Structure-based Inhibitor Design for the Antimalarial Target Plasmepsin

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

dem Fachbereich Pharmazie
der Philipps-Universität Marburg
vorgelegt von

Torsten Luksch
aus Wirges

Marburg/Lahn 2007

Vom Fachbereich Pharmazie der Philipps-Universität Marburg als Dissertation am 08.02.2008 angenommen.

Erstgutachter: Prof. Dr. Gerhard Klebe
Zweitgutachter: Prof. Dr. Christoph Sotriffer

Tag der mündlichen Prüfung: 08.02.2008

Die Untersuchungen zur vorliegenden Arbeit wurden auf Anregung von Herrn Prof. Dr. Gerhard Klebe am Institut für Pharmazeutische Chemie des Fachbereichs Pharmazie der Philipps-Universität Marburg in der Zeit von November 2003 bis Oktober 2007 durchgeführt.

42,195 km

Table of contents

1 Einleitung 1

  1.1 Malaria 1

    1.1.1 Aktuelle Malaria-Wirkstoffe, Resistenzentwicklungen und neue Forschungsansätze 2

    1.1.2 Protozoen der Gattung *Plasmodium* 3

    1.1.3 Der Lebenszyklus von *Plasmodium* 4

    1.1.4 Krankheitsbild (physiologische Folgen) 5

    1.1.5 Hämoglobin-Abbau 6

  1.2 Plasmepsine 7

    1.2.1 Selektivität 8

    1.2.2 Aspartylproteasen 8

    1.2.3 Plasmepsin Inhibitoren 10

  1.3 References 12

2 Aufgabenstellung 17

3 Assay, Kristallisation und Strukturbestimmung 19

  3.1 Assay 19

  3.2 Kristallisation und Strukturbestimmung 21

    3.2.1 Kokristallisation von Plm II mit Pepstatin A 21

    3.2.2 Kristallisation von Plm II mit Azabicyclo[2.2.1]Heptan-Derivaten 22

    3.2.3 Kristallisation von Plm II mit Pyrrolidin-Derivaten 24

    3.2.4 Kristallisation von Plm II mit Primaquin-Statin Kombinationsinhibitoren 25

    3.2.5 Analyse der Proteinreinheit 26

  3.3 References 29

4 Computer-aided Design and Synthesis of non-peptidic Plasmepsin II and IV inhibitors 31

  4.1 Abstract 31

  4.2 Introduction 32

  4.3 Results and Discussion 34

    4.3.1 In silico evaluation of the core structure 34

    4.3.2 Docking starting geometry for Plm IV 37

    4.3.3 Optimization of the P1 and P2’ residues 39

    4.3.4 Analysis of the docked ligands 40

    4.3.5 Synthesis and biological testing of the first design series 42

  4.4 Chemical Synthesis 47

4.5 Summary and Conclusion 49
4.6 Experimental Section 50
4.6.1 Assays and K_i determinations 50
4.6.2 Homology model of Plm IV 50
4.6.3 Molecular Dynamics Simulation 51
4.6.4 Computational Docking 51
4.6.5 Comparison of protein subpockets 52
4.7 References 53
5 The Potential of P1 site alterations in peptidomimetic protease inhibitors as suggested by virtual screening and explored by the use of C-C-Coupling reagents 57
5.1 Abstract 57
5.2 Introduction 58
5.3 Results and Discussion 60
5.3.1 Virtual screening of the plasmepsin II S1 site and design of a norstatine library61
5.3.2 Screening of inhibitors 10a-q against plasmepsin II, cathepsin D and HIV protease 63
5.3.3 Molecular dynamics simulations indicate adaptibility of the S1 pocket 63
5.4 Conclusions 68
5.5 Experimental Section 69
5.5.1 Determination of inhibition constants toward plasmepsin II 69
5.5.2 Determination of inhibition constants toward cathepsin D 69
5.5.3 Virtual screening to select optimal occupancy of the S1 pocket of plasmepsin II69
5.5.4 Molecular dynamics simulation 69
5.6 References 71
6 Pyrrolidine derivatives as Plasmepsin Inhibitors: Binding mode analysis assisted by Molecular Dynamics 75
6.1 Abstract 75
6.2 Introduction 76
6.3 Results and Discussion 76
6.3.1 Enzyme Inhibition Data 76
6.3.2 Computational binding mode generation 80
6.3.3 Molecular Dynamics Simulations 84
6.3.4 Putative binding mode generation for pyrrolidinedimethylene diamines 93
6.3.5 Relevance and implications for inhibitor design 94
6.4 Summary and conclusion 95
6.5 Materials and Methods 96

6.5.1 Assays and K_i determinations ... 96
6.5.2 Molecular Dynamics Simulation ... 96
6.6 References ... 98
7 Lead identification for the antimalarial target Plm by using Ftrees ... 101
7.1 Abstract ... 101
7.2 Introduction ... 102
7.3 Selection of the screening strategy ... 102
7.4 Results and discussion ... 103
7.4.1 Ftree similarity value at level X ... 103
7.4.2 Ftree value at level 0 ... 104
7.4.3 Synthesis ... 106
7.4.4 Biological Evaluation ... 107
7.5 Summary and conclusion ... 110
7.6 Materials and methods ... 111
7.7 Literature ... 113
8 Strukturlösung von Prephenat Dehydratase aus *Methanocaldococcus jannaschii* ... 115
8.1 Einleitung ... 115
8.1.1 Shikimat Biosyntheseweg ... 115
8.1.2 Prephenat Dehydratase ... 116
8.2 Aufgabenstellung ... 118
8.3 Kristallisation und Datensammlung von MjPDT ... 119
8.3.1 Kristallisationsoptimierung ... 119
Erzeugung erster Kristalle ... 119
Entstehung einer zweiten Kristallform ... 120
Micro-seeding und Macro-seeding ... 120
Kokristallisation ... 121
Optimierung der Cryobedingungen ... 121
8.3.2 High Throughput Kristallisation ... 122
8.3.3 Gesammelte Datensätze ... 124
8.4 Strukturlösungsversuche durch Molekularen Ersatz ... 125
8.4.1 Sequenzalignment ... 125
8.4.2 Modelle für den Molekularen Ersatz ... 126
8.4.3 Molekularer Ersatz ... 127
8.5 Experimentelle Phasen via SeMet ... 128
8.5.1 Einfügen von Methioninen durch Mutation ... 128
8.5.2 Kristallisationsoptimierung ... 128

8.5.3 Gesammelte Datensätze 129
8.5.4 Strukturlösungsversuche mit HKL2map 130
8.6 Experimentelle Phasen via Schweratomderivate 132
8.6.1 Geeignete Schweratomderivate 132
8.6.2 Soakingbedingung-Optimierung 132
8.6.3 Druckbegasung mit Xenon 133
8.7 Zusammenfassung und Ausblick 134
8.8 Referenzen 135
9 Zusammenfassung 137
10 Anhang 147
Danksagungen 147
Erklärung 148
Curriculum Vitae 149

1 Einleitung

1.1 Malaria

Abgesehen von bestimmten Alters- und Zivilisationskrankheiten stellt Malaria noch immer die weltweit am weitesten verbreitete Krankheit dar. Besonders in vielen Gebieten der Tropen und teilweise auch in den Subtropen verursacht Malaria jährlich ca. 300 - 500 Millionen Krankheitsfälle, von denen ca. 1,5 bis 2,7 Millionen tödlich verlaufen. Allein im tropischen Afrika sterben daran jährlich ca. eine Million Kinder unter fünf Jahren. Hinzu kommt, dass bezüglich der Anzahl der Krankheitsfälle in den Entwicklungsländern wahrscheinlich eine nicht unerhebliche Dunkelziffer angenommen werden muss. Malaria stellt somit nach Tuberkulose und HIV die bedrohlichste aller bekannten Infektionskrankheiten dar [1].

Obwohl Malaria in vielen Teilen der Welt bereits weitgehend ausgerottet werden konnte, nimmt die Zahl der Malariafälle wieder zu. Dabei tritt Malaria auch wieder in Gebieten auf, in denen die Krankheit bereits als ausgerottet galt. Außerdem erfolgt eine Ausbreitung auf neue Gebiete wie Zentralasien und Osteuropa. Derzeit sterben jährlich mehr Menschen an Malaria als vor 30 Jahren. Schätzungen gehen von einer 40 %igen Zunahme der Infektionsrate zwischen 1970 und 1997 im subtropischen und tropischen Afrika aus, wo ca. 80 % aller Malariafälle auftreten [1]. In Abb. 1 sind die derzeitigen Malaria-Risikogebiete dargestellt.

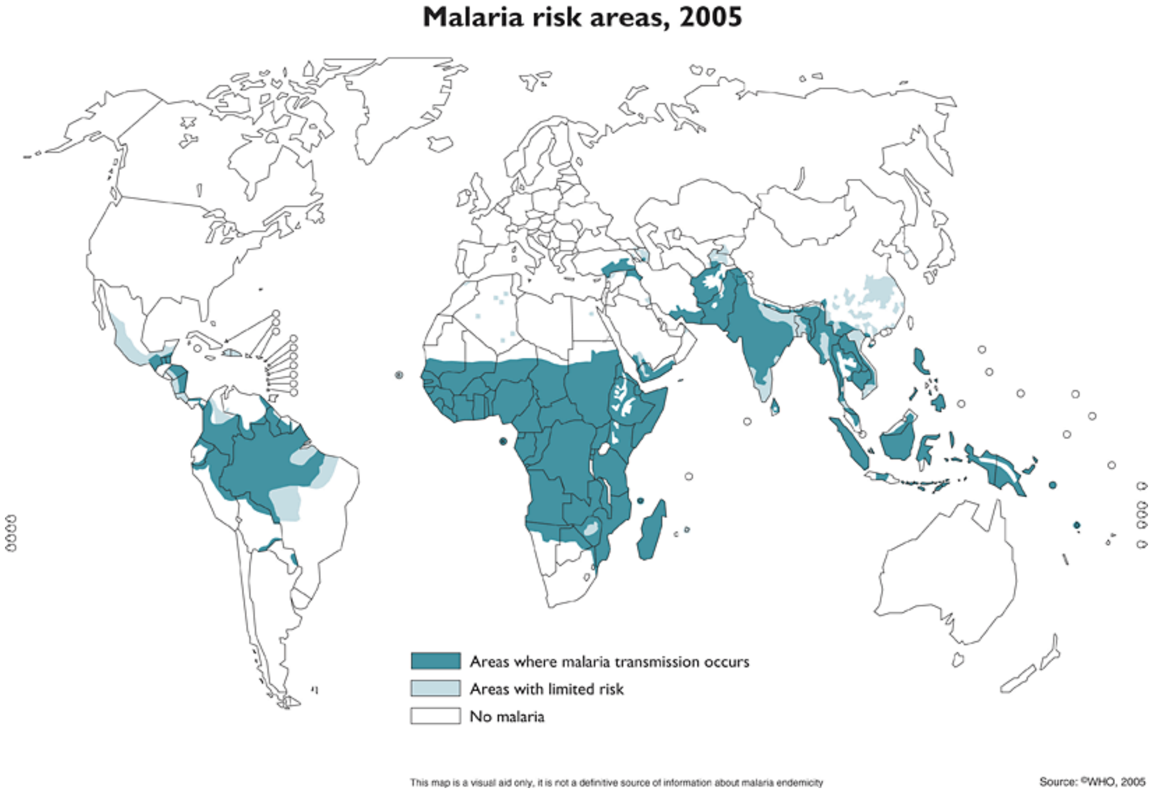

Malaria risk areas, 2005
Areas where malaria transmission occurs
Areas with limited risk
No malaria
This map is a visual aid only, it is not a definitive source of information about malaria endemicity
Source: ©WHO, 2005

Abbildung 1: Derzeitige Verbreitung von Malaria.

1.1.1 Aktuelle Malaria-Wirkstoffe, Resistenzentwicklungen und neue Forschungsansätze

Chinin, ein aus der Rinde des Chinarindenbaums isoliertes Alkaloid, war über eine Dauer von mehreren hundert Jahren das einzige wirksame Mittel gegen Malaria [2]. Die Entwicklung synthetischer Malaria-Wirkstoffe begann zu Zeiten des Ersten und Zweiten Weltkriegs, da auf Grund stark reduzierter Chininlieferungen der Bedarf nicht mehr gedeckt werden konnte. Heute gibt es eine Reihe von Medikamenten, die auf unterschiedliche Weise wirken. Die Entdeckung von Malaria-Medikamenten war in vielen Fällen zufällig. So ist der genaue Wirkmechanismus von vielen Stoffen auch bis heute nicht vollständig aufgeklärt oder komplett unbekannt [3]. Aktuell wichtige Malaria-Wirkstoffe sind in Abbildung 2 dargestellt und umfassen: Chinolin-Methanole (z.B. Chinin und Mefloquin), die vermutlich die parasitäre Entgiftung von Häm zu Hämozoin stören [3-5], 4-Aminochinoline (z.B. Chloroquin), welche ebenfalls mit der Hämozoinbildung interferieren [3-5], 8-Aminochinoline, (z.B. Primaquin), bei denen man glaubt, dass sie die Mitochondrien der Parasiten angreifen [3], Phenantrene (z.B. Halofantrin) mit unbekanntem Mechanismus [6], 2-Hydroxynaphtochinone (z.B. Atovaquon), welche den Elektronentransport der Mitochondrien hemmen sollen [3], Artemisinine (z.B. Artemether und Artesunat), die freie Radikale bilden und dadurch den Erreger schwächen [3, 7], Antifolate (z.B. Proguanil und Pyrimethamin) als Hemmstoffe der Dihydrofolat-Reduktase (DHFR) und Inhibitoren der Dihydropteroat-Synthetase (DHPS), z.B. Sulfadoxin, die ebenfalls den Folatzyklus des Parasiten blockieren und schließlich einige Antibiotika wie die Tetracycline und Doxycycline [3, 7, 8].

