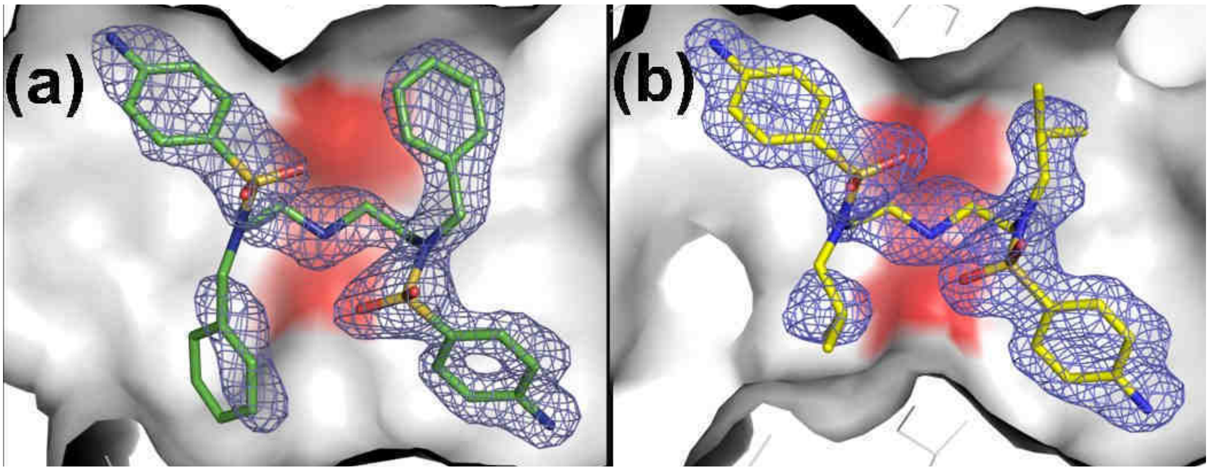


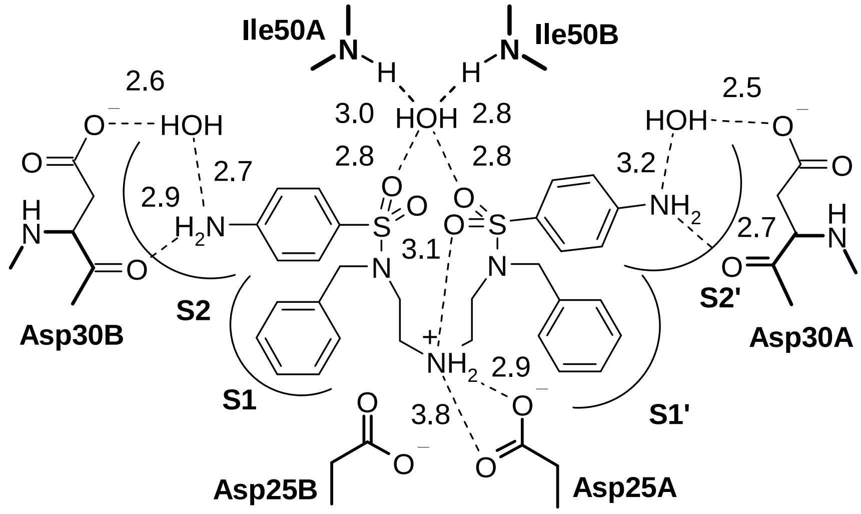
Figure 3: Ligand geometries of **6d** (**a,** color-coded by atom type in green) and **9b** (**b,** color-coded by atom type in yellow) in the cocrystal structures with the HIV-1 protease. The F_o-F_c omit maps for the ligands are displayed at a σ level of 3.0 as blue mesh. The protein surface is schematically represented in grey and the catalytic aspartates Asp25A and Asp25B are highlighted in red.

	6d	**9b**
resolution (Å)	50-1.80	30-1.90
space group	$P2_12_12$	$P2_12_12$
cell dimensions (Å)	a = 57.4 b = 85.7 c = 46.5	a = 56.0 b = 85.9 c = 46.5
highest resolution shell (Å)	1.83 -1.80	1.93 -1.90
no. of measured reflections	49890	65713
no. of independent reflections	21367	18795
completeness (%)[a]	97.4 [97.6]	99.0 [100]
I/σ[a]	16.0 [2.0]	15.6 [2.6]
R sym (%)[a]	7.8 [44.4]	8.0 [49.0]
Rcryst ($F > 4\sigma$ Fo; Fo)	18.4; 20.8	18.2; 20.9
Rfree($F > 4\sigma$ Fo; Fo)	23.0; 26.0	23.9; 27.1
mean B-factor ($Å^2$) protein (chain A; B)	25.9; 22.4	29.5; 27.1
mean B-factor ($Å^2$) ligand	35.1	33.1
mean B-factor ($Å^2$) water	31.0	34.0
Ramachandran plot		
most favored geometry (%)	96.8	95.6
additionally allowed (%)	3.2	4.4
RMSD bonds (Å)	0.006	0.005
RMSD angles (°)	2.0	2.0

Table 3: X-ray data processing and refinement for the HIV-1 Protease complexes of derivative **6d** and **9b**. ([a])Values in brackets refer to the highest resolution shell.

6.2.3.1 Binding mode of 6d

The binding mode is schematically represented in Scheme 4. The pivotal secondary nitrogen of the inhibitor points toward Asp25A forming polar contacts to the terminal oxygen atoms of the side chain and to one sulfonyl oxygen atom of one of its sulfonamide functions.

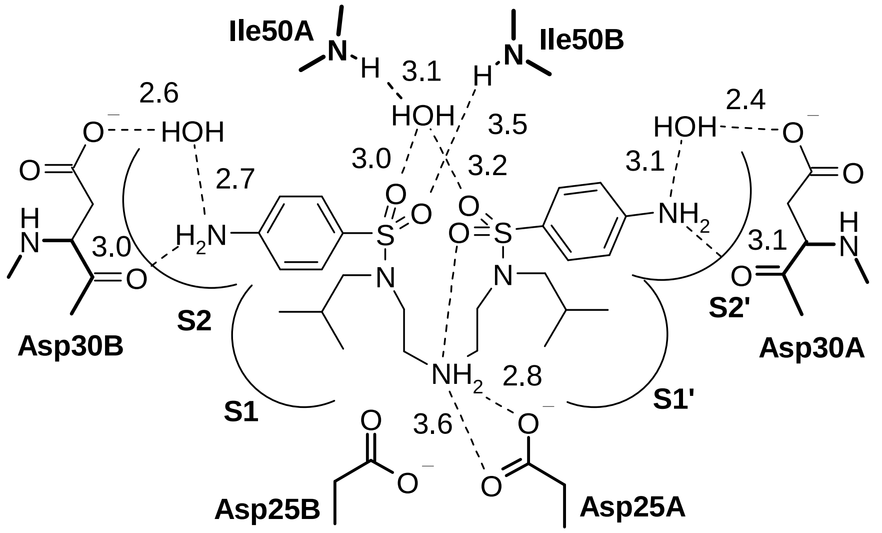


Scheme 4: Schematic representation of the binding mode of **6d**. Hydrogen bonds are indicated by dashed lines and distances are given in Å.

The *N*-benzyl moieties of the inhibitor address the S1 and S1’ pockets, whereas the *p*-amino-phenyl sulfonamide groups occupy the S2 and S2’ pockets. The *p*-amino substituents form a similar hydrogen bond network in both pockets: hydrogen bonds are observed to the main chain carbonyl oxygen atoms of Asp30A and Asp30B and to the corresponding side chain carboxyl oxygen atoms in each case mediated by a water molecule. One sulfonyl oxygen atom of each sulfonamide functionality establishes hydrogen bonds to a tetrahedrally coordinated water molecule mediating polar contacts to the main chain NHs of Ile50A and Ile50B, which are located at the tips of the flaps. One of the sulfonyl oxygen atoms remains unsatisfied showing no direct polar contacts to the protein. The ligand is deeply embedded in the binding pocket with a surface burial of 94%.

6.2.3.2 Binding mode of 9b

The crystal structure of **9b** in complex with HIV-1 protease exhibits a high similarity to the previously described complex. Comparable polar interactions of the secondary amino group to Asp25A are observed (Scheme 5).



Scheme 5: Schematic representation of the binding mode of **9b**. Hydrogen bonds are indicated by dashed lines and distances are given in Å.

The sub-pocket occupancy resembles the binding situation of **6d**: The *iso*-butyl moieties occupy the S1 and S1' pockets, and the *p*-amino-phenyl sulfonamide groups are located in the S2 and S2' pockets. The amino substituents form a similar hydrogen bond network as already described for **6d**: hydrogen bonds to the main chain carbonyl oxygen of Asp30A, Asp30B as well as water-mediated to the corresponding side chain carboxyl oxygens are formed. The most pronounced difference between both complexes is the deviating polar interaction of the sulfonyl groups of the inhibitor to the flap. Although a water molecule is involved in the mediation of polar contacts of the sulfonyl oxygen atoms in both cases, a different coordination is observed. In contrast to **6d**, the water in the enzyme-inhibitor complex of **9b** is not tetrahedrally coordinated. It only forms hydrogen bonds to two sulfonyl oxygen atoms and the main chain NH of Ile50A. One of the sulfonyl oxygen atoms establishes an intramolecular hydrogen bond to the pivotal oligoamine nitrogen atom, as observed in case of **6d**. In contrast to the previous structure, the second sulfonyl oxygen atom directly addresses Ile50 NH of the

B chain, whereas in case of **6d**, this oxygen atom remains without any polar contacts. The overall burial of the inhibitor sums up to 93%.

6.3 Discussion

The biological data as well as the crystal structures of **6d** and **9b** in complex with HIV-1 protease clearly reveal that our initial design concept has been successful. For all of the investigated aspartic proteases inhibitors binding in the micromolar range could be identified. Noteworthy, for most of these enzymes a preference toward ethylene linkers is observed. The sulfonamide acceptor groups allowing additional rotational degrees of freedom are favored compared to the carbonyl groups for all targets except for renin. In the latter case the more rigid structure of *N*-alkylated carboxamides is obviously preferred for inhibition.