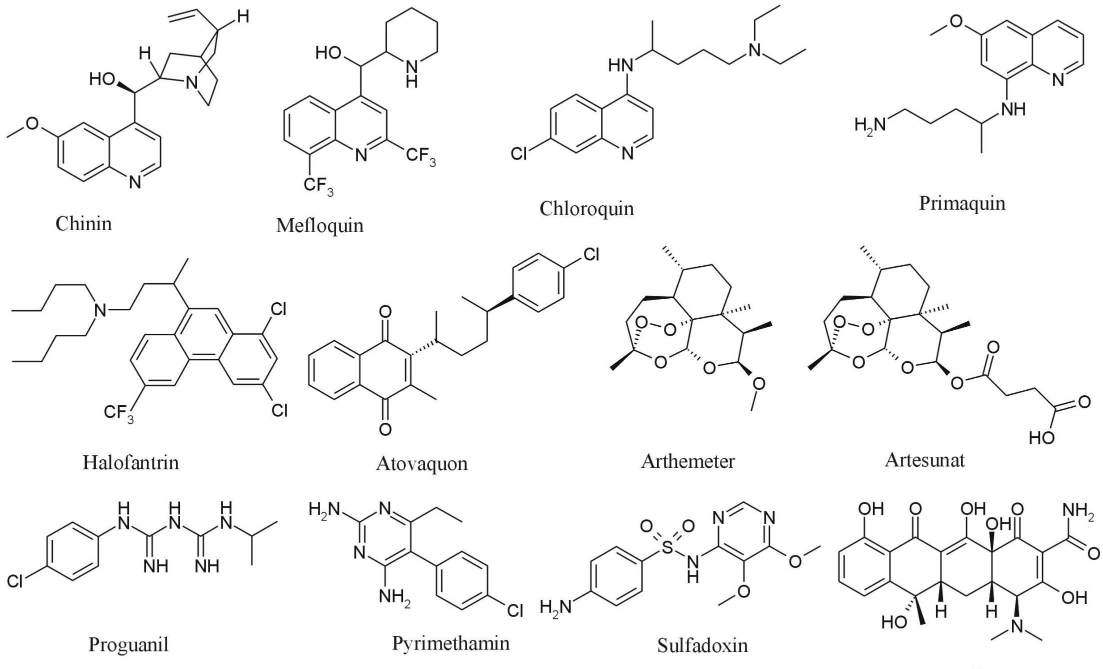

Chinin
Mefloquin
Chloroquin
Primaquin
Halofantrin
Atovaquon
Arthemeter
Artesunat
Proguanil
Pyrimethamin
Sulfadoxin
Tetracyclin

Abbildung 2: Beispiele einiger eingesetzter Malariamedikamente [9].

Es gibt Malaria-Medikamente, die ihre Wirkung zu einem ganz bestimmten Stadium im Lebenszyklus des Parasiten entfalten, wie z.B. Chinin, aber auch Medikamente, die unspezifisch sind und in mehreren Stadien wirken (z.B. Primaquin). Die Mehrzahl der Wirkstoffe interferiert jedoch mit den Blutstadien, die für nahezu alle Symptome verantwortlich sind. Mehrere Jahrzehnte lang war Chloroquin das Mittel der Wahl in Malaria betroffenen Gebieten [10]. Durch aufkommende Resistenzen in vielen Regionen wurde zu einer Kombinationstherapie übergegangen. Sulfadoxin-Pyrimethamin (SP) und auf Artemisin basierende Kombinationen (ACT) setzten sich durch [11]. Unglücklicherweise breiten sich nun Resistenzen auch gegen SP ziemlich schnell aus. Im World Malaria Report 2005 der WHO wird geraten, in Regionen mit weit verbreiteten Resistenzen, ACTs zur Behandlung einzusetzen [12].

1910 wurde in Brasilien die erste Malaria-Resistenz gegen Chinin beobachtet [3]. Seitdem stiegen die Berichte über resistente Stämme weltweit kontinuierlich an [13] und bis auf die Artemisine sind alle Klassen von Wirkstoffen betroffen [14]. Die Wahrscheinlichkeit ist hoch, dass auch gegen Artemisine Resistenzen ausgebildet werden [13].

Neue Arzneimittel, die wirksam gegen „multi-drug“ resistente Parasiten sind, werden dringend benötigt. Abgesehen von den Resistenzen sprechen Nachteile wie schlechte Wirksamkeit, Toxizität und hohe Produktionskosten zusätzlich für die Weiterentwicklung von Malariamitteln.

Ein sehr weit verbreiteter Ansatz ist die Entwicklung von chemisch verwandten Verbindungen zu den existierenden Wirkstoffen [15-17]. Ein alternativer Ansatz besteht aus der Identifizierung neuer Drug Targets und dem anschließenden Design potentieller Inhibitoren dieser Targets [17]. Durch die Sequenzierung des *P. falciparum* Genoms wurden einige potentielle Zielstrukturen entdeckt [15, 17-21], unter denen sich die Plasmepsine befinden, die Hauptgegenstand dieser Arbeit sind.

1.1.2 Protozoen der Gattung *Plasmodium*

Durch die Infektion mit Protozoen der Gattung *Plasmodium* wird Malaria verursacht. Protozoen sind einzellige, bewegliche, animalische Wesen. Charakteristisch im Lebenszyklus pathogener Protozoen ist der parasitäre Aufenthalt in verschiedenen Wirten, bei Malaria im Menschen und der weiblichen Anopheles-Mücke (Abb. 3).

Abbildung 3: weibliche Anopheles Mücke.

Es existieren vier verschiedene humanpathogene *Plasmodium*-Arten, die zu jeweils spezifischen Krankheitsbildern führen. Es handelt sich dabei um *P. falciparum* (Malaria tropica), *P. vivax* (Malaria tertiana), *P. ovale* (Malaria tertiana) und *P. malariae* (Malaria quartana).

1.1.3 Der Lebenszyklus von *Plasmodium*

Der Lebenszyklus der Malaria verursachenden Parasiten ist sehr komplex und besteht aus mehreren Phasen. Er setzt sich aus zwei Grundzyklen, dem asexuellen Zyklus im Menschen und dem sexuellen Zyklus in der weiblichen Anopheles Mücke zusammen (Abbildung 4). Der asexuelle Zyklus kann weiter unterteilt werden in eine präerythrozytäre Phase (Leberschizogenie) und eine erythrozytäre Phase (Abbildung 4) [19, 22, 23]. Beim Stich einer infizierten Anopheles-Mücke werden Sporozoiten aus dem Speichel der Mücke in den menschlichen Wirt übertragen, welche über die Blutbahn zu den Leberzellen gelangen und in diese eindringen.

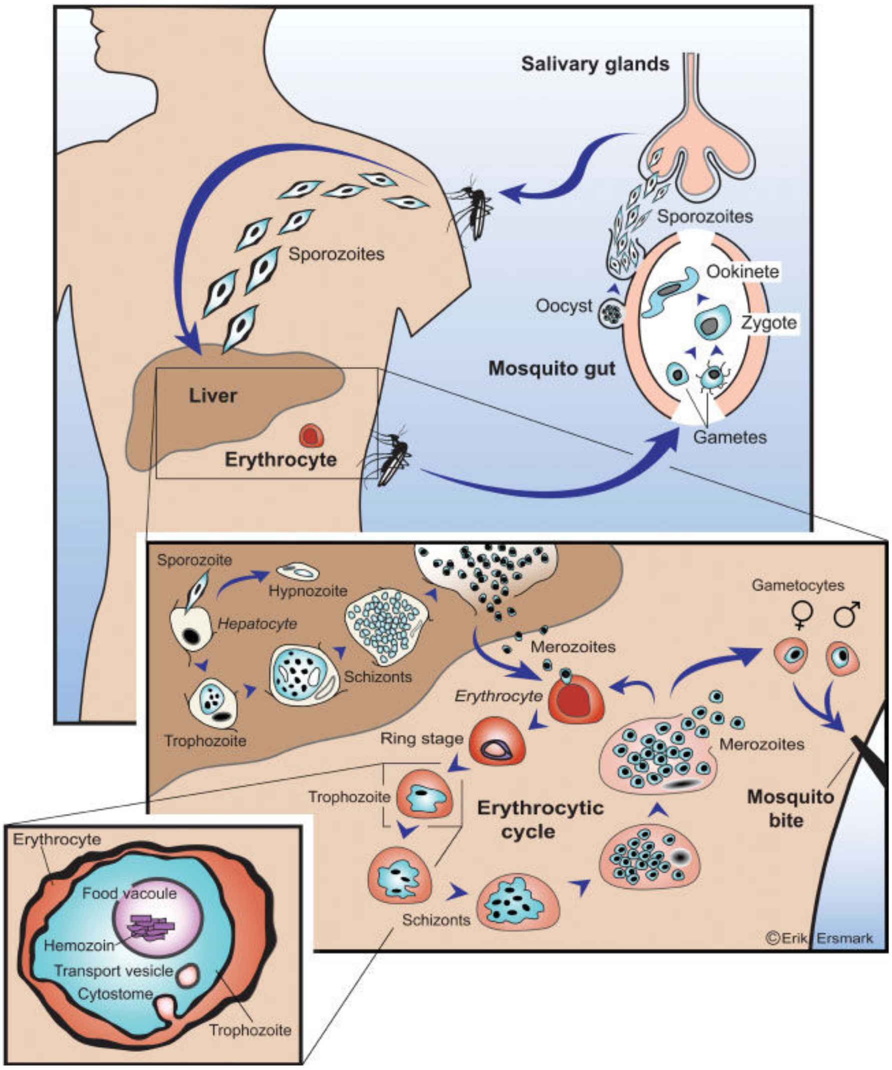

Salivary glands
Sporozoites
Sporozoites
Ookinete
Oocyst
Zygote
Mosquito gut
Gametes
Liver
Erythrocyte
Sporozoite
Hypnozoite
Hepatocyte
Schizonts
Merozoites
Gametocytes
♀ ♂
Erythrocyte
Trophozoite
Ring stage
Merozoites
Trophozoite
Erythrocytic cycle
Mosquito bite
Schizonts
©Erik Ersmark
Erythrocyte
Food vacoule
Hemozoin
Transport vesicle
Cytostome
Trophozoite

Abbildung 4: Der Lebenszyklus des Malariaerregers im Menschen und im Moskito. Das Trophozoiten-Stadium im Erythrozyten ist vergrößert dargestellt [9].

In den Leberzellen werden die Sporozoiten in Trophozoiten umgewandelt, die sich mehrfach teilen, wodurch eine Vielzahl von Schizonten entsteht. Jeder Schizont umgibt sich mit Wirtszell-Zytoplasma, wodurch Merozoiten entstehen, welche freigesetzt werden, in die Blutbahn gelangen und Erythrozyten infizieren. Die präerythrozytäre Phase, die sich aus der Entwicklung und Vermehrung des Parasiten in der Leberzelle zusammensetzt, verläuft asymptotisch und dauert abhängig von der *Plasmodien* Art, zwischen fünf und 16 Tagen.

Die darauf folgende erythrozytäre Phase beginnt mit dem Eindringen der Merozoiten in die Erythrozyten. Es erfolgt eine Differenzierung vom Ringstadium über das Trophozoitenstadium zum Schizonten. Im Segmenter-Stadium teilen sich die Schizonten in meist 16 Merozoiten, die durch Platzen der Erythrozyten freigesetzt werden und neue Erythrozyten infizieren. Die Lyse der Erythrozyten ist verantwortlich für das Fieber. Unter dem Einfluss von Stressfaktoren differenzieren sich einige der erythrozytären Parasiten zu weiblichen und männlichen Gametozyten (Makrogametozyten und Mikrogametozyten). Diese Gametozyten werden bei einem erneuten Mückenstich mit der Blutmahlzeit aufgenommen. Im Mitteldarm der Mücke entwickeln sich die weiblichen Gametozyten zum Makrogameten, und die männlichen Gametozyten bilden jeweils vier bis acht bewegliche Mikrogameten. Die Verschmelzung der Gameten ergibt eine bewegliche Zygote (Ookinet), die das Darmepithel durchdringt und an der Darmaußenwand zur Oocyste heranwächst. In der Oocyste entstehen Tausende von beweglichen Sporozoiten, die nach Platzen der Oocyste frei werden und sich in den Drüsenkanälen der Speicheldrüse sammeln.