The binding modes of **6d** and **9b** in HIV-1 protease reveal how these compounds fulfill the initial pharmacophore hypothesis. As intended, the secondary amino group addresses the catalytic dyad, and two of the sulfonamide oxygen atoms mimic the substrate's amide carbonyl groups. This is illustrated by the C_α-superposition of the complex structures of **6d** and **9b**, respectively, with the crystal structure of a substrate analogue oligopeptide in complex with an inactive HIV-1 protease mutant (PDB code: 1KJH) (Figure 4a and 4b).[20]

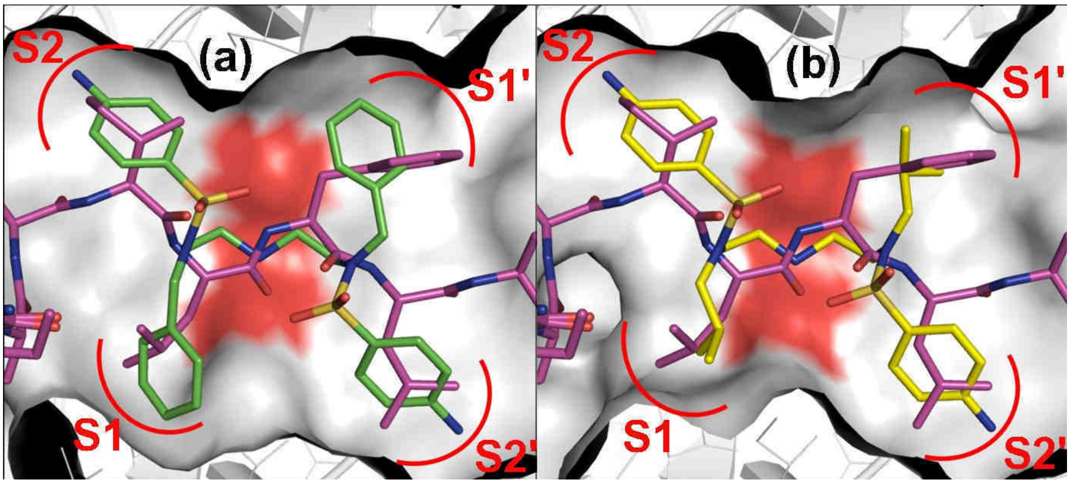


Figure 4: (a, b) C_α-superposition of the cocrystal structures of **6d** (green, color-coded by atom type) and **9b** (yellow, color-coded by atom type) with the structure of a substrate analogue polypeptide (magenta, color-coded by atom type) in complex with an inactive Asp25Asn HIV protease mutant (PDB-code: 1KJH) The protein surface is schematically represented in grey and the catalytic aspartates Asp25A and Asp25B are highlighted in red.

In addition, the specificity pockets of the enzyme are occupied analogously, however, the direction in which the S1/S1’ pocket is addressed deviates from the substrate’s one. The ten-fold lower binding affinity of the inhibitor **6d** compared to **9b** can most likely be attributed to this difference: the benzyl moieties of **6d** are in our case now slightly too large to occupy the S1/S1’ pockets appropriately.

The C_α-superposition of the crystal structures in complex with the HIV-1 protease of the approved peptidomimetic inhibitor Amprenavir (PDB code: 1HPV) and **9b**, which bears the same P1 and P2 substituents, is shown in Figure 5.[43]

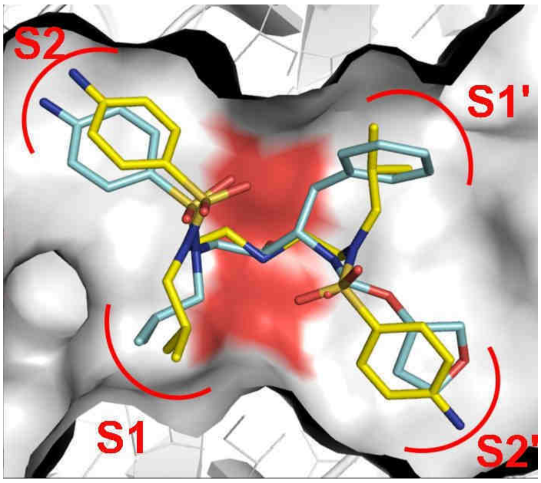


Figure 5: C_α-superposition of the cocrystal structures of **9b** (yellow, color-coded by atom type) with the cocrystal structure of the approved inhibitor Amprenavir (blue, color-coded by atom type, PDB-code: 1HPV). The protein surface is schematically represented in grey and the catalytic aspartates Asp25A and Asp25B are highlighted in red.

Comparing the positions of the *N*-*iso*-butyl-*N*-*p*-amino-phenyl sulfonamide moieties of the inhibitors which address the S1 and S2 pocket, only slight differences can be observed. Although the inhibitors comprise different moieties, the occupation of the S1’ and S2’ pockets appears similar. However, due to their polymethylene linkers, the oligoamine derivatives are very flexible molecules and a high degree of pre-organization prior to binding to the protein is required, resulting in an entropic penalty thus lowering the overall affinity. Additionally, the position of the flap-water molecule present in the complexes of **6d** and **9b** is different in the oligoamine structures when compared to the substrate-like oligopeptide and Amprenavir complexes. In case of **6d**, a similar tetrahedral coordination of the water molecule is observed, but it is shifted about 1 Å away from the catalytic dyad. A deviation from the usually observed coordination is detected in case of the **9b** complex leading to a direct interaction of one sulfonyl oxygen atom with the flap. These observations indicate that the flap interactions of the oligoamines **6d** and **9b** in case of HIV-1 might require some further improvement. Taking all these facts into account, the lower affinity by three orders of magnitude of the oligoamines compared to Amprenavir (K_i = 0.6 nM) could find an explanation.

The observed peptide-like binding mode in case of the ethylene-bearing inhibitors **6d** and **9b** in complex with HIV-1 protease suggests a similar binding situation in case of the other pepsin-like proteases. In this case, the acceptor functionalities of the inhibitors mimic the carbonyl groups of the substrate adjacent to the cleavage site, resulting in the same subsite occupation as observed for **6d** and **9b**. However, in contrast to HIV protease, there is no clear preference for the ethylene linkers. In pepsin-like proteases, amino acids bearing flexible polar side chains are located in the flap region: Thr77 in case of pepsin (Figure 1b), Ser76 and Thr77 in case of renin, Thr72 and Gln73 in case of BACE-1, and Ser79 in case of Plm II and Plm IV. In addition to the peptide recognition motifs, these residues might also be addressed by the acceptor groups of the inhibitors resulting in the preference for longer linkers. The binding of inhibitors with propylene linkers results in a larger distance between the sub-pocket-addressing hydrophobic moieties and could omit occupancy of the S1 or S1’ pockets. This might explain the inhibition of BACE-1 only by the inhibitors **6c** and **6f**, because the S1’-pocket of BACE-1 is featured to recognize an aspartic acid side chain. Most likely, the studied oligoamine inhibitors are not suitable to address this site.

6.4 Summary and Conclusion

In this study, we present a novel rational strategy for the development of aspartic protease inhibitors. This method is based on achiral linear oligoamines, which are easily accessible via a short and flexible synthetic route starting from commercially available precursors. Within the initial compound series comprising eleven inhibitors, several hits for the target enzymes could be identified. Some of them reveal remarkable selectivity, others show broad promiscuity. Noteworthy, for five out of six target enzymes, at least one single-digit micromolar inhibitor could be identified. The crystal structures of two representatives in complex with the HIV-1 protease proved the concept initially used for their design. The central amino group addresses the catalytic dyad, the acceptor groups establish polar interactions to the flap, and the hydrophobic moieties occupy the sub-pockets of the enzyme. The binding mode exhibits high similarity to the binding orientation of substrates as well as to that of other peptidomimetic inhibitors. Taking this into account, the generalization of the observed binding situation to other aspartic proteases appears reasonable thus providing a first insight into the observed structure-activity relationships.

The straightforward synthesis allows the combinatorial introduction of numerous different substituents, thus facilitating the preparation of a plethora of putative aspartic protease inhibitors. The derived SAR data from such libraries can easily be utilized for the selection of promising moieties for selectively addressing the specificity pockets of a target enzyme in order to optimize inhibitors bearing other scaffolds. Hits from these oligoamine libraries can also be utilized as suitable starting point for further optimization. Structural variation of the oligoamine chain based on a certain inhibitor, at best guided by a crystal structure, can lead to novel derived amine-based scaffolds. Increasing the rigidity of the linker chain is the most promising approach, leading to a better preorganization of the inhibitor and a reduction of the entropic penalty costs upon binding hence most likely resulting in a gain of affinity.

6.5 Experimental Section

6.5.1 Enzyme Assays

Enzyme Assays: All enzyme assays were performed at room temperature on a microplate reader (Safire^{2TM}) using black 96-well microtiter plates purchased from Nunc. The assay volume was 200 µl, and inhibitors and substrates were previously dissolved in dimethyl sulfoxide (final concentration 4%). The hydrolysis of the substrates was recorded as increase in fluorescence intensity. IC_{50} values were generated by nonlinear regression analysis from plots of v_i/v_0 versus inhibitor concentration, in which v_i is the velocity in presence, and v_0 the velocity in the absence of an inhibitor. The kinetic constants were determined by the method of Lineweaver and Burk and K_i values were consecutively calculated from the following equation: $K_i = IC_{50} / [1 + (S/K_m)]$. The overall error of the assays is estimated to be ±40%.

HIV-1 Protease: Recombinant HIV-1 protease was expressed from *Escherichia coli* and purified as previously described.[44] The fluorogenic substrate Abz-Thr-Ile-Nle-(p-NO_2-Phe)-

Gln-Arg-NH_2 was purchased from Bachem and exhibited a K_m of 14.6 µM. The assays were performed in 100 mM MES, 300 mM KCl, 5 mM EDTA, 1 mg/mL BSA, pH 5.5, and a substrate concentration of 20 µM, (excitation wavelength 337 nm, emission wavelength 410 nm).[45]

Plasmepsin II and Plasmepsin IV: Recombinant plasmepsin II and plasmepsin IV were expressed from *Escherichia coli* and purified following the protocol described by Hill *et al.*[46] The fluorogenic substrate Abz-Thr-Ile-Nle-(*p*-NO_2-Phe)-Gln-Arg-NH_2 was purchased from Bachem (K_m (Plm II) = 63 µM and K_m (Plm IV) = 28 µM). The assays were performed in 100 mM NaAc, pH 4.5, 1 mg/mL BSA and a substrate concentration of 18 µM (excitation wavelength 337 nm, emission wavelength 410 nm).

Renin: The fluorogenic substrate Arg-Glu(EDANS)-Ile-His-Pro-Phe-His-Leu-Val-Ile-His-Thr-Lys(DABCYL)-Arg and recombinant renin were purchased from Sigma-Aldrich (K_m = 3.3µM). Assays were performed in 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 8.0 and a substrate concentration of 5 µM (excitation wavelength 340 nm, emission wavelength 500 nm).[47]

β-Secretase: The fluorogenic substrate MOCAc-Ser-Glu-Val-Asn-Leu-Asp-Ala-Glu-Phe-Arg-(2,4-dinitrophenyl)-Lys-Arg-Arg-NH_2 and recombinant β-secretase were purchased from Sigma (K_m = 7.9 µM). Assays were performed in 100 mM NaAc, pH 4.5, and a substrate concentration of 5 µM (excitation wavelength 320 nm, emission wavelength 405 nm).