1.1.4 Krankheitsbild (physiologische Folgen)

Nach dem Stich einer infizierten Mücke (Abb. 3) kommt es zunächst zu einer symptomfrei verlaufenden Infektion der Leber. Nach einer Inkubationszeit von durchschnittlich einer Woche werden Blutstadien gebildet, die sich in den Erythrozyten vermehren. Ein Teil der Parasiten entwickelt sich zu geschlechtlichen Stadien, sog. Gametozyten, die zur erneuten Infektion einer Mücke führen, wenn sie bei einer Blutmahlzeit aufgenommen werden. Für die klinische Manifestation der Malaria sind ausschließlich die erythrozytären Stadien verantwortlich. Mit der Freisetzung der Parasiten aus den infizierten Erythrozyten gelangen Zellbestandteile, die für die Entstehung des Fiebers verantwortlich sind, in die Blutbahn. Die meisten tödlich verlaufenden Infektionen werden durch *P. falciparum* verursacht. Durch diesen Parasiten werden spezielle Proteine in die Membran der infizierten Erythrozyten eingelagert. Dadurch heften sich die infizierten Erythrozyten an den Wänden prävenöser Kapillaren an, so dass es zur Blockade der Blutgefäße kommt. Über diesen Mechanismus kann die Obstruktion von Gehirnmikrogefäßen zur cerebralen Malaria führen, die in komatösen Zuständen endet und ohne unverzügliche Behandlung tödlich verläuft [21]. Bei den

Spezies *P. vivax* und *P. ovale* können die Leberstadien als sog. Hypnozoiten jahrelang überdauern und regelmäßige Krankheitsausbrüche verursachen.

1.1.5 Hämoglobin-Abbau

Während der erythrozytären Phase baut *P. falciparum* ca. 80 % des in der Wirtszelle vorhandenen Hämoglobins ab, vermutlich um seinen Stoffwechselbedarf an Aminosäuren zu decken [24]. Der Abbau ist während der einzelnen Phasen unterschiedlich stark ausgeprägt, wobei der größte Teil im Trophozoiten-Stadium verstoffwechselt wird [25].

Der Trophozoit nimmt das Zellplasma der Wirtszelle durch eine kreisförmige Struktur, dem Zytostom, entlang der Plasmamembran gelegen, auf (Abb. 4). Das Zellplasma der Wirtszelle gelangt in den zystostomalen Hohlraum und wird in Form einer Ausbuchtung an der zytostomalen Öffnung zusammen mit der zystostomalen Membran abgeschnitten, wodurch eine Nahrungsvakuole im Zellplasma des Parasiten entsteht. Die Verdauung des Hämoglobins erfolgt in der Nahrungsvakuole, einem Kompartiment mit einem pH-Wert von ungefähr 5, wo es schließlich abgebaut wird [24] (Abb.4).

Die Gründe für den Hämoglobinabbau sind nicht vollständig geklärt, wobei einige Hypothesen im Raum stehen: Der Parasit besitzt nur begrenzte Möglichkeiten sich selbständig Aminosäuren herzustellen. Daher könnten die aus dem Hämoglobin gewonnenen Aminosäuren, dem Erreger zur Proteinbiosynthese dienen [26, 27]. Weiterhin könnten sie aber auch für dessen Energieversorgung wichtig sein [27].

Unabhängig von den Gründen, konnten Studien mit Proteaseinhibitoren belegen, dass der Hämoglobinabbau essentiell für das Überleben des Parasiten ist [28-31].

Der Abbau in *P. falciparum* erfolgt entlang einer Enzymkaskade durch das Zusammenwirken von Aspartylproteasen (Plasmepsin (Plm) I [29, 32], II [33, 34], IV [35] und der nahe verwandten Histidin-Asparagin Protease (HAP) [35-37]) sowie Cystein Proteasen (Falcipain-1 [38, 39], -2 [40], -2' [41] und -3 [42]), einer Metalloprotease (Falcilysin) [43] und der kürzlich entdeckten Dipeptidyl Aminopeptidase 1 (DPAP1) [44]. Der exakte Ablauf ist nur schwierig zu bestimmen, vor allem die Frage, ob die initiale Spaltung des Hämoglobins durch ein Falcipain oder Plasmepsin katalysiert wird [33, 42, 45]. Das grundlegende Schema ist in Abbildung 5 zu sehen [45]. Die initiale Spaltung des Hämoglobins erfolgt in der $\alpha-$Kette des Proteins zwischen Phe33 und Leu34. Dadurch entfaltet sich das Tetramer und kann durch Falcipaine und Plasmepsine weiter in kleinere Peptide zerlegt werden [46]. Die Metalloprotease Falcilysin baut kleine Peptide von bis zu 20 Aminosäuren zu noch kleineren Peptiden ab [43]. DPAP1 spaltet Dipeptide von Oligopeptiden ab [44], bis letztendlich die Hydrolyse zu den freien Aminosäuren vermutlich durch Aminopeptidasen im Cytoplasma des Parasiten erfolgt [47].

Das bei dem Abbau entstehende, freie toxische Häm wird durch Polymerisation und Ablagerung in Form einer inerten kristallinen Substanz, dem Hämazoin (auch bekannt als Malaria-Pigment), vom Parasiten entgiftet [4, 25]. Dieser Vorgang wird durch die Chinolin-Antimalariamittel gestört, die somit die intraparasitäre Hämkonzentration erhöhen und eine Intoxikation des Parasiten durch sein eigenes Stoffwechselprodukt herbeiführen.

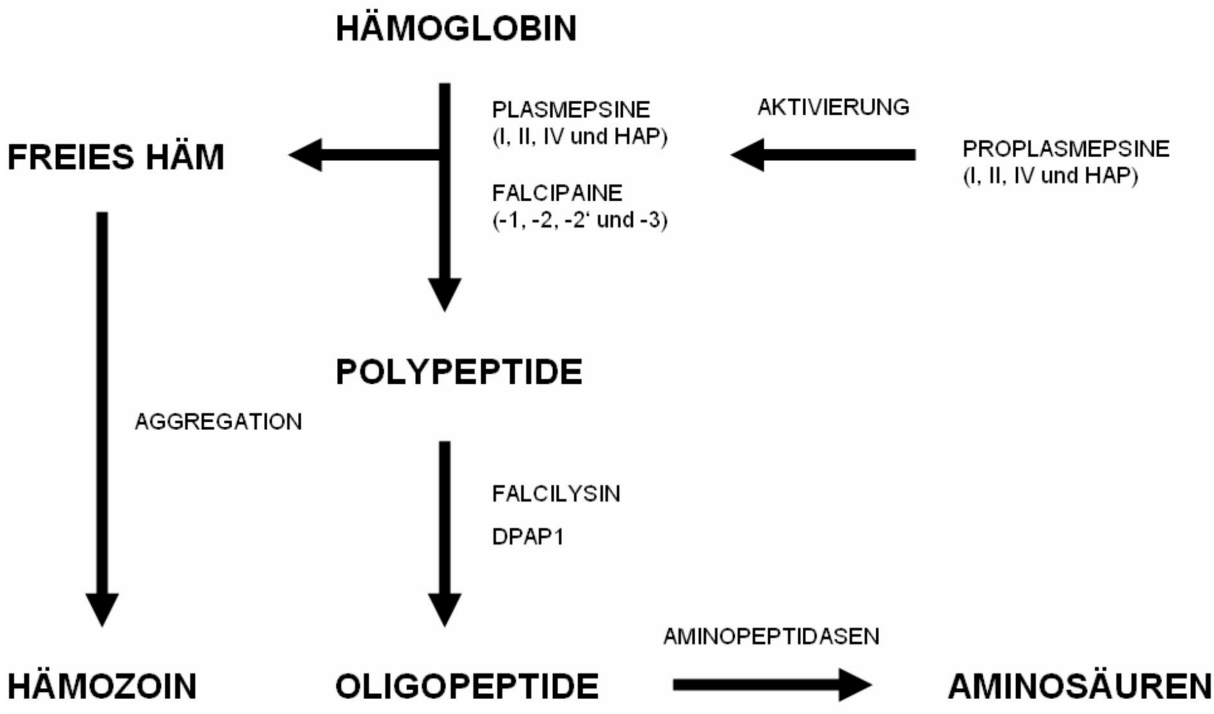

HÄMOGLOBIN
PLASMEPSINE
(I, II, IV und HAP)
FALCIPAINE
(-1, -2, -2' und -3)
AKTIVIERUNG
PROPLASMEPSINE
(I, II, IV und HAP)
FREIES HÄM
POLYPEPTIDE
AGGREGATION
FALCILYSIN
DPAP1
AMINOPEPTIDASEN
HÄMOZOIN
OLIGOPEPTIDE
AMINOSÄUREN

Abbildung 5: Hämoglobinabbau in *P. falciparum* [9].

1.2 Plasmepsine

Die Aspartylproteasen der *Plasmodien* werden Plasmepsine genannt. In *Plasmodium falciparum* gibt es mindestens zehn Gene, die Aspartylproteasen kodieren (Plm I, II, IV, V, VI, VII, VIII, IX, X und die nahe verwandte Histidin-Asparagin Protease) [45]. Die exakte Rolle jedes einzelnen Enzyms im Metabolismus, ist bisher nicht bekannt. Die Expression von Plm I, II, IV, V, IX, X und HAP findet während der erythrozytären Phase statt. Plm VI, VII und VIII werden dagegen während der exo-erythrozytären Phase exprimiert [37]. Plm I, II, IV und HAP sind die bisher am besten untersuchten Enzyme: Sie wurden im Gegensatz zu den anderen Plasmepsinen in der Nahrungsvakuole nachgewiesen und sind direkt am Hämoglobinabbau beteiligt [37].

Die Plasmepsine liegen zunächst als Zymogene vor, werden dann aber durch Abspaltung der Prosequenz von einer Proplasmepsin-Konvertase in ihre aktive Form überführt (Abbildung 5) [25, 48]. Da die Proplasmepsin-Konvertase die Aktivierung der Plasmepsine steuert, könnte eine Inhibierung dieses Proteins ein weiterer guter Ansatzpunkt zu einem potentiellen Wirkstoff sein [48]. Die aktivierten Plasmepsine unterscheiden sich in ihrer Spezifität bezüglich der

Spaltstelle und in ihrem pH-Optimum [37]. Zusätzlich zum Hämoglobinabbau wird vermutet, dass Plm II und IV an der Spaltung der Wirts-Erythrozytenmembran beteiligt sind [37, 49, 50].
HAP hat ein Histidin an der Position eines der katalytischen Aspartate. Der genaue Reaktionsmechanismus dieses Enzyms ist nicht geklärt. Die Ähnlichkeit zu den Plasmepsinen wird aber durch die hohe Sequenzhomologie offensichtlich (60% - 70%). Im Vergleich zu Plm II sind die Aminosäuren in den Bindetaschen von Plm I, IV und HAP zu 84%, 68% und 39% identisch. Dagegen zeigen die Plms V-X eine sehr viel geringere Sequenzidentität.

1.2.1 Selektivität

Bei der Entwicklung von Plasmepsin-Inhibitoren spielt die Selektivität gegenüber humanen Aspartylproteasen eine wichtige Rolle. Bei nicht ausreichendem Selektivitätsprofil können einerseits toxische Effekte in Erscheinung treten, weiterhin wird beim Binden an menschliche Proteine aber auch die Konzentration des Wirkstoffes herabgesetzt [45]. Die Plasmepsine zeigen unterschiedliche Sequenzhomologien zu menschlichen Aspartylproteasen. Das ähnlichste Enzym ist das lysosomale Cathepsin D (Cat D), mit einer Sequenzidentität von 35 % zu Plm II und einer sogar noch höheren Sequenzidentität im Bereich der Bindetaschen [51]. Andere wichtige menschliche Aspartylproteasen wie Cathepsin E, Renin und Pepsin A besitzen eine niedrigere Sequenzähnlichkeit [52, 53]. Um mögliche Kreuzinhibierungen auszuschließen, werden Plasmepsin-Inhibitoren daher häufig auf Cathepsin D-Aktivität getestet.
Das Design von affinen Plasmepsin-Inhibitoren, die nicht an Cathepsin D binden stellt eine große Herausforderung dar.

1.2.2 Aspartylproteasen

Aspartylproteasen sind eine Unterklasse der Proteasen. Auf Grund von strukturellen Unterschieden werden die Unterklassen noch weiter in Familien unterteilt, wobei ähnliche Familien zu Clans zusammengefasst werden [54]. Die meisten Aspartylproteasen, die Plasmepsine und Cathepsin D eingeschlossen, gehören zu der Pepsin-Familie, die nur in Eukaryonten vorkommen. Zusammen mit den viralen Retropepsinen, zu denen die HIV-1 Protease gehört, bilden sie einen Clan [55]. Aspartylproteasen bestehen aus zwei Domänen. An jeder Domäne befindet sich jeweils ein Aspartat der katalytischen Diade [56]. Eine weitere Wechselwirkung zu Substrat oder zu einem gebundenen Inhibitor wird über eine β-Haarnadel Schleife, die so genannte Flap-Region, gebildet, die die Bindetasche bedeckt [56]. Das Substrat bindet an das Enzym, indem die Aminosäureseitenketten des Substrats (P1-Pn / P1’-Pn’) Wechselwirkungen zu den Aminosäuren der entsprechenden Subtaschen (S1-Sn / S1’-Sn’) ausbilden [57]. Die Nomenklatur folgt dabei dem Vorschlag von Schechter und Berger: Von der Spaltstelle in Richtung des N-Terminus werden die Aminosäurereste des Substrats mit P1-Pn, die korrespondierenden Bindetaschen mit S1-Sn benannt, während in Richtung des C-Terminus die Reste mit P1'-Pn' und die Taschen mit S1'-Sn' bezeichnet werden [57].