Pepsin: Pepsin from porcine gastric mucosa was purchased from Sigma-Aldrich and the HIV protease substrate Abz-Thr-Ile-Nle-(*p*-NO_2-Phe)-Gln-Arg-NH_2 was used for kinetic measurements (K_m = 13.3 µM). Assays were performed in 10 mM NaHCOO, pH 3.5 and a substrate concentration of 20 µM (excitation wavelength 337 nm, emission wavelength 410 nm).

6.5.2 Structural analysis

Crystallization of the HIV-1 protease inhibitor complexes: The HIV-1 protease inhibitor complexes were crystallized at 18°C in 0.1 M BisTris, pH 6.5, 2.5-3.0 M NaCl and a protein concentration of 7 mg/mL in the space group $P2_12_12$ (crystal data, Table 4). The crystals were obtained by cocrystallization of the enzyme with inhibitor concentrations ranging from 20 to 100 fold the K_i value. Crystals were further optimized using streak-seeding techniques. For cryoprotection the crystals were briefly soaked in mother liquor containing 25% glycerol.

Data collection, phasing and refinement: The data sets were collected on a Rigaku R-AXIS IV image plate detector using Cu K_α radiation from an in-house RU-H3R Rigaku rotating anode. Data were processed and scaled with Denzo and Scalepack as implemented in HKL2000.[48] The structures were determined by the molecular replacement method using Phaser,[49] one monomer of the 1.50 Å HIV-1 protease in complex with a pyrrolidine based inhibitor (PDB code 2PQZ) was used as the search model. The structure refinement was continued with SHELXL-97,[50] for each refinement step at least 10 cycles of conjugate gradient minimization were performed, with restraints on bond distances, angles and B-values. Intermittent cycles of model building were done with the program COOT.[51] The coordinates have been deposited in the PDB (http://www.rcsb.org/pdb/) with access codes: 3BGC and 3BGB.

6.6 References

1. Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. R., Binding of a reduced peptide inhibitor to the aspartic proteinase from Rhizopus chinensis: implications for a mechanism of action. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, 84, (20), 7009-7013.
2. Northrop, D. B., Follow the Protons: A Low-Barrier Hydrogen Bond Unifies the Mechanisms of the Aspartic Proteases. *Acc. Chem. Res.* **2001**, 34, (10), 790-797.
3. Schechter, I.; Berger, A., On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, 27, (2), 157-162.
4. Dunn, B. M., Structure and Mechanism of the Pepsin-Like Family of Aspartic Peptidases. *Chem. Rev.* **2002**, 102, (12), 4431-4458.
5. Cooper, J. B., Aspartic proteinases in disease: a structural perspective. *Current Drug Targets* **2002**, 3, (2), 155-173.
6. Rawlings, N. D.; Morton, F. R.; Barrett, A. J., MEROPS: the peptidase database. *Nucleic Acids Res.* **2006**, 34, D270-272.

7. Eder, J.; Hommel, U.; Cumin, F.; Martoglio, B.; Gerhartz, B., Aspartic Proteases in Drug Discovery. *Curr. Pharm. Des.* **2007**, 13, 271-285.
8. Greenlee, W. J., Renin Inhibitors. *Med. Res. Rev.* **1990**, 10, (2), 173-236.
9. Tice, C. M.; Anthony, W., Renin Inhibitors. In *Annu. Rep. Med. Chem.*, Academic Press: 2006; Vol. Volume 41, pp 155-167.
10. Hardy, J.; Selkoe, D. J., The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. *Science* **2002**, 297, (5580), 353-356.
11. Durham, T. B.; Shepherd, T. A., Progress toward the discovery and development of efficacious BACE inhibitors. *Curr Opin Drug Discov Devel* **2006**, 9, (6), 776-91.
12. Silva, A. M.; Lee, A. Y.; Gulnik, S. V.; Majer, P.; Collins, J.; Bhat, T. N.; Collins, P. J.; Cachau, R. E.; Luker, K. E.; Gluzman, I. Y.; Francis, S. E.; Oksman, A.; Goldberg, D. E.; Erickson, J. W., Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, 93, (19), 10034-10039.
13. Ersmark, K.; Samuelsson, B.; Hallberg, A., Plasmepsins as potential targets for new antimalarial therapy. *Med. Res. Rev.* **2006**, 26, (5), 626-666.
14. Fujinaga, M.; Chernaia, M. M.; Tarasova, N. I.; Mosimann, S. C.; James, M. N. G., Crystal structure of human pepsin and its complex with pepstatin. *Protein Sci.* **1995**, 4, (5), 960-972.
15. Wlodawer, A.; Erickson, J. W., Structure-Based Inhibitors of HIV-1 Protease. *Annual Review of Biochemistry* **1993**, 62, (1), 543-585.
16. Meadows, D. C.; Gervay-Hague, J., Current Developments in HIV Chemotherapy. *ChemMedChem* **2006**, 1, (1), 16-29.
17. Randolph, J. T.; DeGoey, D. A., Peptidomimetic Inhibitors of HIV Protease. *Curr. Top. Med. Chem.* **2004**, 4, 1079-1095.
18. Chrusciel, R. A.; Strohbach, J. W., Non-Peptidic HIV Protease Inhibitors. *Curr. Top. Med. Chem.* **2004**, 4, 1097-1114.
19. Fitzgerald, P. M.; McKeever, B. M.; VanMiddlesworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C.; Herber, W. K.; Dixon, R. A.; Darke, P. L., Crystallographic Analysis of a Complex between Human Immunodeficiency Virus Type 1 Protease and Acetyl-Pepstatin at 2.0-Å Resolution. *J. Biol. Chem.* **1990**, 265, (24), 14209-14219.
20. Prabu-Jeyabalan, M.; Nalivaika, E.; Schiffer, C. A., Substrate Shape Determines Specificity of Recognition for HIV-1 Protease: Analysis of Crystal Structures of Six Substrate Complexes. *Structure* **2002**, 10, (3), 369-381.

21. Pomerantz, R. J.; Horn, D. L., Twenty years of therapy for HIV-1 infection. *Nature Medicine* **2003**, 9, (7), 867-873.

22. Göschke, R.; Stutz, S.; Rasetti, V.; Cohen, N. C.; Rahuel, J.; Rigollier, P.; Baum, H. P.; Forgiarini, P.; Schnell, C. R.; Wagner, T.; Gruetter, M. G.; Fuhrer, W.; Schilling, W.; Cumin, F.; Wood, J. M.; Maibaum, J., Novel 2,7-Dialkyl-Substituted 5(S)-Amino-4(S)-hydroxy-8-phenyl-octanecarboxamide Transition State Peptidomimetics Are Potent and Orally Active Inhibitors of Human Renin. *J. Med. Chem.* **2007**, 50, (20), 4818-4831.

23. Cohen, N. C., Structure-Based Drug Design and the Discovery of Aliskiren (Tekturna): Perseverance and Creativity to Overcome a R&D Pipeline Challenge. *Chem. Biol. Drug. Des.* **2007**, 70, (6), 557-565.

24. Jensen, C.; Herold, P.; Brunner, H. R., Aliskiren: the first renin inhibitor for clinical treatment. *Nat Rev Drug Discov* **2008**, 7, (5), 399-410.

25. Thompson, L. A.; Tebben, A. J., Chapter 24. Pharmacokinetics and design of aspartyl protease inhibitors. In *Annu. Rep. Med. Chem.*, Academic Press: 2001; Vol. Volume 36, pp 247-256.

26. Arrowsmith, R. J.; Carter, K.; Dann, J. G.; Davies, D. E.; Harris, C. J.; Morton, J. A.; Lister, P.; Robinson, J. A.; Williams, D. J., Novel renin inhibitors: synthesis of aminostatine and comparison with statine-containing analogs. *Chem. Soc., Chem. Commun.* **1986**, (10), 755-7.

27. Yang, W.; Lu, W.; Lu, Y.; Zhong, M.; Sun, J.; Thomas, A. E.; Wilkinson, J. M.; Fucini, R. V.; Lam, M.; Randal, M.; Shi, X. P.; Jacobs, J. W.; McDowell, R. S.; Gordon, E. M.; Ballinger, M. D., Aminoethylenes: A Tetrahedral Intermediate Isostere Yielding Potent Inhibitors of the Aspartyl Protease BACE-1. *J. Med. Chem.* **2006**, 49, (3), 839-842.

28. Baxter, E. W.; Conway, K. A.; Kennis, L.; Bischoff, F.; Mercken, M. H.; DeWinter, H. L.; Reynolds, C. H.; Tounge, B. A.; Luo, C.; Scott, M. K.; Huang, Y.; Braeken, M.; Pieters, S. M. A.; Berthelot, D. J. C.; Masure, S.; Bruinzeel, W. D.; Jordan, A. D.; Parker, M. H.; Boyd, R. E.; Qu, J.; Alexander, R. S.; Brenneman, D. E.; Reitz, A. B., 2-Amino-3,4-dihydroquinazolines as Inhibitors of BACE-1 (bSite APP Cleaving Enzyme): Use of Structure Based Design to Convert a Micromolar Hit into a Nanomolar Lead. *J. Med. Chem.* **2007**, 50, (18), 4261-4264.

29. Edwards, P. D.; Albert, J. S.; Sylvester, M.; Aharony, D.; Andisik, D.; Callaghan, O.; Campbell, J. B.; Carr, R. A.; Chessari, G.; Congreve, M.; Frederickson, M.; Folmer,

R. H. A.; Geschwindner, S.; Koether, G.; Kolmodin, K.; Krumrine, J.; Mauger, R. C.; Murray, C. W.; Olsson, L. L.; Patel, S.; Spear, N.; Tian, G., Application of Fragment-Based Lead Generation to the Discovery of Novel, Cyclic Amidine &#x03B2;-Secretase Inhibitors with Nanomolar Potency, Cellular Activity, and High Ligand Efficiency. *J. Med. Chem.* **2007**, 50, (24), 5912-5925.

30. Coburn, C. A.; Stachel, S. J.; Jones, K. G.; Steele, T. G.; Rush, D. M.; DiMuzio, J.; Pietrak, B. L.; Lai, M.-T.; Huang, Q.; Lineberger, J.; Jin, L.; Munshi, S.; Katharine Holloway, M.; Espeseth, A.; Simon, A.; Hazuda, D.; Graham, S. L.; Vacca, J. P., BACE-1 inhibition by a series of Y[CH_2NH] reduced amide isosteres. *Bioorg. Med. Chem. Lett.* **2006**, 16, (14), 3635-3638.

31. Stauffer, S. R.; Stanton, M. G.; Gregro, A. R.; Steinbeiser, M. A.; Shaffer, J. R.; Nantermet, P. G.; Barrow, J. C.; Rittle, K. E.; Collusi, D.; Espeseth, A. S.; Lai, M.-T.; Pietrak, B. L.; Holloway, M. K.; McGaughey, G. B.; Munshi, S. K.; Hochman, J. H.; Simon, A. J.; Selnick, H. G.; Graham, S. L.; Vacca, J. P., Discovery and SAR of isonicotinamide BACE-1 inhibitors that bind [beta]-secretase in a N-terminal 10s-loop down conformation. *Bioorg. Med. Chem. Lett.* **2007**, 17, (6), 1788-1792.

32. Oefner, C.; Binggeli, A.; Breu, V.; Bur, D.; Clozel, J.-P.; D'Arcy, A.; Dorn, A.; Fischli, W.; Grüninger, F.; Güller, R.; Hirth, G.; Märki, H. P.; Mathews, S.; Müller, M.; Ridley, R. G.; Stadler, H.; Vieira, E.; Wilhelm, M.; Winkler, F. K.; Wostl, W., Renin inhibition by substituted piperidines: a novel paradigm for the inhibition of monomeric aspartic proteinases? *Chem. Biol.* **1999**, 6, (3), 127-131.

33. Vieira, E.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Güller, R.; Hirth, G.; Märki, H. P.; Müller, M.; Oefner, C.; Scalone, M.; Stadler, H.; Wihelm, M.; Wostl, W., Substituted Piperidines - Highly Potent Renin Inhibitors due to Induced Fit Adaptation of the Active Site. *Bioorg. Med. Chem. Lett.* **1999**, 9, (10), 1397-1402.

34. Güller, R.; Binggeli, A.; Breu, V.; Bur, D.; Fischli, W.; Hirth, G.; Jenny, C.; Kansy, M.; Montavon, F.; Müller, M.; Oefner, C.; Stadler, H.; Vieira, E.; Wilhelm, M.; Wostl, W.; Märki, H. P., Piperidine-Renin Inhibitors Compounds with Improved Physicochemical Properties. *Bioorg. Med. Chem. Lett.* **1999**, 9, (10), 1403-1408.

35. Prade, L.; Jones, A. F.; Boss, C.; Richard-Bildstein, S.; Meyer, S.; Binkert, C.; Bur, D., X-ray Structure of Plasmepsin II Complexed with a Potent Achiral Inhibitor. *J. Biol. Chem.* **2005**, 280, (25), 23837-23843.

36. Carcache, D. A.; Hörtner, S. R.; Bertogg, A.; Binkert, C.; Bur, D.; Märki, H. P.; Dorn, A.; Diederich, F., De Novo Design, Synthesis, and In Vitro Evaluation of a New Class

of Nonpeptidic Inhibitors of the Malarial Enzyme Plasmepsin II. *ChemBioChem* **2002**, 3, (11), 1137-1141.

37. Hof, F.; Schütz, A.; Fäh, C.; Meyer, S.; Bur, D.; Liu, J.; Goldberg, D. E.; Diederich, F., Starving the Malaria Parasite: Inhibitors Active against the Aspartic Proteases Plasmepsins I, II, and IV. *Angew. Chem., Int. Ed.* **2006**, 45, (13), 2138-2141.
38. Boss, C.; Corminboeuf, O.; Grisostomi, C.; Meyer, S.; Jones, A. F.; Prade, L.; Binkert, C.; Fischli, W.; Weller, T.; Bur, D., Achiral, Cheap, and Potent Inhibitors of Plasmepsins I, II, and IV. *ChemMedChem* **2006**, 1, (12), 1341-1345.
39. John, V.; Beck, J. P.; Bienkowski, M. J.; Sinha, S.; Heinrikson, R. L., Human b-Secretase (BACE) and BACE Inhibitors. *J. Med. Chem.* **2003**, 46, (22), 4625-4630.
40. Blum, A.; Böttcher, J.; Heine, A.; Klebe, G.; Diederich, W. E., Structure-Guided Design of C_2-Symmetric HIV-1 Protease Inhibitors Based on a Pyrrolidine Scaffold. *J. Med. Chem.* **2008**, 51, (7), 2078-2087.
41. Czodrowski, P.; Sotriffer, C. A.; Klebe, G., Atypical Protonation States in the Active Site of HIV-1 Protease: A Computational Study. *J. Chem. Inf. Model.* **2007**, 47, (4), 1590-1598.
42. Koščová, S.; Buděšínský, M.; Hodačová, J., A Facile Synthesis of Selectively Protected Linear Oligoamines. *Collect. Czech. Chem. Commun.* **2003**, 68, (4), 744-750.
43. Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A., Crystal Structure of HIV-1 Protease in Complex with VX-478, a Potent and Orally Bioavailable Inhibitor of the Enzyme. *J. Am. Chem. Soc.* **1995**, 117, (3), 1181-1182.
44. Taylor, A.; Brown, D. P.; Kadam, S.; Maus, M.; Kohlbrenner, W. E.; Weigl, D.; Turon, M. C.; Katz, L., High-level expression and purification of mature HIV-1 protease in *Escherichia coli* under control of the *araBAD* promoter. *Appl. Microbiol. Biotechnol.* **1992**, 37, (2), 205-210.
45. Toth, M. V.; Marshall, G. R., A simple, continuous fluorometric assay for HIV protease. *Int. J. Pept. Protein Res.* **1990**, 36, (6), 544-550.
46. Hill, J.; Tyas, L.; Phylip, L. H.; Kay, J.; Dunn, B. M.; Berry, C., High level expression and characterisation of Plasmepsin II, an aspartic proteinase from *Plasmodium falciparum*. *FEBS Lett.* **1994**, 352, (2), 155-158.
47. Wang, G. T.; Chung, C. C.; Holzman, T. F.; Krafft, G. A., A Continuous Fluorescence Assay of Renin Activity. *Anal. Biochem.* **1993**, 210, (2), 351-359.

48. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode. In *Methods Enzymol.*, Carter Jr., C. W., Ed. Academic Press: 1997; Vol. 276, pp 307-326.

49. Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2004**, 60, (3), 432-438.

50. Sheldrick, G. M.; Schneider, T. R., *SHELXL*: High-resolution refinement. In *Methods Enzymol.*, Charles, W. C. J.Robert, M. S., Eds. Academic Press: 1997; Vol. 277, pp 319-343.

51. Emsley, P.; Cowtan, K., *Coot*: model-building tools for molecular graphics. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2004**, 60, (12), 2126-2132.

7. Chasing Binding Modes in HIV Protease: From Seemingly Perturbed to Seemingly Relaxed Pose Without Altering Affinity *

7.1 Introduction

Inhibition of HIV protease has been successfully established as a therapeutic strategy to combat AIDS.[1, 2] Hardly any other drug target has been studied with similar intense by structural biology and medicinal chemistry over the last 20 years. Nowadays, rational approaches to drug discovery heavily exploit crystal structure analyses to learn about binding modes and the effectiveness of protein-ligand binding in order to draw conclusions for further drug optimization.[3] Innumerable reports discuss binding modes in terms of protein-ligand complementarity and frequently the argument is used that a ligand “sits perfectly in the binding pocket” or the structure “suggests optimal mutual complementarity”. This observation is translated into arguments to explain effectiveness of binding. However, is this intuitively very welcome concept of a perfect-fit-to-high-affinity correlation really justified?

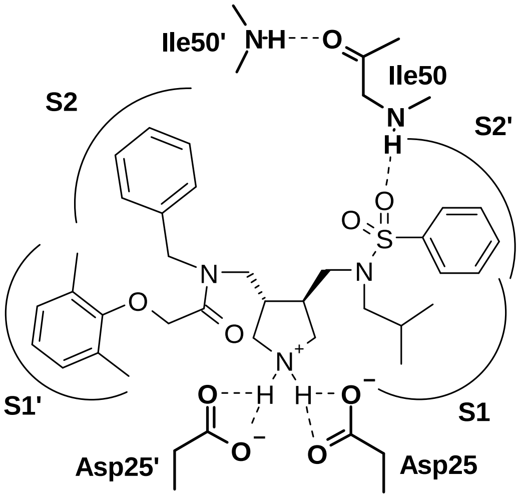
In the present report we want to discuss the binding of five related HIV protease inhibitors. Instinctively, one would argue that for the first in the series all goes wrong that could go wrong. Both, protein and ligand should feel quite “unhappy” with the achieved binding mode. For the last example in the series, a nice and relaxed binding mode is suggested and both interaction partners should feel “comfortable”. However, with respect to affinity, hardly any difference could be detected.

---

* Taken from a manuscript in preparation, Andreas Blum, Jark Böttcher, Stefanie Dörr, Andreas Heine, Gerhard Klebe, Wibke E. Diederich. Chasing binding modes in HIV Protease: From seemingly perturbed to seemingly relaxed pose without altering affinity.

7.2 Results and Discussion

Recently, we introduced a novel privileged skeleton to successfully address the catalytic dyad of HIV protease, a C_2-symmetric viral protease.[4, 5] In its core, the novel compound class features a pyrrolidine moiety (**1**, Scheme 1). Confirmed by crystal structure analysis, the basic nitrogen binds most likely in double protonated and positively charged state at the pivotal position between both aspartates.[6] Next to the central building block but spaced by two methylene amino linkers, the inhibitor skeleton has been decorated with two acceptor functionalities planned to interact via the unique structural water in HIV protease with the backbone NH groups of Ile50 and Ile50' in the flap region. Further hydrophobic substituents connected to the linking amide nitrogens and terminal to the carboxy or sulfone ends were thought to address the four specificity pockets S2, S1, S1', and S2' of the protease. Although the inhibitor skeleton of **1** had been thoroughly designed, the crystal structure in complex with the protease provided a surprising binding mode (Scheme 1).

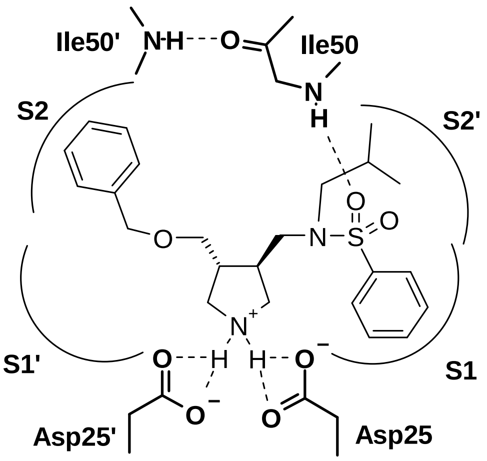


1

Scheme 1: Molecular structure and schematic representation of the observed binding modes of **1** in complex with the HIV-1 protease.