Der katalytische Mechanismus der Aspartylproteasen wurde intensiv durch kinetische Messungen, Isotopenmarkierung, theoretische Rechnungen und Röntgenkristallographie untersucht [56, 58-60]. In Abbildung 6 ist der Mechanismus schematisch dargestellt: Die Spaltung der Amidbindung des Peptidsubstrats wird durch zwei Aspartylreste katalysiert. Die Peptidspaltung verläuft über den nukleophilen Angriff eines katalytischen Wassermoleküls auf den Carbonylkohlenstoff der zu spaltenden Amidbindung. Die Elektrophilie dieses Kohlenstoffs wird durch eine Wasserstoffbrücke des gebundenen Sauerstoffs zu dem räumlich benachbarten Aspartat erhöht. Das zweite Aspartat polarisiert und fixiert das angreifende Wassermolekül im Raum. Intermediär tritt ein tetraedrischer Übergangszustand mit zwei Hydroxylgruppen am zentralen Kohlenstoff auf. Unter Ausbildung einer freien Säure- und Aminfunktion zerfällt das Peptidsubstrat.

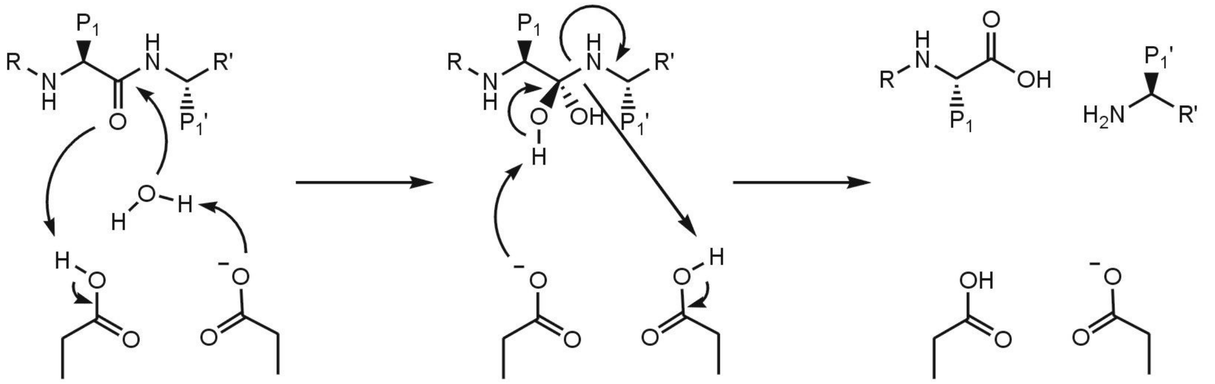


Abbildung 6: Schematischer Mechanismus von Aspartylproteasen.

Eine allgemeine Strategie zum Design von Inhibitoren verfolgt den Entwurf von Verbindungen zur Stabilisierung des intermediären tetraedrischen Übergangszustands. Diese Verbindungen, die Übergangszustand-Analoga genannt werden, binden stärker an das Protein als das Substrat [61, 62]. Ein Übergangszustand-Isoster ist eine funktionelle Gruppe, die der tetraedrischen Zwischenstufe strukturell gleicht, jedoch durch das Enzym nicht gespalten werden kann [63]. Einige funktionelle Gruppen, die diese Kriterien erfüllen, wurden identifiziert, so z.B. reduzierte Amide, Statine, Hydroxyethylamine, Hydroxypropylamine, Norstatine, Dihydroxyethylene, Phosphinate und Difluorketone [60, 64, 65] (Abbildung 7).

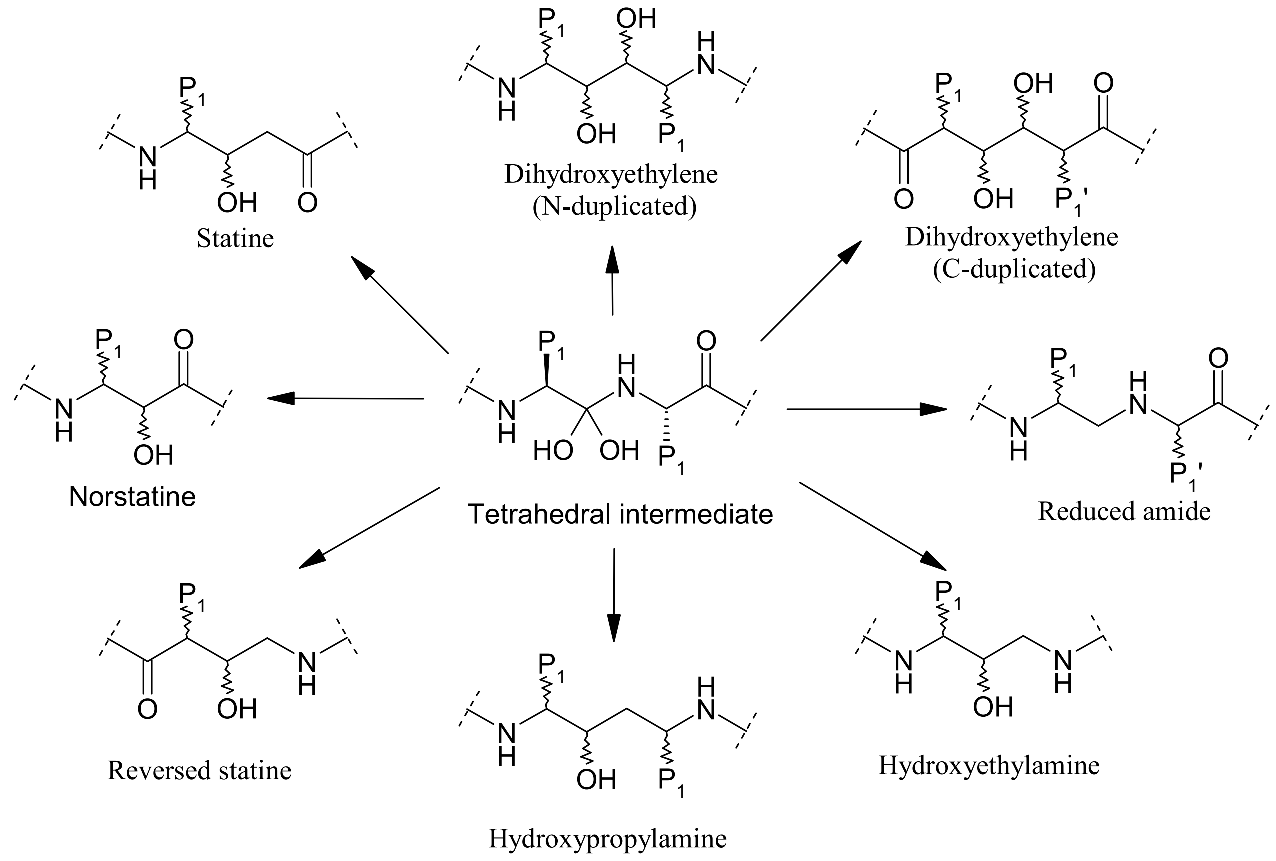

Statine
Dihydroxyethylene
(N-duplicated)
Dihydroxyethylene
(C-duplicated)
Norstatine
Tetrahedral intermediate
Reduced amide
Reversed statine
Hydroxypropylamine
Hydroxyethylamine

Abbildung 7: Übergangszustands Mimetika peptidischer Plasmepsin Inhibitoren [9].

1.2.3 Plasmepsin Inhibitoren

Die Plasmepsine wurden als Drug Targets validiert, indem gezeigt wurde, dass Malaria-Parasiten, sowohl in Kultur als auch in Tiermodellen durch Plasmepsin-Inhibitoren absterben [28, 29, 66, 67]. Die Analyse von verschiedenen Substraten und Mutationen im Bereich der Bindetasche, gaben Aufschlüsse über Spezifitäten der verschiedenen Plasmepsine [68-75]. Weiterhin wurden Substanzbibliotheken gescreent und Ähnlichkeiten zu anderen Aspartylproteasen berücksichtigt, um neue affine Plasmepsin-Inhibitoren zu finden. Sowohl substratähnliche, peptidomimetische, als auch nichtpeptidische Inhibitoren wurden gefunden [67, 76-80].

Zu Beginn der neunziger Jahre waren Plm I und II die einzigen charakterisierten Plasmepsine. Dadurch, dass es Schwierigkeiten bei der rekombinanten Produktion von Plm I gab, von Plm II allerdings recht schnell eine Kristallstruktur im Komplex mit Pepstatin A erhalten werden konnte, wurde der Schwerpunkt beim Inhibitoren-Design überwiegend auf Plm II gelegt [51, 66, 81]. Im Laufe der Jahre fand die Charakterisierung der weiteren Plasmepsinen statt, was zu einer Erweiterung der Palette an denkbaren Drug Targets führte [45]. Heute spielen neben Plm I und Plm II hauptsächlich Plm IV und HAP eine bedeutende Rolle [82]. Trotz der Tatsache, dass Aspartylproteasen *P. falciparum* in Kultur zum Absterben bringen, ist nicht geklärt, welches bzw. welche Plasmepsine essentiell für das Überleben des Parasiten sind [45]. Kürzlich durchgeführte Versuche an *P. falciparum* Plm Knockout-Klonen zeigten die Notwendigkeit des gleichzeitigen Hemmens mehrerer Plasmepsine, um den Malariaerreger zu bezwingen [83, 84]. Erfolgversprechend scheint daher insbesonders das Design von Inhibitoren zu sein, die mehrere Plms hemmen können.

In der letzten Zeit wurden auch bekannte Cat D-, Renin- und HIV-1-Protease-Inhibitoren auf ihre Fähigkeit Plasmepsine zu hemmen, getestet und dabei eine Vielzahl interessanter Verbindungen gefunden [51, 66, 85-89]. Sogar einige auf dem Markt zugelassene HIV-1-Protease-Inhibitoren zeigten hohe Affinitäten zu Plm II und Aktivität in *P.falciparum* infizierten Erythrozyten [90].

Die größte Anzahl an Plasmepsin-Inhibitoren gehört zu den Peptidomimetika. Die peptidomimetischen Verbindungen werden entsprechend ihrem Übergangszustand-Isosters unterteilt [9] (Abbildung 7).

1.3 References

1. Report on Infectious Diseases. Geneva, Switzerland: World Health Organisation, Office of Information, 1999.
2. Greenwood D. The quinine connection. *J Antimicrob Chemother* 1992;**30**(4):417-27.
3. Rosenthal P, ed. *Antimalarial chemotherapy: Mechanisms of action, resistance, and new directions in drug discovery*. Totowa, New Jersey: Human Press, 2001.
4. Egan T. Haemozoin (malaria pigment): A unique crystalline drug target. *Targets* 2003;**2**:115-124.
5. Fitch CD. Ferriprotoporphyrin IX, phospholipids, and the antimalarial actions of quinoline drugs. *Life Sci* 2004;**74**(16):1957-72.
6. Casteel D. *Antimalaria Agents*. Vol 5. New York: John Wiley and Sons, 1997.
7. Tang Y, Dong Y, Wang X*, et al.* Dispiro-1,2,4-trioxane analogues of a prototype dispiro-1,2,4-trioxolane: mechanistic comparators for artemisinin in the context of reaction pathways with iron(II). *J Org Chem* 2005;**70**(13):5103-10.
8. Gregson A, Plowe CV. Mechanisms of resistance of malaria parasites to antifolates. *Pharmacol Rev* 2005;**57**(1):117-45.
9. Ersmark K, Samuelsson B, Hallberg A. Plasmepsins as potential targets for new antimalarial therapy. *Medicinal Research Reviews* 2006;**26**(5):626-666.
10. Winstanley P, Ward S, Snow R*, et al.* Therapy of falciparum malaria in sub-saharan Africa: from molecule to policy. *Clin Microbiol Rev* 2004;**17**(3):612-37, table of contents.
11. Kremsner PG, Krishna S. Antimalarial combinations. *Lancet* 2004;**364**(9430):285-94.
12. Korenromp E, Miller J, Nahlen B*, et al.* World Malaria Report. World Health Organization (WHO), Roll Back Malaria (RBM) Department, and the United Nations Children's Fund (UNICEF) 2005.
13. Wongsrichanalai C, Pickard AL, Wernsdorfer WH*, et al.* Epidemiology of drug-resistant malaria. *Lancet Infect Dis* 2002;**2**(4):209-18.
14. White NJ. Antimalarial drug resistance. *J Clin Invest* 2004;**113**:1084-1092.
15. Dominguez JN. Chemotherapeutic agents against malaria: what next after chloroquine? *Curr Top Med Chem* 2002;**2**(11):1173-85.
16. Waters N, Dow G, Kozar M. Antimalarial compound identification and design: Advances in the patent literature, 2000-2003. *Expert Opin Ther Patents* 2004;**14**:1125-1138.
17. Woster P. New therapies for malaria. *Ann Rep Med Chem* 2003;**38**:203-211.
18. Mehlin C. Structure-based drug discovery for Plasmodium falciparum. *Comb Chem High Throughput Screen* 2005;**8**(1):5-14.
19. Young JA, Winzeler EA. Using expression information to discover new drug and vaccine targets in the malaria parasite Plasmodium falciparum. *Pharmacogenomics* 2005;**6**(1):17-26.
20. Olliaro PL, Yuthavong Y. An overview of chemotherapeutic targets for antimalarial drug discovery. *Pharmacol Ther* 1999;**81**(2):91-110.
21. Wiesner J, Ortmann R, Jomaa H*, et al.* New antimalarial drugs. *Angew Chem Int Ed Engl* 2003;**42**(43):5274-93.