The inhibitor repells the structural water molecule from the binding site and uses its sulfone group to directly form a hydrogen bond to the NH of Ile50 in the flap. The carbonyl oxygen at the opposite side remains uncoordinated without forming any H-bond to the protein. The unbalanced H-bonding inventory concerning the Ile50'NH leads to a perturbance of the flap loop thus resulting in a direct contact with the neighbouring flap. Considering the well-planned and accurately balanced occupancy of the four specificity pockets, the inhibitor only addresses S1 and S2' properly whereas S2 remains virtually unoccupied and S1' appears overcrowded by the sterically demanding *o,o'*-dimethylphenoxy group. The structure even suggests that the latter bulky substituent remains partly solvent-exposed. It sterically interferes with the flap loop likely causing some of its perturbed spatial orientation. In summary, the crystal structure of **1** implies that the designed inhibitor should feel rather "unhappy" in the protein binding pocket. The more the binding affinity of K_i = 1.5 μM, measured for the racemate, is surprising.

In order to probe the hypothesis whether the bulky and obviously misplaced *o,o'*-dimethylphenoxy substituent is responsible for the perturbed binding mode, a model compound lacking this sterically demanding side chain has been synthesized. This goal of preparing an inhibitor with only three sidechains could be accomplished by replacing the amide nitrogen by an ether oxygen. The obtained compound **2** shows a K_i = 52 μM (again determined for the racemate) against the protease. Interestingly, its crystal structure provides another surprise (Scheme 2).

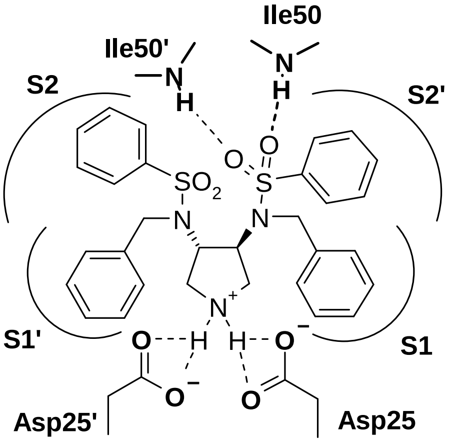


2

Scheme 2: Molecular structure and schematic representation of the observed binding modes of **2** in complex with the HIV-1 protease.

Still, the structural water is repelled from the binding site and the sulfonyl group forms a direct hydrogen bond to Ile50NH. The other flap loop remains in its perturbed geometry establishing a direct contact to the facing C_2-symmetrical counterpart. Occupancy of the S1 site appears to be similar to the complex with **1**, even though **2** places its benzene sulfonamide group instead of the *iso*-butyl substituent into this pocket. The unsatisfactory occupancy of the S2 pocket remains quite similar compared to the previous complex. Finally, the *iso*-butyl substituent of **2** is hosted in a cavity between S2' and S1' leaving both pockets only partially filled. In conclusion, the bulky *o,o'*-dimethylphenoxy substituent is not causing the perturbed binding mode. More likely the stereochemical arrangement of the five-membered ring and the attached SO_2 group are the determinant elements inducing the observed binding mode.

As a consequence of this observation, we anticipated to move the two branching substituents on either sides of the heterocycle closer together. At the pyrrolidine skeleton the methylene amino groups were replaced by amino groups to furnish the core structure of **3**. Furthermore, we decided to move to a C_2-symmetrical inhibitor skeleton to better match with the imposed constraints of the protein binding pocket.[7] Decorated with a benzyl moiety at nitrogen and a benzene sulfone group at both terminal ends, the produced inhibitor experiences a binding affinity of K_i = 2.2 μM (enantiomerically pure). Even though the inhibitor was endowed with C_2 symmetry, the crystal structure shows deviations from a proper C_2-symmetrical binding mode (Scheme 3).

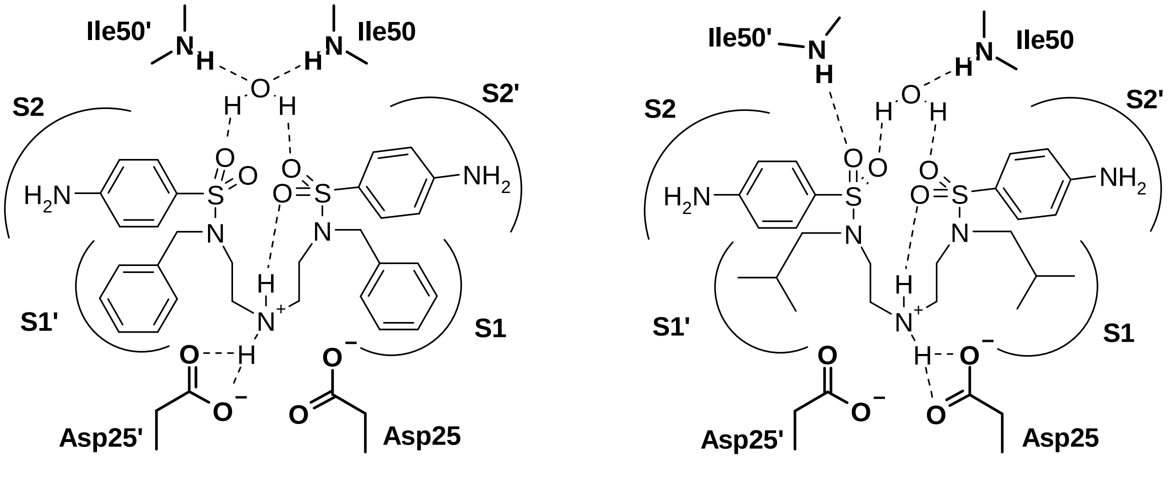


3

Scheme 3: Molecular structure and schematic representation of the observed binding modes of **3** in complex with the HIV-1 protease.

Besides this unexpected observation it provides some further surprises: The structural water is still repelled from the binding site. One of the sulfone oxygens of one sulfonamide group interacts with Ile50NH and Ile50’NH. The second oxygen of the same group forms an H-bond to Ile50’NH. In contrast, the facing SO_2 group remains uninvolved in any polar interactions. The orientation of the flap loops appears rather unperturbed. The phenyl substituents on both ends of the inhibitor are nicely hosted in the four recognition pockets. The *N*-benzyl groups are found nearly C_2-symmetrically in S1 and S1’ whereas the benzene sulfone moieties are placed into the S2 and S2’ sites.

Finally, we decided to cut the C3 to C4 bond of the five-membered pyrrolidine moiety resulting in symmetrical open chain compounds.[8] The central basic nitrogen and both sulfonamide nitrogens are bridged by an ethylene linker on either side. Two derivatives **4** and **5** were prepared both featuring *p*-amino benzene sulfone groups. At the amide nitrogens either a benzyl (**4**) or an *iso*-butyl group (**5**) were examined. Both open chain derivatives inhibit the protease with low micromolar affinity (**4**: K_i = 9.6 μM, **5**: K_i = 0.9 μM). With both ligands a crystal structure could be determined. Interestingly, they exhibit a binding mode with bound structural water (Scheme 4).



4

5

Scheme 4: Molecular structures and schematic representation of the observed binding modes of the inhibitors **4** and **5** in complex with the HIV-1 protease.

Both SO_2 groups involve one of their oxygens in a direct hydrogen bond to this water. The basic nitrogen is placed between the two aspartates. However, it does not occupy the pivotal position between both residues. The structures suggest that they preferentially interact with one of the two aspartates. Interestingly, a second contact, likely performed by the other polar hydrogen at the basic central nitrogen, is formed via an intramolecular H-bond to one of the sulfone oxygens. In the complex with **4**, the corresponding oxygen at the other SO_2 group remains unsatisfied. In the *iso*-butyl derivative **5**, the second sulfone oxygen is involved in a hydrogen bond to Ile50’NH. As a consequence, the tetrahedral arrangement around the structural water is slightly perturbed. Apart from these differences both inhibitors occupy, similar to **3**, all four recognition pockets. As in **3**, the benzyl (or *iso*-butyl) substituents are placed into the S1/S1’ pockets and the *p*-amino benzene sulfone moieties are hosted in S2/S2’. There they form hydrogen bonds to Asp30 and Asp30’, located at the far end of these pockets.

7.3 Summary and Conclusion

What conclusions can be drawn from this study? On first sight, **1** with a unsymmetrically decorated 3,4-dimethyleneamino pyrrolidine core shows an “ugly” and rather perturbed binding mode. The protein appears perturbed next to the bound inhibitor and the occupancy of the four binding pockets appears quite unsatisfactory (Figure 1G). Nevertheless, **1** achieves single digit micromolar inhibition. To trace possible reasons for this at first glance “perturbed” binding mode, the inhibitor **2** with only three substituents was synthesized. It looses in affinity by a factor of ~30. Despite these changes, **2** adopts a binding mode which even suggests a more unpleasant binding. In consequence, the size of the core fragment and the mutual distance between the branching side chains attached to address S2, S1, S1’ and S2’ were made responsible for the ”ugly” binding modes of **1** and **2**. Moving the side chains closer together and producing a C_2-symmetrical inhibitor results in a more “convincing” binding mode. The four specificity pockets are nicely filled, however, the structural water is still repelled from the binding pocket and deviations from a C_2-symmetrical binding are detected. Even though the binding mode appears intuitively more “relaxed” and the ligand exhibits better shape complementarity with the protein, the binding affinity remains in the one digit micromolar range (Figure 1H). Possibly, as one of the SO_2 groups remains uninvolved in any polar contacts, a rather high prize has to be paid for this unbalanced solvation/desolvation inventory. Finally, we opened the central pyrrolidine core to endow the inhibitor with a possibly required adaptivity to address properly the binding pocket.