22. Volk W, Gebhardt B, M-L H*, et al.* Essentials of medicinal microbiology. Philadelphia: Lippincott-Raven Publishers, 1996.

23. Khan SM, Waters AP. Malaria parasite transmission stages: an update. *Trends Parasitol* 2004;**20**(12):575-80.

24. Goldberg DE. Hemoglobin degradation in Plasmodium-infected red blood cells. *Seminars in Cell Biology* 1993;**4**(5):355-61.

25. Francis SE, Sullivan DJ, Jr., Goldberg DE. Hemoglobin metabolism in the malaria parasite Plasmodium falciparum. *Annual Review of Microbiology* 1997;**51**:97-123.

26. Sherman IW, Tanigoshi L. Incorporation of ^{14}C-labeled amino acids by the malaria[parasite] *Plasmodium lophurea*. IV. In vivo utilization of host-cell hemoglobin. *Int J Biochem* 1970;**1**:635-637.

27. Sherman IW. Amino acid metabolism and protein synthesis in malarial parasites. *Bull WHO* 1977;**55**:265-276.

28. Bailly E, Jambou R, Savel J*, et al.* Plasmodium falciparum: differential sensitivity in vitro to E-64 (cysteine protease inhibitor) and Pepstatin A (aspartyl protease inhibitor). *J Protozool* 1992;**39**(5):593-9.

29. Francis SE, Gluzman IY, Oksman A*, et al.* Molecular characterization and inhibition of a Plasmodium falciparum aspartic hemoglobinase. *EMBO Journal* 1994;**13**(2):306-17.

30. Rosenthal PJ. Plasmodium falciparum: effects of proteinase inhibitors on globin hydrolysis by cultured malaria parasites. *Exp Parasitol* 1995;**80**(2):272-81.

31. Olson JE, Lee GK, Semenov A*, et al.* Antimalarial effects in mice of orally administered peptidyl cysteine protease inhibitors. *Bioorganic & Medicinal Chemistry* 1999;**7**(4):633-8.

32. Goldberg DE, Slater AF, Beavis R*, et al.* Hemoglobin degradation in the human malaria pathogen Plasmodium falciparum: a catabolic pathway initiated by a specific aspartic protease. *Journal of Experimental Medicine* 1991;**173**(4):961-9.

33. Gluzman IY, Francis SE, Oksman A*, et al.* Order and specificity of the Plasmodium falciparum hemoglobin degradation pathway. *Journal of Clinical Investigation* 1994;**93**(4):1602-8.

34. Dame JB, Reddy GR, Yowell CA*, et al.* Sequence, expression and modeled structure of an aspartic proteinase from the human malaria parasite Plasmodium falciparum. *Molecular and Biochemical Parasitology* 1994;**64**(2):177-90.

35. Humphreys MJ, Moon RP, Klinder A*, et al.* The aspartic proteinase from the rodent parasite Plasmodium berghei as a potential model for plasmepsins from the human malaria parasite, Plasmodium falciparum. *Febs Letters* 1999;**463**(1-2):43-48.

36. Berry C, Humphreys MJ, Matharu P*, et al.* A distinct member of the aspartic proteinase gene family from the human malaria parasite Plasmodium falciparum. *FEBS Lett* 1999;**447**(2-3):149-54.

37. Banerjee R, Liu J, Beatty W*, et al.* Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. *Proceedings of the National Academy of Sciences of the United States of America* 2002;**99**(2):990-995.

38. Salas F, Fichmann J, Lee GK*, et al.* Functional Expression of Falcipain, a Plasmodium-Falciparum Cysteine Proteinase, Supports Its Role as a Malarial Hemoglobinase. *Infection and Immunity* 1995;**63**(6):2120-2125.

39. Sijwali PS, Kato K, Seydel KB*, et al.* Plasmodium falciparum cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites. *Proceedings of the National Academy of Sciences of the United States of America* 2004;**101**(23):8721-8726.

40. Shenai BR, Sijwali PS, Singh A*, et al.* Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of Plasmodium falciparum. *Journal of Biological Chemistry* 2000;**275**(37):29000-29010.
41. Sijwali PS, Rosenthal PJ. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by Plasmodium falciparum. *Proceedings of the National Academy of Sciences of the United States of America* 2004;**101**(13):4384-4389.
42. Sijwali PS, Shenai BR, Gut J*, et al.* Expression and characterization of the Plasmodium falciparum haemoglobinase falcipain-3. *Biochemical Journal* 2001;**360**(Pt 2):481-9.
43. Eggleson K, K. , Duffin KL, Goldberg DE. Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite Plasmodium falciparum. *Journal of Biological Chemistry* 1999;**274**(45):32411-32417.
44. Klemba M, Gluzman I, Goldberg DE. A Plasmodium falciparum dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation. *Journal of Biological Chemistry* 2004;**279**(41):43000-7.
45. Coombs GH, Goldberg DE, Klemba M*, et al.* Aspartic proteases of Plasmodium falciparum and other parasitic protozoa as drug targets. *Trends in Parasitology* 2001;**17**(11):532-537.
46. Francis SE, Banerjee R, Goldberg DE. Biosynthesis and maturation of the malaria aspartic hemoglobinases plasmepsins I and II. *Journal of Biological Chemistry* 1997;**272**(23):14961-8.
47. Gavigan CS, Dalton JP, Bell A. The role of aminopeptidases in haemoglobin degradation in Plasmodium falciparum-infected erythrocytes. *Mol Biochem Parasitol* 2001;**117**(1):37-48.
48. Banerjee R, Francis SE, Goldberg DE. Food vacuole plasmepsins are processed at a conserved site by an acidic convertase activity in Plasmodium falciparum. *Molecular and Biochemical Parasitology* 2003;**129**(2):157-165.
49. Bonniec Le S, Deregnaucourt C, Redeker V*, et al.* Plasmepsin II, an Acidic Hemoglobinase from Plasmodium falciparum Food Vacuole, is active at Neutral pH on the Host Erythrozyte Membrane Skeleton. *Journal of Biological Chemistry* 1999;**274**(20):14218-14223.
50. Wyatt DM, Berry C. Activity and inhibition of plasmepsin IV, a new aspartic proteinase from the malaria parasite, Plasmodium falciparum. *FEBS Lett* 2002;**513**(2-3):159-62.
51. Silva AM, Lee AY, Gulnik SV*, et al.* Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from Plasmodium falciparum. *Proc Natl Acad Sci U S A* 1996;**93**(19):10034-9.
52. Silva AM, Lee AY, Erickson JW*, et al.* Structural analysis of plasmepsin II. A comparison with human aspartic proteases. *Advances in Experimental Medicine and Biology* 1998;**436**:363-73.
53. Bernstein NK, Cherney MM, Yowell CA*, et al.* Structural insights into the activation of P. vivax plasmepsin. *J Mol Biol* 2003;**329**(3):505-24.
54. Rawlings ND, Tolle DP, Barrett AJ. MEROPS: the peptidase database. *Nucleic Acids Res* 2004;**32**(Database issue):D160-4.
55. Rawlings ND, Tolle DP, Barrett AJ. Evolutionary families of peptidase inhibitors. *Biochem J* 2004;**378**(Pt 3):705-16.
56. Davies DR. The structure and function of the aspartic proteinases. *Annu Rev Biophys Biophys Chem* 1990;**19**:189-215.
57. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* 1967;**27**(2):157-62.

58. Northrop DB. Follow the protons: a low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Acc Chem Res* 2001;**34**(10):790-7.
59. Dunn BM. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem Rev* 2002;**102**(12):4431-58.
60. Brik A, Wong CH. HIV-1 protease: mechanism and drug discovery. *Org Biomol Chem* 2003;**1**(1):5-14.
61. Wolfenden R. Analog approaches to the structure of the transition state in enzyme reactions. *Acc Chem Res* 1972;**5**:10-18.
62. Lienhard GE. Enzymatic catalysis and transition-state theory. *Science* 1973;**180**(82):149-54.
63. Dash C, Kulkarni A, Dunn B*, et al.* Aspartic peptidase inhibitors: implications in drug development. *Crit Rev Biochem Mol Biol* 2003;**38**(2):89-119.
64. Babine RE, Bender SL. Molecular Recognition of Proteinminus signLigand Complexes: Applications to Drug Design. *Chem Rev* 1997;**97**(5):1359-1472.
65. Cooper JB. Aspartic proteinases in disease: a structural perspective. *Curr Drug Targets* 2002;**3**(2):155-73.
66. Moon RP, Tyas L, Certa U*, et al.* Expression and characterization of plasmepsin I from Plasmodium falciparum. *European Journal of Biochemistry* 1997;**244**(2):552-560.
67. Boss C, Richard-Bildstein S, Weller T*, et al.* Inhibitors of the Plasmodium falciparum parasite aspartic protease plasmepsin II as potential antimalarial agents. *Current Medicinal Chemistry* 2003;**10**(11):883-907.
68. Westling J, Yowell CA, Majer P*, et al.* Plasmodium falciparum, P. vivax, and P. malariae: a comparison of the active site properties of plasmepsins cloned and expressed from three different species of the malaria parasite. *Experimental Parasitology* 1997;**87**(3):185-93.
69. Berry C. Plasmepsins as antimalarial targets. *Current Opinion in Drug Discovery & Development* 2000;**3**(5):624-629.
70. Gulnik SV, Suvorov LI, Majer P*, et al.* Sensitive fluorogenic substrates for plasmepsin 2. *Protein and Peptide Letters* 2000;**7**(4):219-223.
71. Dunn BM, Hung S. The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases. *Biochemical and Biophysical Acta* 2000;**1477**(1-2):231-40.
72. Li H, Bleriot Y, Chantereau C*, et al.* The first synthesis of substituted azepanes mimicking monosaccharides: a new class of potent glycosidase inhibitors. *Org Biomol Chem* 2004;**2**(10):1492-9.
73. Siripurkpong P, Yuvaniyama J, Wilairat P*, et al.* Active Site Contribution to Specificity of the Aspartic Proteases Plasmepsins I and II. *J. Biol. Chem.* 2002;**277**(43):41009-41013.
74. Tyas L, Gluzman I, Moon RP*, et al.* Naturally-occurring and recombinant forms of the aspartic proteinases plasmepsins I and II from the human malaria parasite Plasmodium falciparum. *FEBS Lett* 1999;**454**(3):210-4.
75. Istvan ES, Goldberg DE. Distal Substrate Interactions Enhance Plasmepsin Activity. *J. Biol. Chem.* 2005;**280**(8):6890-6896.
76. Wiesner J, Kettler K, Sakowski J*, et al.* Farnesyltransferase inhibitors inhibit the growth of malaria parasites in vitro and in vivo. *Angewandte Chemie, International Edition* 2003;**43**(2):251-254.
77. Werbovetz KA. Target-based drug discovery for malaria, leishmaniasis, and trypanosomiasis. *Current Medicinal Chemistry* 2000;**7**(8):835-860.
78. Moon RP, Bur D, Loetscher H*, et al.* Studies on plasmepsins I and II from the malarial parasite Plasmodium falciparum and their exploitation as drug targets. *Adv Exp Med Biol* 1998;**436**:397-406.

79. Berry C. Plasmepsins as antimalarial targets. *Current Opinion in Drug Discovery & Development* 2000;**3**(5):624-629.

80. Abdel-Rahman HM, Kimura T, Hidaka K*, et al.* Design of inhibitors against HIV, HTLV-I, and Plasmodium falciparum aspartic proteases. *Biol Chem* 2004;**385**(11):1035-9.

81. Luker KE, Francis SE, Gluzman IY*, et al.* Kinetic analysis of plasmepsins I and II, aspartic proteases of the Plasmodium falciparum digestive vacuole. *Molecular and Biochemical Parasitology* 1996;**79**(1):71-78.

82. Nezami A, Kimura T, Hidaka K*, et al.* High-affinity inhibition of a family of Plasmodium falciparum proteases by a designed adaptive inhibitor. *Biochemistry* 2003;**42**(28):8459-64.

83. Omara-Opyene AL, Moura PA, Sulsona CR*, et al.* Genetic Disruption of the Plasmodium falciparum Digestive Vacuole Plasmepsins Demonstrates Their Functional Redundancy. *J. Biol. Chem.* 2004;**279**(52):54088-54096.

84. Liu J, Gluzman IY, Drew ME*, et al.* The Role of Plasmodium falciparum Food Vacuole Plasmepsins. *J. Biol. Chem.* 2005;**280**(2):1432-1437.

85. Haque TS, Skillman AG, Lee CE*, et al.* Potent, low-molecular-weight non-peptide inhibitors of malarial aspartyl protease plasmepsin II. *Journal of Medicinal Chemistry* 1999;**42**(8):1428-40.

86. Nezami A, Luque I, Kimura T*, et al.* Identification and characterization of allophenylnorstatine-based inhibitors of plasmepsin II, an antimalarial target. *Biochemistry* 2002;**41**(7):2273-80.

87. Asojo OA, Afonina E, Gulnik SV*, et al.* Structures of Ser205 mutant plasmepsin II from Plasmodium falciparum at 1.8 A in complex with the inhibitors rs367 and rs370. *Acta Crystallographica. Section D, Biological Crystallography* 2002;**58**(Pt 12):2001-8.

88. Oscarsson K, Oscarson S, Vrang L*, et al.* New potent C2-Symmetric malaria plasmepsin I and II inhibitors. *Bioorganic & Medicinal Chemistry* 2003;**11**(7):1235-1246.

89. Ersmark K, Feierberg I, Bjelic S*, et al.* C2-symmetric inhibitors of Plasmodium falciparum plasmepsin II: synthesis and theoretical predictions. *Bioorganic and Medicinal Chemistry* 2003;**11**(17):3723-33.

90. Parikh S, Gut J, Istvan E*, et al.* Antimalarial activity of human immunodeficiency virus type 1 protease inhibitors. *Antimicrob Agents Chemother* 2005;**49**(7):2983-5.

2 Aufgabenstellung

Die Hemmung der Plasmepsine stellt einen Ansatzpunkt für die Entwicklung neuer Malaria-Medikamente dar. Zu Beginn der Arbeit waren bereits eine Reihe potenter, peptidomimetischer Inhibitoren und eine Kokristallstruktur von Plasmepsin II in Komplex mit Pepstatin A in der Literatur beschrieben. Um einen Beitrag im Bereich des strukturbasierten Designs von Plasmepsin Inhibitoren leisten zu können, sollten zunächst die erforderlichen Rahmenbedingungen geschaffen werden, wozu die Produktion von Plasmepsin II und IV, die Etablierung eines Assays und das Züchten zur Strukturbestimmung geeigneter Kristalle gehörten.

Im Rahmen der Diplomarbeit wurden bereits ein Expressionssystem und ein Assay zur Messsung von Affinitätsdaten etabliert. Dieser Assay spiegelt die Situation im Parasiten sehr gut wider, da Hämoglobin das natürliche Substrat des Parasiten ist. Nachteilig erwies sich allerdings die hohe Komplexität des Systems, da das Substrat mehrere Spaltstellen enthält.

Aufbauend auf den Ergebnissen der Diplomarbeit sollte ein funktionsfähiges Kristallisationssystem für Plasmepsin II entwickelt werden und ein zweiter Assay, der die Gesetze der Michaelis-Menten Kinetik befolgt und zu Kontrollmessungen eingesetzt werden kann.

Nach Etablierung der für das strukturbasierte Liganden-Design erforderlichen Methoden galt es in Zusammenarbeit mit synthetisch arbeitenden Gruppen Inhibitoren mit neuartigen, nichtpeptidomimetischen Eigenschaften zu entwerfen und zu optimieren. Da es in der Literatur bereits Hinweise darauf gab, dass es sich bei den Plasmepsinen um sehr flexible Proteine handelt, sollten Flexibilitäten beim computergestützten Design berücksichtigt werden und Affinitätsstudien für einzelne Subtaschen angefertigt werden.

Durch das Zusammenspiel der einzelnen Teilaspekte wurde erwartet einiges über die Plasmepsine zu lernen, was einen Beitrag zur Entwicklung eines neuen Malaria-Medikamentes leisten könnte.

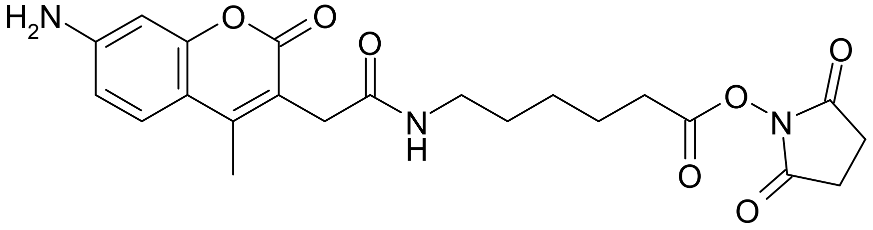
In einem Nebenprojekt sollte die Kristallstruktur für die Prephenatdehydratase aus *Methanocaldococcus jannaschii* gelöst werden, die ein vielversprechendes Target für die Entwicklung von Antibiotika darstellt.

3 Assay, Kristallisation und Strukturbestimmung

3.1 Assay

Die rekombinante Produktion der Aspartylproteasen Plm II und Plm IV konnte in einer vorausgehenden Arbeit in der AG Klebe etabliert werden (Diplomarbeit Torsten Luksch). Durchschnittliche Proteinausbeuten lagen bei ca. 3 mg Protein aus 2 l Bakterienkultur.

Die Aspartylproteasen Plasmepsin II bzw. IV katalysieren den Abbau von Hämoglobin an spezifischen Spaltstellen. Bereits in der Diplomarbeit wurde ein Verfahren zur Aktivitätsbestimmung der Plasmepsine II und IV mittels Fluoreszenzspektroskopie beschrieben: Als Fluorophor wurde AMCA (Abb. 1) verwendet, das voraussichtlich an die α-Aminogruppen der N-terminalen Reste der Hämoglobinketten gebunden wird. Dies geschieht, indem die freien α-Aminogruppen der N-terminalen Reste die aktivierte Carboxylgruppe des AMCA in einer S_N-Reaktion angreifen, eine Amidbindung bilden und dadurch 1-Hydroxy-pyrrolidin-2,5-dion als Abgangsgruppe abgespalten wird. Alternativ können auch die ε-Aminogruppen der Lysine mit AMCA reagieren.



MW = 443,46
Formel des verwendeten Fluorophors = $C_{22}H_{25}N_3O_7$

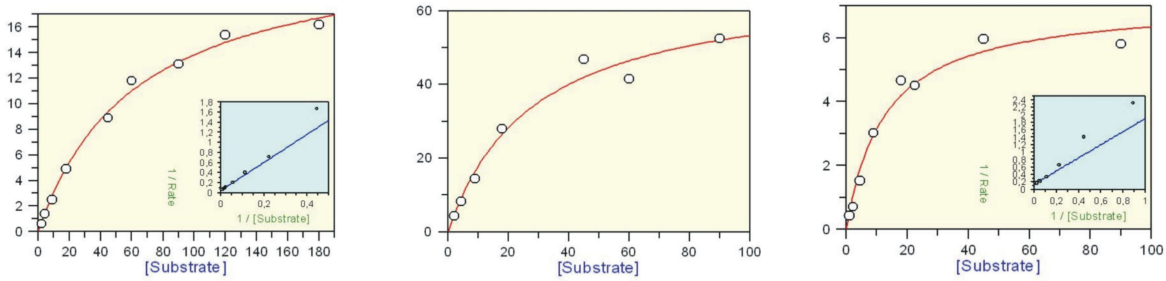
Abbildung 1: 6-((7-amino-4-methylcoumarin-3-acetyl)amino)hexanoic acid (AMCA-X).

Basierend auf diesem Assay konnte die Aktivierung von Proplasmepsin II zu aktivem Plm II nachvollzogen werden. Ebenso gelang es zu zeigen, dass Pepstatin A ein sehr potenter Inhibitor ist (IC_{50}=1 nM), was im Einklang mit Literaturergebnissen steht [1].

Da bei der Bestimmung von IC_{50}-Werten für eine Serie von Azepinderivaten einige unerklärliche Struktur-Wirkbeziehungen festgestellt wurden, entstand das Bestreben, einen Kontroll-Assay zu etablieren. Eine Hypothese für die Entstehung fehlerhafter Assay-Ergebnisse bei nur einigen Verbindungen, könnten Restspuren von ungebundenem AMCA in der Substratlösung sein, die mit manchen Inhibitoren Wechselwirkungen eingehen könnten.

Mit dem fluorogenen Substrat Arg-Gln-Phe*Nphe-Ile-Thr von Bachem, das für HIV-1 Protease-Assays eingesetzt wird, wurde ein Substrat entdeckt, das sich zur Aktivitätsbestimmung von Plm II, Plm IV und Cathepsin D als ebenfalls geeignet erwies. Ein

Vorteil dieses Substrats im Vergleich zu AMCA-Hämoglobin ist das Vorhandensein nur einer Spaltstelle, was eine Auswertung nach Michaelis-Menten ermöglicht. Die Spaltstelle des Substrats liegt zwischen Phenylalanin und p-Nitrophenylalanin (s. Stern). Für alle drei Enzyme wurden die K_M-Werte dieses Substrats bestimmt (Abb.2).



Parameter	Value	Std. Error
Vmax	22,5736	1,1297
Km	63,0175	7,5566

Parameter	Value	Std. Error
Vmax	68,5669	7,0009
Km	28,5464	7,5398

Parameter	Value	Std. Error
Vmax	7,1439	0,4765
Km	12,6477	2,4715

Abb. 2: K_M-Werte von links nach rechts für Plm II, Plm IV und Cat D.

K_i-Werte wurden für aus der Literatur bekannte Plm II-Inhibitoren gemessen, um den Assay weiter zu validieren. Bei der Verwendung von unterschiedlichen Substraten und Geräten innerhalb verschiedener Arbeitsgruppen wird eine Reproduktion der absoluten Werte häufig nicht erreicht. Von größerer Relevanz ist, ob Trends der Affinitäts-Abstufung von Inhibitoren innerhalb einer Verbindungsserie bei unterschiedlichen Assays übereinstimmen. Im Vergleich zu den Literaturwerten sind die hier experimentell bestimmten Messwerte durchschnittlich um einen Faktor ~10 schlechter, wobei die Trends unter den Inhibitoren jedoch erhalten bleiben (Tab. 1). Der K_i-Wert für Pepstatin A zeigt die größte Abweichung, was darauf zurückzuführen ist, dass auf Grund der im Assay eingesetzten Enzymkonzentration (2 nM) keine Affinitätsdaten niedriger als 1 nM gemessen werden können.

Inhibitor	K_i-Wert (exp.)	K_i-Wert (Lit.)	Rel. Unterschied (Faktor)
Pepstatin A	1 nM	0.006 nM [1]	166
GR34	7.7 nM	0.59 nM [2]	13
Saquinavir	7500 nM	750 nM [3]	10
CF-1-25	840 nM	45 nM [4]	18

Tabelle 1: Experimentell bestimmte K_i-Werte ausgewählter Inhibitoren für Plm II im Vergleich zu Literaturwerten.

3.2 Kristallisation und Strukturbestimmung

3.2.1 Kokristallisation von Plm II mit Pepstatin A

Bevor das Ziel einer Kokristallisation von Plm II mit neuartigen Inhibitoren in Angriff genommen wurde, stand zunächst die Reproduktion der in der Literatur bekannten Kristallisationsbedingungen für Plm II in Komplex mit Pepstatin A an [1]. Dadurch sollten Erkenntnisse über ausreichende Reinheit des Proteins gewonnen und zugleich ein für die Kristallisation geeigneter Proteinkonzentrationsbereich gefunden werden. Ein *Screening*, ausgehend von den publizierten Bedingungen, bei denen vor allem die Präzipitanzkonzentration, die Proteinkonzentration und der pH-Wert variiert wurden, führte nicht zum Erfolg.

Von der Literatur abweichende Kristallisationsbedingungen, die aus persönlichen Mitteilungen[1] in Erfahrung gebracht wurden, führten jedoch in wenigen Tagen zu hexagonalen stabförmigen Kristallen (Abb. 3).

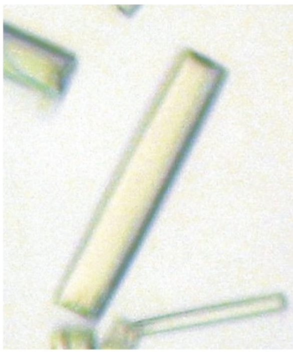


Abb.3: Stabförmige Kristalle von Plasmepsin II im Komplex mit Pepstatin A

Zur Kristallisation wurde die „hanging drop“ Methode angewandt. Plm II in einem 0.015 M Tris Puffer (pH 8.5), wurde auf 10 mg/ml ankonzentriert. Zum Protein wurde Pepstatin A, gelöst in DMSO gegeben, so dass die finale Inhibitorkonzentration 0.7 mM und die End-DMSO Konzentration 10 % betrug. Die Reservoirlösung bestand aus einem 0.1 M MES Puffer (pH 6.3) und 2.7 M $(NH_4)_2SO_4$. Die Kristallisationstropfen setzten sich aus 1 µl der Protein-Inhibitor Lösung, gemischt mit dem gleichen Volumen der Reservoirlösung zusammen. Für die Datensammlung unter Cryobedingungen wurden die Kristalle in eine Cryolösung überführt, die zu 75 % aus Reservoirlösung und zu 25 % Glycerin bestand. Ein In-House Datensatz konnte bis zu einer Auflösung von 2.7 Å aufgenommen werden (Abbildung 4). Das Protein kristallisierte in der Raumgruppe $P3_121$ mit Zellkonstanten von a = b = 104.1 Å und c = 97.6 Å. Zwei Proteinmoleküle befanden sich in der asymmetrischen Einheit.