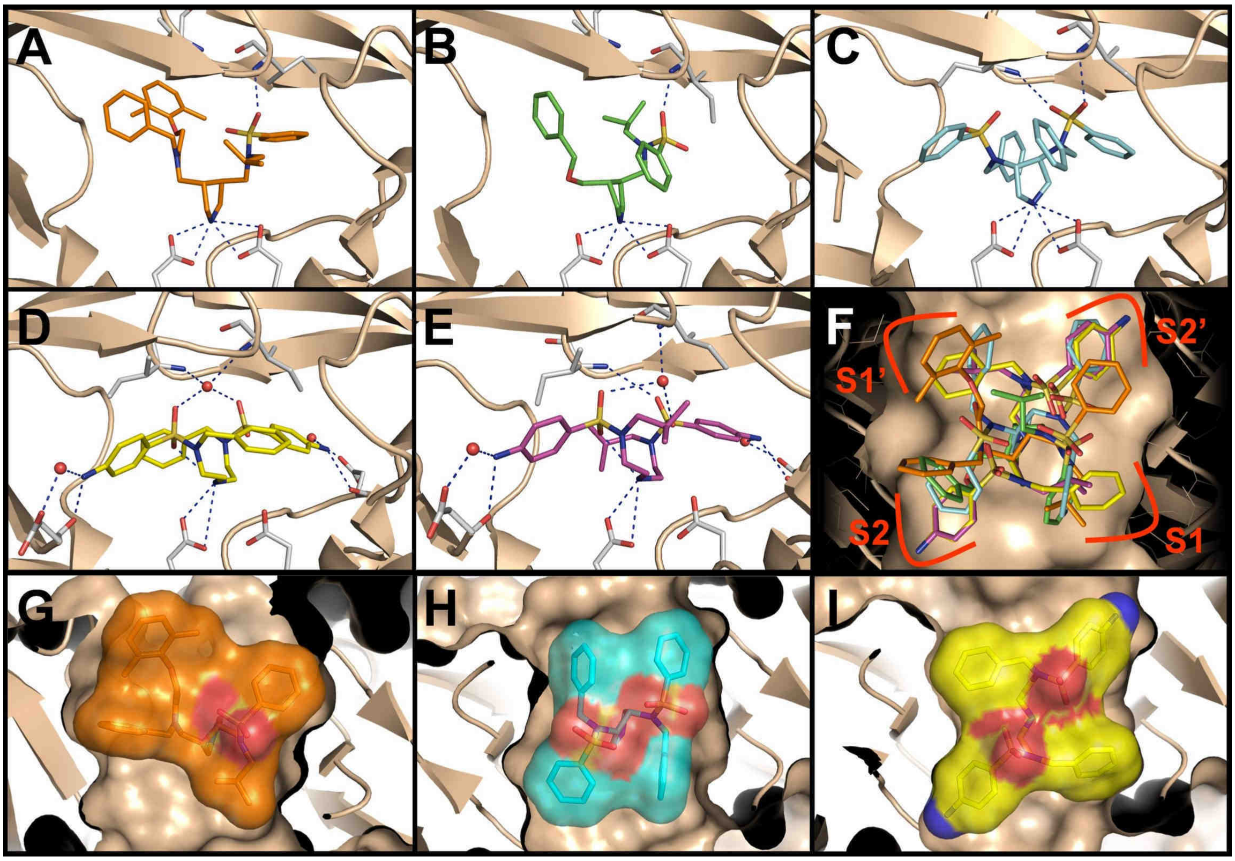


Figure 1: Crystallographically observed binding modes of **1**-**5** in HIV protease.[9] In **A** – **E** a side-on view is given showing the hydrogen-bonding pattern of **1** – **5** with the protein. **F** shows a view from the top with the flap region clipped off; all five inhibitors are superimposed and the four specificity pockets are indicated. In **G** – **I** the same top view is used on **1**, **3** and **4** as in **F**, showing the space filling and the achieved occupancy in the four subpockets. The ligand's surface color corresponds to the color used for the ligand skeleton in **A**, **C** and **D**, surface portions next to sulphur, oxygen and nitrogen are color-coded by atom-type. (**A**, **G**): **1** (orange) binds to the enzyme repelling the flap water. The bulky *o,o'*-dimethylphenoxy substituent penetrates only partly into S1' and seems to perturb the flap region. S2 seems to remain nearly entirely unoccupied. (**B**): **2** (green) fills the S2' and S1' site quite inefficiently, also here the flap water is discarded from the complex. (**C**, **H**): **3** (cyan) binds nearly C_2 symmetrically and places the benzene sulfone groups in S2 and S2' whereas the *N*-benzyl moieties are found in S1, S1'. The flap water is also repelled from the complex. (**D**, **I**): **4** (yellow) places its aminobenzene sulfone moieties in S2, S2' and the *N*-benzyl substituents occupy the S1, S1' pocket. Both SO_2 groups form a hydrogen bond to the flap water present in the complex. (**E**): **5** (magenta) binds very similarly to the protease, the *N*-*iso*-butyl substituents are placed into the S1, S1' pockets. (**F**): Superposition of **1**-**5** (**1** orange, **2** green, **3** cyan, **4** yellow, **5** magenta) shows the deviating distances between the acceptor groups thought in the design to address the flap water. However, in the observed binding mode they are placed in rather different positions in the catalytic site.

Interestingly enough, the structural water returns back into the binding pocket. Both acceptor functionalities, represented by the two SO_2 groups, are properly involved in mediating contacts to both flaps. The distance between the two sulfonamide nitrogens at the branching positions is the longest in these complexes (5.2 Å). In the pyrrolidine derivatives **1** and **2** bearing the methylene linkers, this distance shrinks to 4.3 Å (**1**) and 5.0 Å (**2**) respectively, whereas **3** brings the branching positions with 3.7 Å closest together. Hence, the SO_2 groups (in **3** – **5** or SO_2/C=O in **1**) are spatially placed differently. Obviously, this leads to the deviating addressing of the flap region (Figure. 1F). In **4** and **5**, the achieved occupancy of the four recognition pockets S2, S1, S1' and S2' appears very satisfactory (Figure 1I). It is tempting to describe the binding modes of the open-chain inhibitors **4** and **5** as "almost perfect" and hardly perturbed. Nevertheless, their binding affinity still remains in the one digit micromolar range, even though they are decorated at the benzene sulfone termini with amino groups that get involved in further hydrogen bonds to the protein. This fact has to be faced with the corresponding amino-substituted ligand from the series of **3** which shows a ten-fold improved affinity. Most likely, the open-chain derivatives pay to some extent an entropic prize as they have to freeze an increased number of rotational degrees of freedom. The pronounced pre-organization of the structural arrangement of the side chains required to fill the subpockets of the protein is not satisfactorily given in these ligands. In summary, this comparative study shows that tracing the "ugly" and "beauty" of binding modes is not a question of energy and symmetry. Also obvious arguments concerning the perturbance of parts of the protein structure or rough estimates in terms of degrees of binding pocket filling are not convincing. Apparently, the picture is much more complex and results as a detailed composition of multiple contributions that finally determine affinity and binding geometry of a protein-ligand complex.

7.4 References

1. Pomerantz, R. J.; Horn, D. L., Twenty years of therapy for HIV-1 infection. *Nat. Med.* **2003**, 9, (7), 867-873.
2. Flexner, C., HIV drug development: the next 25 years. *Nat. Rev. Drug Discovery* **2007**, 6, (12), 959-966.
3. Babine, R. E.; Bender, S. L., Molecular Recognition of Protein-Ligand Complexes: Applications to Drug Design. *Chem. Rev.* **1997**, 97, (5), 1359-1472.

4. Specker, E.; Böttcher, J.; Lilie, H.; Heine, A.; Schoop, A.; Müller, G.; Griebenow, N.; Klebe, G., An Old Target Revisited: Two New Privileged Skeletons and an Unexpected Binding Mode For HIV-protease Inhibitors. *Angew. Chem., Int. Ed.* **2005**, 44, (20), 3140-4.
5. Specker, E.; Böttcher, J.; Brass, S.; Heine, A.; Lilie, H.; Schoop, A.; Muller, G.; Griebenow, N.; Klebe, G., Unexpected Novel Binding Mode of Pyrrolidine-Based Aspartyl Protease Inhibitors: Design, Synthesis and Crystal Structure in Complex with HIV Protease. *ChemMedChem* **2006**, 1, (1), 106-17.
6. Czodrowski, P.; Sotriffer, C. A.; Klebe, G., Atypical Protonation States in the Active Site of HIV-1 Protease: A Computational Study. *J. Chem. Inf. Model.* **2007**, 47, (4), 1590-1598.
7. Blum, A.; Böttcher, J.; Heine, A.; Klebe, G.; Diederich, W. E., Structure-Guided Design of C_2-Symmetric HIV-1 Protease Inhibitors Based on a Pyrrolidine Scaffold. *J. Med. Chem.* **2008**, 51, (7), 2078-2087.
8. Blum, A.; Böttcher, J.; Sammet, B.; Luksch, T.; Heine, A.; Klebe, G.; Diederich, W. E., Achiral Oligoamines as Versatile Tool for the Development of Aspartic Protease Inhibitors. *Bioorg. Med. Chem.* **in press**
9. The structure data of **1-5** are available from the Protein Data Bank (PDB) under the codes 1XL2, 3BHE, 2PQZ, 3BGB and 3BGC, respectively.

8. Summary/Zusammenfassung

8.1 Summary

Infections with the HI virus, which inevitably lead to the development of AIDS, are still among the most serious global health problems causing more than 2.5 million deaths per year. In the patho-physiological processes of this pandemic, HIV protease has proven to be an invaluable drug target due to its essential role in the virus' replication process. Currently, nine HIV protease inhibitors are approved by the FDA and are being applied within the *highly active antiretroviral therapy* (HAART). However, the occurrence of multi-drug resistant HIV variants diminishes the efficacy of all approved protease inhibitors. In addition to that, the pharmacokinetic properties of the mostly peptidomimetic compounds are often not optimal and the chemical access to these complex and chiral inhibitors is challenging and resource-intensive. This demands a continuous and persistent search for new inhibitors of the protease. Novel inhibitors should preferably exhibit a different mutation profile compared to that of the marketed drugs to circumvent the development of cross resistance. As novel promising scaffold for HIV protease inhibition, pyrrolidine-derived inhibitors have recently been reported. However, the reported compounds exhibited only a moderate affinity and the thorough investigation of the X-ray structure of the protease in complex with the most potent derivative revealed an unsatisfactory occupation of the sub-pockets and an asymmetric flap interaction, resulting in strong deformations of the protein structure. In this thesis, the stepwise improvement of this compound class to potent inhibitors of wildtype as well as selected mutant proteases utilizing rational drug discovery methods is reported.

The in-depth analysis of the crystal structure of a (*rac*)-3,4-dimethyleneamino-pyrrolidine in complex with HIV-1 protease provided a promising starting point for further inhibitor design: As initial approach symmetric pyrrolidine-diesters possessing the same stereochemistry as indicated in the cocrystal structure of the initial 3,4-dimethyleneamino-pyrrolidines were synthesized following a chiral-pool approach. The structure-guided design

of these inhibitors is described in Chapter 2. The compounds were designed taking the essential pharmacophore requirements for a HIV protease inhibitor into account: In addition to the cyclic amino functionality addressing the catalytic dyad, the compounds were equipped with hydrophobic moieties intended to occupy the protease's subpockets. Furthermore, two acceptor groups allowing appropriate flap-interactions were introduced. Compared to the 3,4-dimethyleneamino-pyrrolidine derivatives, the acceptor groups are attached closer to the central pyrrolidine ring to address the flap in a geometrically better-suited way. The most potent compounds of the series achieve one-digit micromolar inhibition towards wild type as well as two mutant proteases (Ile50Val and Ile84Val). The cocrystal structure of one derivative in complex with the Ile84Val HIV protease could be determined with a resolution of 1,82 Å. To our surprise, this complex structure revealed that two inhibitor molecules are bound in the large active site cavity comprising an area encompassed by the catalytic dyad and the flaps in the open conformation. This crystal structure is the first HIV protease cocrystal structure in which the open-flap conformation of the enzyme is stabilized by an inhibitor that concomitantly addresses the catalytic dyad. Thus, this cocrystal structure provides a valuable novel starting point for the further development of HIV protease inhibitors possessing a different mode of binding compared to known drugs.