---

[1] Actelion Pharmaceuticals Ltd.

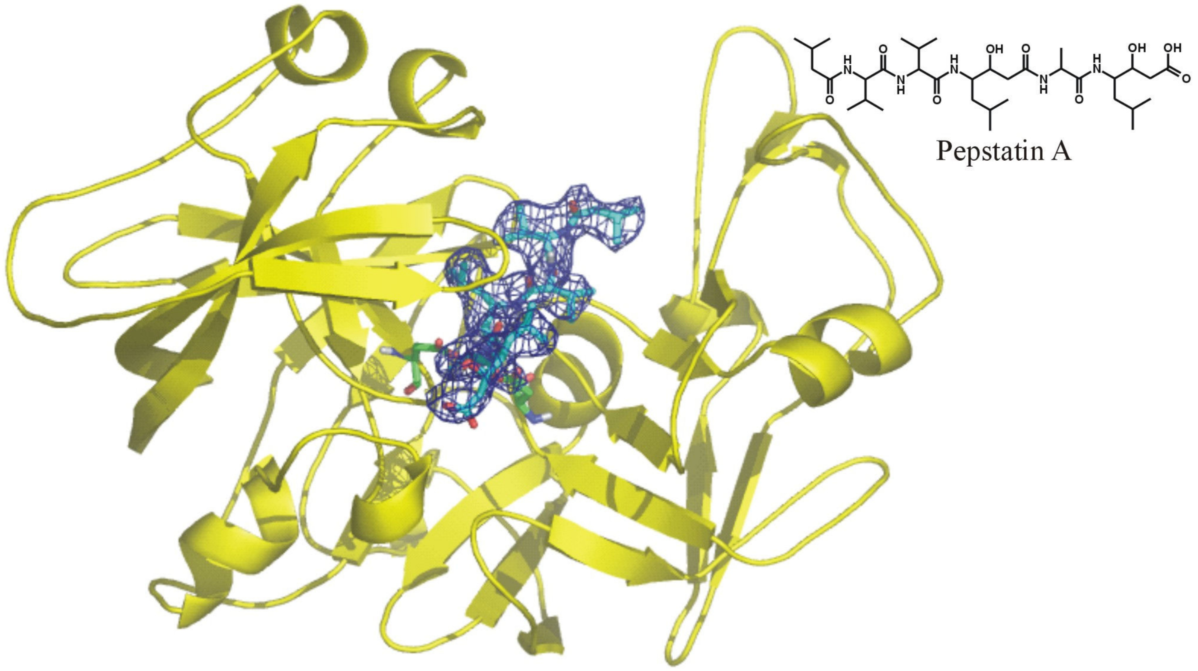


Abbildung 4: Plasmepsin II dargestellt als Cartoon-Repräsentation (gelb) mit Pepstatin A als Ligand (cyan, Stab-Repräsentation). Die Dichte des Inhibitors ist in blau gezeigt. Die beiden Aspartate der katalytischen Diade in grün.

Nachdem Bedingungen zur Kristallisation von Plm II etabliert waren, wurden Kokristallisationsexperimente mit Inhibitoren durchgeführt, die von Kooperationspartnern synthetisiert worden waren. Sie zeigten hohe Affinitäten und Dockingversuche legten das Auftreten neuer Bindungsmodi nahe.

3.2.2 Kristallisation von Plm II mit Azabicyclo[2.2.1]Heptan-Derivaten

In der Arbeitsgruppe von Prof. Dr. F. Diederich wurden Inhibitoren mit einem 7-Azabicyclo[2.2.1]Heptan Grundgerüst entwickelt [4-7]. Die zu Plm II affinsten Verbindungen zeigten IC_{50}-Werte im niedrigen mikromolaren bis hohen nanomolaren Bereich [4]. Unter den oben beschriebenen Bedingungen, die zur Kristallistion von Plm II mit Pepstatin A führten, wuchsen mit diesen Hemmstoffen keine Kristalle. Auch die Verwendung kommerzieller Screens zum Finden neuer Kristallisationsbedingungen (z.B. Hampton Research) verlief erfolglos. Daraufhin wurde ein Screen mit 96 Kristallisationsbedingungen für Plm II zusammengestellt (Tabelle 2), der einerseits alle in der Literatur bekannten Kristallisationsbedingungen einschließt [1, 8-10], zudem aber auch Erfahrungsberichte von mit Plasmepsin arbeitenden Kristallographen berücksichtigt[1]. Dieser speziell für Plm II erstellte Screen lieferte erste Kristalle (Abb. 5), die jedoch stark verwachsen waren und daher nicht zur

---

[1] Actelion Pharmaceuticals Ltd.; J. Yuvaniyama, Mahidol University Bangkok

1 MES 0,1M pH 5,0 PEG3350 30% NH_4OAc 0,3 M	**2** MES 0,1M pH 5,5 PEG3350 30% NH_4OAc 0,3 M	**3** MES 0,1M pH 6,0 PEG3350 30% NH_4OAc 0,3 M	**4** MES 0,1M pH 6,5 PEG3350 30% NH_4OAc 0,3 M	**5** MES 0,1M pH 5,0 PEG3350 25% NH_4OAc 0,3 M
6 MES 0,1M pH 5,5 PEG3350 25% NH_4OAc 0,3 M	**7** MES 0,1M pH 6,0 PEG3350 25% NH_4OAc 0,3 M	**8** MES 0,1M pH 6,5 PEG3350 25% NH_4OAc 0,3 M	**9** MES 0,1M pH 5,0 PEG3350 20% NH_4OAc 0,3 M	**10** MES 0,1M pH 5,5 PEG3350 20% NH_4OAc 0,3 M
11 MES 0,1M pH 6,0 PEG3350 20% NH_4OAc 0,3 M	**12** MES 0,1M pH 6,5 PEG3350 20% NH_4OAc 0,3 M	**13** MES 0,1M pH 5,0 PEG3350 35% NH_4OAc 0,3 M	**14** MES 0,1M pH 5,5 PEG3350 35% NH_4OAc 0,3 M	**15** MES 0,1M pH 6,0 PEG3350 35% NH_4OAc 0,3 M
16 MES 0,1M pH 6,5 PEG3350 35% NH_4OAc 0,3 M	**17** $(NH_4)_2SO_4$ 2M MES 0,1M pH 5,0	**18** $(NH_4)_2SO_4$ 2M MES 0,1M pH 5,5	**19** $(NH_4)_2SO_4$ 2M MES 0,1M pH 6,0	**20** $(NH_4)_2SO_4$ 2M MES 0,1M pH 6,5
21 $(NH_4)_2SO_4$ 2,4M MES 0,1M pH 5,0	**22** $(NH_4)_2SO_4$ 2,4M MES 0,1M pH 5,5	**23** $(NH_4)_2SO_4$ 2,4M MES 0,1M pH 6,0	**24** $(NH_4)_2SO_4$ 2,4M MES 0,1M pH 6,5	**25** $(NH_4)_2SO_4$ 2,8M MES 0,1M pH 5,0
26 $(NH_4)_2SO_4$ 2,8M MES 0,1M pH 5,5	**27** $(NH_4)_2SO_4$ 2,8M MES 0,1M pH 6,0	**28** $(NH_4)_2SO_4$ 2,8M MES 0,1M pH 6,5	**29** $(NH_4)_2SO_4$ 3,2M MES 0,1M pH 5,0	**30** $(NH_4)_2SO_4$ 3,2M MES 0,1M pH 5,5
31 $(NH_4)_2SO_4$ 3,2M MES 0,1M pH 6,0	**32** $(NH_4)_2SO_4$ 3,2M MES 0,1M pH 6,5	**33** 0,2M $(NH_4)_2SO_4$ 0,1M BisTris pH6,5 25% PEG 3350	**34** PEG3350 30% 0,1M BisTris pH6,5 0,2M $(NH_4)_2SO_4$	**35** PEG3350 20% 0,1M BisTris pH6,5 0,2M $(NH_4)_2SO_4$
36 Citronensäure 0,1M pH4,5 0,3 M NH_4OAc PEG 3350 25%	**37** Citronensäure 0,1M pH4,5 0,3 M NH_4OAc PEG 3350 30%	**38** Citronensäure 0,1M pH4,5 0,3 M NH_4OAc PEG 3350 35%	**39** Citronensäure 0,1M pH4,5 0,3 M NH_4OAc PEG 3350 20%	**40** PEG4000 25% 0,3 M Phosphate-buffer 5,0
41 PEG4000 30% 0,3 M Phosphate-buffer 5,0	**42** PEG4000 20% 0,3 M Phosphate-buffer 5,0	**43** PEG4000 35% 0,3 M Phosphate-buffer 5,0	**44** PEG4000 25% 0,3 M Phosphate-buffer 5,5	**45** PEG4000 30% 0,3 M Phosphate-buffer 5,5
46 PEG4000 20% 0,3 M Phosphate-buffer 5,5	**47** PEG4000 35% 0,3 M Phosphate-buffer 5,5	**48** PEG4000 25% 0,3 M Phosphate-buffer 6,0	**49** PEG4000 30% 0,3 M Phosphate-buffer 6,0	**50** PEG4000 20% 0,3 M Phosphate-buffer 6,0
51 PEG4000 35% 0,3 M Phosphate-buffer 6,0	**52** PEG-Ionen-Screen 1	**53** PEG-Ionen-Screen 2	**54** PEG-Ionen-Screen 3	**55** PEG-Ionen-Screen 4
56 PEG-Ion-Screen 5	**57** PEG-Ion-Screen 6	**58** PEG-Ion-Screen 7	**59** PEG-Ion-Screen 8	**60** PEG-Ion-Screen 9
61 PEG-Ion-Screen 10	**62** PEG-Ion-Screen 11	**63** PEG-Ion-Screen 12	**64** PEG-Ion-Screen 13	**65** PEG-Ion-Screen 14
66 PEG-Ion-Screen 15	**67** PEG-Ion-Screen 16	**68** PEG-Ion-Screen 17	**69** PEG-Ion-Screen 19	**70** PEG-Ion-Screen 20
71 PEG-Ion-Screen 21	**72** PEG-Ion-Screen 22	**73** PEG-Ion-Screen 23	**74** PEG-Ion-Screen 24	**75** PEG-Ion-Screen 25
76 PEG-Ion-Screen 26	**77** PEG-Ion-Screen 27	**78** PEG-Ion-Screen 28	**79** PEG-Ion-Screen 29	**80** PEG-Ion-Screen 30
81 PEG-Ion-Screen 31	**82** PEG-Ion-Screen 32	**83** PEG-Ion-Screen 33	**84** PEG-Ion-Screen 34	**85** PEG-Ion-Screen 35
86 PEG-Ion-Screen 36	**87** PEG-Ion-Screen 37	**88** PEG-Ion-Screen 38	**89** PEG-Ion-Screen 39	**90** PEG-Ion-Screen 41
91 PEG-Ion-Screen 43	**92** PEG-Ion-Screen 44	**93** PEG-Ion-Screen 45	**94** PEG-Ion-Screen 46	**95** PEG-Ion-Screen 47
96 PEG-Ion-Screen 48				

Tabelle 2: Plm II Kristallisations-Screen

Strukturbestimmung verwendbar waren. Die Kokristallisationsbedingungen waren reproduzierbar, die Kristallqualität konnte jedoch durch weiteres Screenen nicht verbessert werden. Besonders die lange Zeit von drei Monaten, die die Kristalle zum Wachstum benötigten in der Kombination mit geringen Proteinausbeuten waren im Optimierungsprozess nicht förderlich. Die optisch besten Kristalle wuchsen bei einer Proteinkonzentration von 7.4 mg/ml und einer Inhibitorkonzentration (Inhibitor gelöst in DMSO) von 20 mg/ml. Die Reservoirlösung bestand aus 12 % PEG3350 als Präzipitanz, 0.2 M MES pH 6.5 als Puffer und 0.6 M NH_4OAc. 1 µl Protein/Inhibitorlösung wurde mit 1 µl Reservoirlösung vermischt.

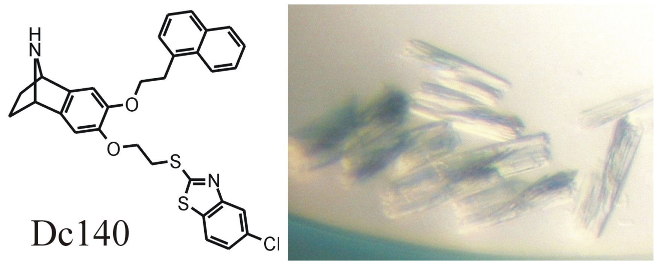


Abbildung 5: Kristalle, die aus der Kokristallisation von PlmII mit DC140 entstanden.