In Chapter 3, as an alternative approach towards HIV protease inhibitors, the development of symmetric 3,4-bis *N*-alkyl sulfonamide-pyrrolidines is described. Similar to the diesters, these inhibitors also possess a 3*S*,4*S*-disubstituted pyrrolidine, which addresses the catalytic dyad via its secondary amino functionality. Utilizing the analogous 3,4-diaminopyrrolidine as core structure, arylsulfonyl moieties were introduced not only intended to occupy the S2- and S2'-subpockets but also to address the flap region of the enzyme. Via alkylation of the secondary sulfonamides, two additional hydrophobic moieties were introduced designed to address the S1- and S1'-subsites of the enzyme. The compounds are accessible following the same chiral pool approach, however, utilizing D-(-)-tartaric acid thus rendering the putative inhibitors in a 7- or 8-step synthesis in high overall yield. The initial lead structure possessing benzene sulfonamide groups and benzyl substituents exhibited a K_i of 2.2 µM and the X-ray structure in complex with the protease was determined with a resolution of 1.55 Å. Surprisingly, the C_2-symmetric inhibitor adopts an asymmetric binding mode in the complex with the C_2-symmetric protein. This crystal structure enabled the rational design of a second series of inhibitors and revealed three promising symmetric substitution patterns for further lead optimization: (A) Elongation of the P1/P1'-benzyl moieties with hydrophobic substituents in para-position, (B) *ortho*-substitution at the P2/P2'-phenyl ring systems, and (C) *para*-substitution at the P2/P2'-phenyl moieties. All three strategies were pursued and resulted in inhibitors with improved affinities up to 260 nM. To elucidate the underlying factors accounting for the SAR, consecutively the crystal structures of four representatives, at least one of each modification type, in complex with HIV protease were determined (1.5-2.3Å resolution). These structures provided deeper insights into the protein–ligand interactions and the underlying principles of the SAR thus enabling us to choose the most promising combination of substituents in the next design cycle. As *ortho*-substitution at the P2/P2'-phenyl led to a slightly different binding mode, *ortho*-substitution at the P2-residue seemed to be less attractive for the further inhibitor design. As substituents for the combined inhibitor, a trifluoromethyl group as P1/P1'(strategy A) and a carboxamido moiety as P2/P2'-substituent (strategy B) were selected. The combination rendered a final inhibitor showing a significantly improved affinity of K_i = 74 nM. The cocrystal structure of this inhibitor in complex with the HIV protease (1.55Å resolution) confirmed the successful application of the pursued optimization strategy.

The extremely successful development of the symmetric pyrrolidine-3,4-bis-*N*-benzyl-sulfonamides as inhibitors of the wildtype protease suggested also an investigation of their potential towards active site mutants. In Chapter 4 the influence of the active site mutations Ile50Val and Ile84Val on these inhibitors is investigated by structural and kinetic analysis. Both single-point mutants were selected due to their importance and the remarkable differences in binding of the pyrrolidine-based inhibitors, particularly in the contact area next to Ile50 and Ile84. Whereas the Ile50Val mutation leads to a significant decrease in affinity for all compounds in this series, they retain or even show increased affinity towards the crucial Ile84Val mutation. By detailed analysis of the crystal structures of two representatives in complex with wild-type and mutant proteases (1.58-1.92Å), we were able to elucidate the structural basis of this phenomenon. In case of the Ile50Val mutant the reduced affinity towards PR_{I50V} could be rationalized in terms of reduced Van der Waals contacts, an observation already described in literature for the approved inhibitor Amprenavir. In case of the Ile84Val mutation multiple contributions could be identified for the improved affinities. In addition to amplified Van der Waals contacts and relief of steric strain likely given in the wild-type complexes, enhanced polar interactions mediated by an interstitial water molecule that fills the gap created by the mutational exchange could be observed. The observation of the change in physicochemical properties of the binding pocket upon Ile84Val mutation can be exploited for further inhibitor design.

In Chapter 5 a remarkably observation, which was made during the further development of the symmetrically disubstituted 3,4-amino-pyrrolidines as HIV-1 protease inhibitors, is investigated in depth. The determination of the binding affinities towards the Ile50Val and Ile84Val mutants of inhibitors bearing smaller *N*-alkyl substituents revealed a selectivity profile not being explicable with the initial SAR: In contrast to the *N*-benzyl-substituted derivatives, a reduced affinity is observed towards both active site mutants. By cocrystallization of the most potent derivative of a small series with HIV-1 protease, astonishingly two different crystal forms, $P2_12_12_1$ and $P6_122$, were obtained. Structural analysis revealed two completely different binding modes, the interaction of the pyrrolidine nitrogen atom to the catalytic aspartates being the only similarity. The binding mode observed in the complex in space group $P2_12_12_1$ closely resembles the one observed for all examples of the *N*-benzyl series (Chapter 3). However, due to the similar surface complementarities of these complexes, particularly in the contact area to Ile84Val, this binding mode does not provide a convincing explanation for the deviating selectivity profile. The binding mode observed in $P6_122$ resembles more that one adopted by all approved peptidomimetic inhibitors, which are known to show high susceptibility towards the Ile84Val substitution. It can not be predicted which of the two binding orientations is be preferred under biological conditions. The nature of the substituent and the different binding-pocket shape of the mutant enzymes might induce different preferences thus explaining the overall non-uniform SAR observed for derivatives bearing smaller *N*-alkyl substituents.

Encouraged by the successful utilization of cyclic secondary amines as anchoring group in the development of HIV protease inhibitors, this strategy was expanded into a general approach for lead structure identification for aspartic proteases in Chapter 6. Due to the important role that aspartic proteases play in many patho-physiological processes, they have already been targeted intensively by modern drug development. However, up to now, only for two family members, renin and HIV protease, inhibitors have been approved for disease therapy. The inhibitor development, so far mostly guided by mimicking the natural peptide substrates, resulted in very potent inhibitors for several targets, but the pharmacokinetic properties of these compounds were often not optimal. For a more efficient development of novel aspartic protease inhibitors, an easily accessible and achiral core structure would be preferable. Therefore, an initial library comprising eleven inhibitors based on easily accessible achiral linear oligoamines was developed and screened against six selected aspartic proteases (HIV-1 protease, plasmepsin II, plasmepsin IV, renin, BACE-1, and pepsin). The compounds bear a central secondary amino function aimed to address the catalytic dyad, acceptor groups to address the flap regions of the respective target enzyme, and hydrophobic moieties to address the specificity pockets. Several hits could be identified, among them selective as well as rather promiscuous inhibitors. The design concept was consecutively confirmed by determination of the crystal structure of two derivatives in complex with HIV-1 protease (1.80/1.90Å). The binding mode exhibits high similarity to the binding orientation of substrates as well as to that of peptidomimetic inhibitors. Using this information, a generalization of this binding situation to other aspartic proteases appears reasonable thus providing a first insight into the observed structure-activity relationships. Hits from this oligoamine library can be easily exploited as suitable starting point for the further development of aspartic protease inhibitors.

In Chapter 7 the binding modes of the structurally related secondary amine-based HIV protease inhibitors presented in Chapters 1-6 are discussed with respect to their affinity. It is commonly assumed that a perfect protein-ligand complementarity is one of the main driving forces for a high binding affinity. The asymmetrically decorated 3,4-dimethyleneamino-pyrrolidine (Chapter 1.5) achieves one-digit micromolar inhibition, even though the protein structure appears to be perturbed next to the bound inhibitor and the occupancy of the specificity pockets appears to be quite unsatisfactory. In contrast, the initial lead structure of the 3,4-bis-*N*-alkylsulfonamides (Chapter 3) show approximately the same inhibitory activity, however, without protein deformations and a virtually good occupation of the selectivity pockets as revealed by X-ray crystallography. The oligoamine derivatives presented in Chapter 6 show a high degree of adaptivity and possess a binding mode with a nearly ideal protein-ligand complementarity. However, their binding affinity still remains in the one-digit micromolar range. In case of the 3,4-bis-*N*-alkylsulfonamides one of the sulfonyl groups remains uninvolved in any polar contacts and the resultant negative solvation/desolvation inventory could be the reason for the comparably low affinity with respect to the better surface complementarity. In case of the oligoamines, an entropic price has to be paid for the freezing of an increased number of rotational degrees of freedom. This points out that the obvious arguments concerning the perturbance of parts of the protein structure or rough estimates in terms of degrees of binding pocket filling are not sufficient to explain a certain affinity. The picture is much more complex and the final affinity is a result of a detailed composition of multiple contributions certainly involving the binding geometry of a protein-ligand complex.

In summary, this thesis clearly points out that structural biology has to be integral part of drug discovery projects and, if applied, has to escort the entire optimization process. It can provide novel strategies for ligand development, e.g. by the discovery of unexpected protein-ligand interaction patterns. The discovery of a completely new class of inhibitors targeting an alternative protein conformation of HIV protease would have been undiscovered without the structural investigation of the protein-ligand complex. It has been shown that the cooperation between synthetic medicinal chemistry and structural biology on equal footing and in short time intervals can lead to a highly efficient optimization of lead structures. The detailed structural analysis of related protein-ligand complexes revealed that large affinity changes might result from very small deviations in the protein-ligand complexes. The limitation of structural biology has been shown by the identification of two complexes of identical composition exhibiting two completely different protein-inhibitor interaction patterns. However, in combination with the biological evaluation, hints for such multiple binding orientations can be gained and even provide novel opportunities to optimize ligands towards a particular binding profile. Using the knowledge gained within the development of pyrrolidine-based HIV protease inhibitors, a general approach for lead structure identification for aspartic proteases has successfully been established. The structural and kinetic information gained within the different projects display the difficulty to estimate the potency of a certain lead compound in terms of degrees of binding pocket filling and observed binding geometry.

8.2 Zusammenfassung

Im Rahmen dieser Arbeit wurden neuartige Leitstrukturen für die Inhibition von Aspartylproteasen, insbesondere der HIV-1-Protease, unter Verwendung strukturbiologischer Methoden entwickelt. Als Startpunkt diente die Kristallstruktur eines Komplexes der HIV-1-Protease mit einem 3,4-Dimethylenamino-pyrrolidin-Derivat, welches mit der zyklischen Aminofunktion die katalytische Dyade der Protease adressiert. Ausgehend von der Beobachtung starker Proteinverzerrung und einer unbefriedigenden Taschenbesetzung in dieser Struktur wurden Verbindungen entwickelt, die nur die essentiellen Pharmacophor-Erfordernisse für einen HIV-Protease-Inhibitor in sich vereinen. Die symmetrischen Pyrrolidindiester-Derivate, ausgestattet mit einer zyklischen sekundären Aminofunktion zur Adressierung der katalytischen Aspartate, zwei hydrophoben Resten für die Besetzung der

Erkennungstaschen und zwei Akzeptorfunktionen zur Adressierung der flexiblen Flapregion, zeigten einstellig mikromolare Inhibitionswerte. Die Kristallstrukturanalyse eines Derivates im Komplex mit der Protease offenbarte eine einzigartige Bindungssituation, in der zwei Inhibitormoleküle an ein Protease-Dimer in der geöffneten Konformation binden. Diese Beobachtung ermöglichte nun das rationale Design von Inhibitoren, die gezielt diese Konformation der Protease adressieren. Parallel zu diesem Ansatz wurden Verbindungen auf Basis eines 3*S*,4*S*-Diamino-pyrrolidin-Gerüstes entwickelt, die über vier Reste zur Adressierung der Subtaschen verfügen. Auch hier zeigten die ersten Verbindungen einstellig mikromolare Inhibition. Aus dem mittels Röntgenkristallographie aufgeklärten Bindungsmodus wurden drei Strategien zur Affinitätssteigerung postuliert und umgesetzt. Alle drei Strategien führten zu Inhibitoren mit einem deutlichen Affinitätsgewinn. Mindestens ein Vertreter jeder Strategie wurde anschließend im Komplex mit der HIV-Protease kristallisiert und analysiert. Aus den beobachteten Bindungsmodi wurden die beiden vielversprechendsten Strategien zur Kombination ausgewählt und in dem finalen Inhibitor vereint. Dieser wies mit 74 nM eine weitere deutliche Affinitätssteigerung auf, und die Analyse der Kokristallstruktur bestätigte den methodischen Ansatz. Mit den deutlich affineren Substanzen wurden kinetische Untersuchungen an zwei ausgewählten, klinisch relevanten Punktmutanten der HIV-Protease durchgeführt. Während ein Affinitätsverlust im Falle der Ile50Val-Mutation zu beobachten ist, zeigen die Verbindungen eine erhebliche Affinitätssteigerung im Falle der Ile84Val-Mutation. Zwei Verbindungen, bei denen dieses Phänomen besonders zu Tage tritt, wurden anschließend im Komplex mit den Protease-Varianten strukturell untersucht. Aus der Analyse ergab sich, dass im Falle der Ile84Val Mutation eine Kombination aus mehreren Faktoren zur beobachteten Affinitätssteigerung beiträgt. Besonders die Beobachtung veränderter physikochemischer Eigenschaften in der Kontaktregion zur Aminosäure 84 ist hierbei von erheblicher Relevanz für die Entwicklung zukünftiger Generationen von HIV-Protease-Inhibitoren, mit einer wahrscheinlich deutlich geringeren Empfindlichkeit gegen diese Mutation. Aus den kinetischen Studien ergab sich zudem, dass Verbindungen mit kleineren Alkylresten ein deutlich verändertes Aktivitätsprofil gegenüber den Mutanten aufweisen. Bei dem Versuch, diese Beobachtung strukturell zu untersuchen, wurden Kokristalle eines in der Zusammensetzung identischen Protein-Ligand Komplexes in zwei verschieden Raumgruppen erhalten. In der orthorhombischen Kristallform ist ein Bindungsmodus vergleichbar zu den vorher untersuchten Verbindungen zu beobachten, wohingegen in der hexagonalen Kristallform ein Bindungsmodus zu beobachten ist, der sich fundamental von diesem unterscheidet. Eine uneinheitliche Struktur-Aktivitäts-Beziehung in

Bezug auf die Protease-Varianten legt ein paralleles Vorliegen dieser beiden Bindungsmodi nahe. Anhand der in den Projekten gesammelten Erfahrungen wurde ein genereller Ansatz für die Leitstrukturfindung für Aspartylprotease-Inhibitoren auf der Basis eines Oligoamin-Grundgerüstes entwickelt und umgesetzt. In einer 11 Verbindungen umfassenden Bibliothek konnten mikromolare Inhibitoren für 6 der 7 untersuchten Proteasen identifiziert werden. Mit der Hilfe der Kokristallstrukturen zweier Derivate im Komplex mit der HIV-Protease konnte eine generelle Bindungssituation für die untersuchten Aspartylproteasen postuliert und mit Affinitätsdaten bestärkt werden. In dieser Arbeit wurden vielfältige Ansätze für die Entwicklung sekundärer Amine als Aspartylprotease-Inhibitoren erarbeitet. Besonders hervorzuheben ist die Entwicklung von Verbindungen, die an die geöffnete Konformation der HIV-Protease binden und somit die Entwicklung einer neuen Klasse von Inhibitoren ermöglichen. Zudem ist es gelungen, eine neue, potente Klasse von Inhibitoren zu etablieren, die im Vergleich zu den literaturbekannten Inhibitoren über einzigartige Eigenschaften gegenüber der Ile84Val Punktmutation verfügt. Diese Arbeit untermauert die Bedeutung, die der Strukturbiologie in der Entwicklung von biologisch aktiven Verbindungen zukommt.

Erklärung

Ich versichere, dass ich meine Dissertation

„Structure-based Development of Secondary Amines as Aspartic Protease 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 bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den . .2008

Jark Böttcher

Danksagung

- Herrn *Prof. Dr. Gerhard Klebe* danke ich für die gute Betreuung und die interessante Themenstellung meiner Doktorarbeit. Ich danke ihm für das Vertrauen, das er mir entgegengebracht hat und die Möglichkeit, die Themen eigenständig weiterentwickeln zu können. Viele Konzepte dieser Arbeit entstanden in persönlichen Gesprächen und Diskussionen, für die Sie immer Zeit gefunden haben. Der Erfolg dieser Arbeit basiert auf der außerordentlich guten Ausstattung und Organisation unserer Arbeitsgruppe. Zudem danke ich für die Möglichkeit unsere Ergebnisse auf internationalen Tagungen präsentieren zu können.

- Frau *Dr. Wibke E. Diederich* danke ich für die gelungenen Kooperationen, ohne die diese Arbeit niemals hätte durchgeführt werden können, und die permanente Hilfsbereitschaft bei der Publikation der durchgeführten Arbeiten.

- Das in dieser Arbeit behandelte Themengebiet wurde mit *Dr. Andreas Blum* gemeinsam entwickelt. Durch diese gemeinsame wissenschaftliche und freundschaftliche Zeit wurde diese Arbeit maßgeblich beeinflusst.

- Ich danke *Benedikt Sammet*, der im Rahmen seiner Diplomarbeit durch Fleiß und Ideenreichtum die Erforschung der Oligoamine als Aspartylprotease-Inhibitoren ermöglichte.

- *Steffi Dörr* trägt einen großen Anteil an dieser Arbeit, indem sie Andreas mit der Synthese diverser Inhibitoren, die in dieser Arbeit untersucht wurden, unterstützt hat.

- Ohne *Dr. Torsten Luksch* wäre ich vielleicht nie ein Teil der Arbeitsgruppe Klebe geworden. Ich danke ihm für die enge Zusammenarbeit auf dem Gebiet der Aspartylproteasen und eine schöne Zeit in unserer gemeinsamen Arbeitsgruppe.

- *Dr. Andreas Heine* danke ich für die gewissenhafte und fundierte Einführung in die Kristallographie. Seine Erfahrung hat maßgeblich zu der Qualität dieser Arbeit beigetragen.

- Ein außerordentlicher Dank gilt Herrn *Christian Sohn,* der durch seine fachmännische Betreuung den Röntgengenerator unserer Arbeitsgruppe immer einsatzbereit gehalten hat.

- Ich danke *Angela Scholz* für viele schöne Momente und die Unterstützung in vielen organisatorischen Aspekten meiner Arbeit und meines Lebens. Sie war immer eine der tragenden Säulen der Arbeitsgruppe Klebe.

- Ich danke allen Mitgliedern der AG Klebe für eine schöne Zeit und vielfältige freundschaftliche und wissenschaftliche Momente.

- Mein Dank gilt der *ZEDIRA GmbH,* insbesondere *Dr. Kai Oertel* und *Dr. Edgar Specker*, die mir Einblicke in die Transglutaminase-Forschung und industrielle Arbeitsabläufe ermöglicht haben. An dieser Stelle möchte ich auch *Ina Lindemann* für die gelungene Zusammenarbeit auf diesem Gebiet danken.

- Ich danke der Arbeitsgruppe von *Prof. Manfred Schubert-Zsilavecz* für die interessante und fruchtbare Zusammenarbeit auf dem Gebiet der nukleären Rezeptoren.

- Meiner Liebe, Frau und Freundin *Anita* möchte ich für viele schöne Momente in der Zeit meiner Doktorarbeit danken. Kocham Cię.

- Meinen Freunden, insbesondere *David Pereira Vaz,* danke ich für die Lebenskraft, die sie mir durch ihre Freundschaft geben.

- Meiner *Rugby Mannschaft* danke ich für viele schöne gemeinsame Siege, Niederlagen und Feste.

- Ein besonderer Dank gilt meinen Eltern, die mich in meinem Leben immer unterstützen und mit Rat und Tat zur Seite stehen.

Curriculum Vitae

Full name: Jark Böttcher
Date of birth: 20.05.1978
Place of birth: Winsen (Luhe)
Nationality: German

09/2004-09/2008	PhD, Philipps-University Marburg, Institute for Pharmaceutical Chemistry, Germany, Adviser: Prof. Dr. Gerhard Klebe
08/2004	Diploma, Chemistry, Philipps-University Marburg, Department of Chemistry, Germany, Adviser: Prof. Dr. Lars Oliver Essen.
03/2002-07/2002	Research internship, University of Bologna, Department of Chemistry, Italy, Adviser: Prof. Dr. Magda Monari
09/2001- 03/2001	Studies of Chemistry, Università di Bologna, Italy, EU-funded scholarship
10/2000	Vordiplom (intermediate diploma), Chemistry, Philipps University Marburg, Department of Chemistry, Germany
10/1999-10/2004	Studies of Chemistry, Philipps-Universität Marburg, Germany
09/1997 -09/1998	Zivildienst (compulsory social service), Elisabethenstift, Darmstadt, Germany
08/1988-06/1997	Abitur (majors: chemistry, social studies), Georg Büchner Schule, Darmstadt, Germany