3.2.3 Kristallisation von Plm II mit Pyrrolidin-Derivaten

Von einem weiteren Inhibitor (EA, K_i=630 nM für Plm II), der aus einem neuartigen Pyrrolidin-Grundgerüst bestand und in der AG Klebe synthethisiert worden war, gelang es, Plm II-Kokristalle zu erhalten (Abb.6). Dabei wurde Plm II in einer Konzentration von 7.8 mg/ml eingesetzt und mit dem Inhibitor (20 mg/ml DMSO) für zehn Minuten inkubiert. Protein/Inhibitor-Lösung und Reservoirlösung wurden im gleichen Verhältnis zusammen pipettiert. Die Reservoirlösung setzte sich aus 0.1 M MES pH 6.0, PEG3350 30 % und 0.3 M NH_4OAc zusammen. Die Kristalle entstanden nach ca. drei Monaten und besaßen klar definierte Kantenlängen, die in allen drei Dimensionen eine Länge von 150 µm hatten.

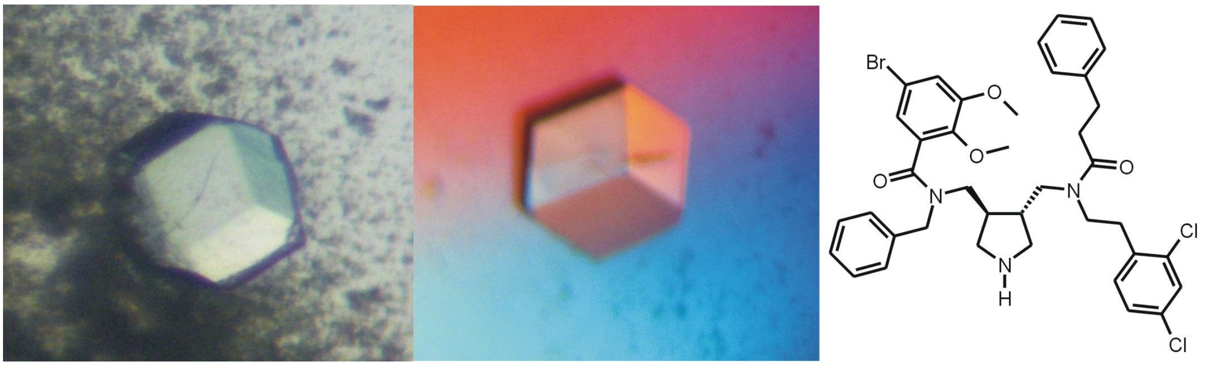


Abbildung 6: Kristalle (links und Mitte), die aus der Kokristallisation von Plm II mit einem Pyrrolidin-Derivat (rechts) (EA) entstanden.

Die Reflexe zeigten eine sehr hohe Mosaizität und eine Auflösung von 3.5 Å konnte nicht übertroffen werden, was insgesamt dazu führte, dass der gesammelte Datensatz sich nicht indizieren ließ (Abb.7).

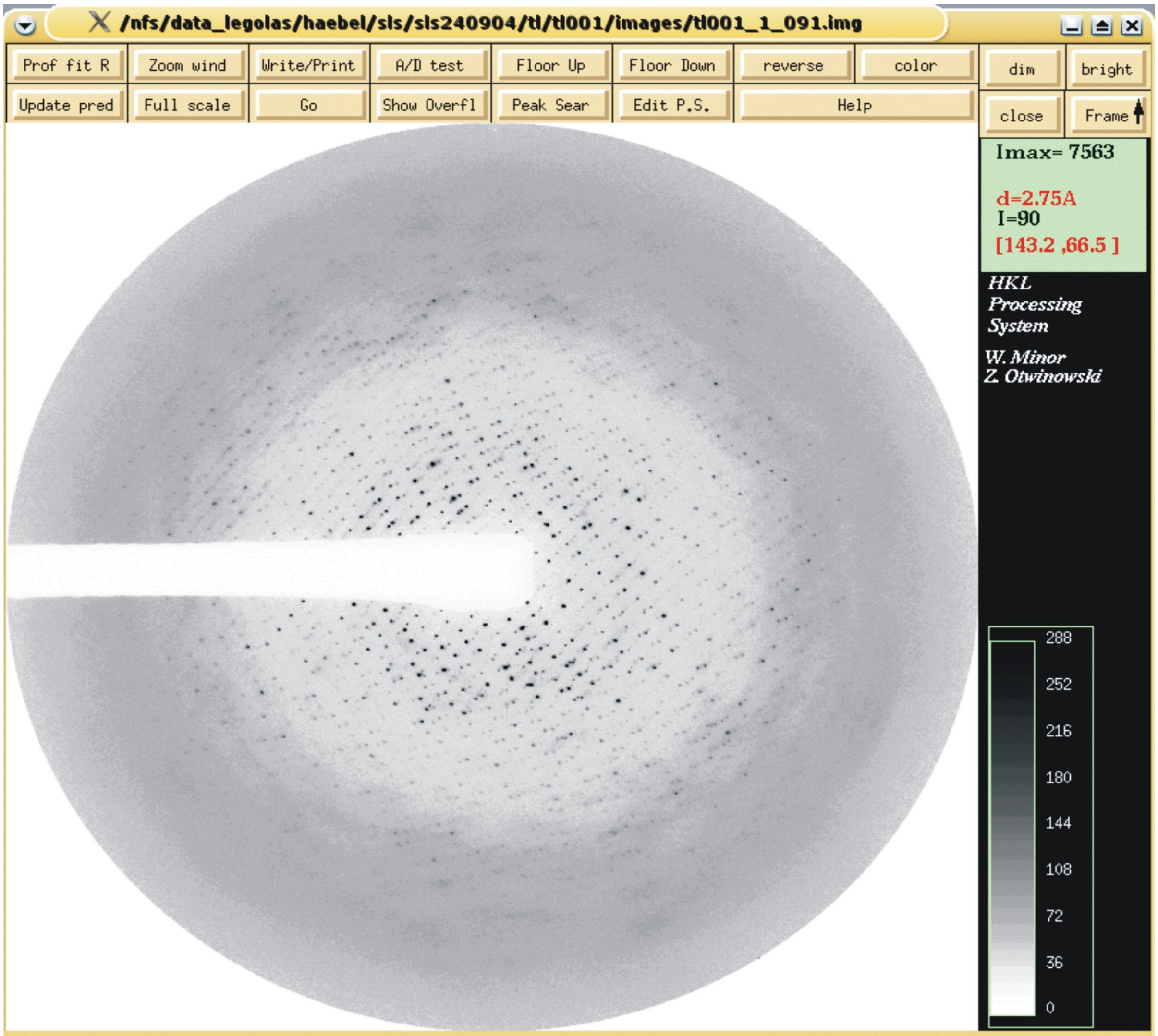

/nfs/data_legolas/haebel/sls/sls240904/tl/tl001/images/tl001_1_091.img
Prof fit R
Zoom wind
Write/Print
A/D test
Floor Up
Floor Down
reverse
color
dim
bright
Update pred
Full scale
Go
Show Overfl
Peak Sear
Edit P.S.
Help
close
Frame
Imax= 7563
d=2.75A
I=90
[143.2 ,66.5]
HKL
Processing
System
W. Minor
Z. Otwinowski
288
252
216
180
144
108
72
36
0

Abbildung 7: Diffraktionsmuster Plasmepsin II in Komplex mit EA.

3.2.4 Kristallisation von Plm II mit Primaquin-Statin Kombinationsinhibitoren

Romeo et al. stellten Plm II Inhibitoren mit einem Primaquin-Statin Grundgerüst her, die im Assay Affinitäten im subnanomolarem Bereich zeigten. Durch Zugabe von mikromolaren Inhibitor-Konzentrationen wurde das *Plasmodium falciparum*-Wachstum in vitro gestoppt.

Um Kristallisationsbedingungen zu finden, wurde ein Screen mit einem Cartesian Kristallisationsroboter angesetzt. Bei einer Bedingung aus insgesamt 288 Ansätzen entstanden nach einem Monat Kristalle, die wie bei der Kokristallisation mit Pepstatin A, stabförmig aussahen (Abb. 8). Die Kristallisationsbedingungen unterschieden sich dennoch deutlich: PEG3350 30 % und 0.2 M $(NH_4)_2SO_4$ dienten als Präzipitanz, 0.1 M BisTris pH 6.5 wurde als Puffer eingesetzt. Die Proteinkonzentration lag bei 9.5 mg/ml, die finale

Inhibitorkonzentration bei 0.8 mg/ml. Da die Tropfengröße bei einem Volumen von 100 nl lag, resultierten daraus kleine Kristalle, deren Volumen für Röntgenbeugungsexperimente nicht ausreichte. Bei Versuchen, die Kristallisationsbedingungen in größerem Maßstab zu reproduzieren, gelang es trotz Mikro- und Makro-Seedings nicht, erneut Kristalle zu gewinnen.

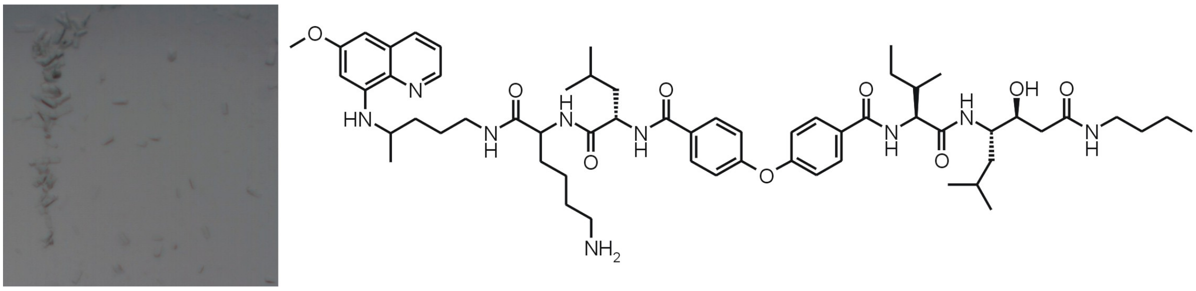


Abbildung 8: Kristalle von Plasmepsin II im Komplex mit einem Primaquin-Statin Kombinationsinhibitor.

3.2.5 Analyse der Proteinreinheit

Durch dankenswerten Einblick in unveröffentlichte Dokumente aus der Firma Roche ließ sich in Erfahrung bringen, dass Proplasmepsin II als Monomer, Dimer und Oligomer vorliegen kann[1]. Bei der Aktivierung zu aktivem Plasmepsin II und der anschließenden Kristallisation könnten die Dimer- und Oligomer-Fraktionen stören, weshalb sie vorher abgetrennt werden sollten. Dieser Aspekt wurde im Folgenden detailliert untersucht:

Bereits während der Anionen-Austausch-Chromatographie entstanden zwei Peaks, die Proplasmepsin II enthielten (Abb. 9 links). Fraktionen aus beiden Peaks wurden per SDS-Gel analysiert, es erschien jeweils eine Bande bei 42 kDa, was der Größe von Proplasmepsin II entspricht. Verunreinigungen waren keine erkennbar (Abb. 9 rechts).

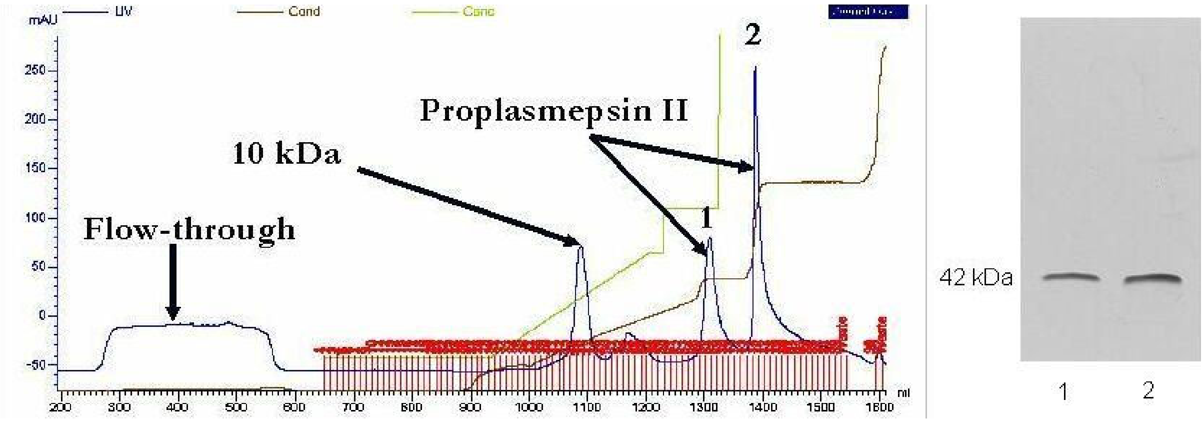


Abbildung 9: Bei der Durchführung der Anionen-Austausch-Chromatographie, wurden zwei Fraktionen erhalten, die Proplasmepsin II enthielten. Die anderen Peaks entsprechen dem Durchfluss und einem Protein mit einer Größe von ca. 10 kDa. Eine anschließende SDS-Gel Analyse verdeutlicht die Proteinreinheit.

---

[1] Roche: unveröffentlichte Ergebnisse

Fraktion 1 wurde anschließend durch eine analytische Größenaustausch-Chromatographie weiter aufgereinigt (Abb. 10). Eine einzelne Fraktion wurde erhalten, die durch Kalibrierung im Vergleich zu Referenzsubstanzen, einem Protein mit einer Größe von ca. 40 kDa entspricht. Somit lag Proplasmepsin II in dieser Fraktion als Monomer vor, Dimer und Oligomer-Fraktionen konnten nicht beobachtet werden.

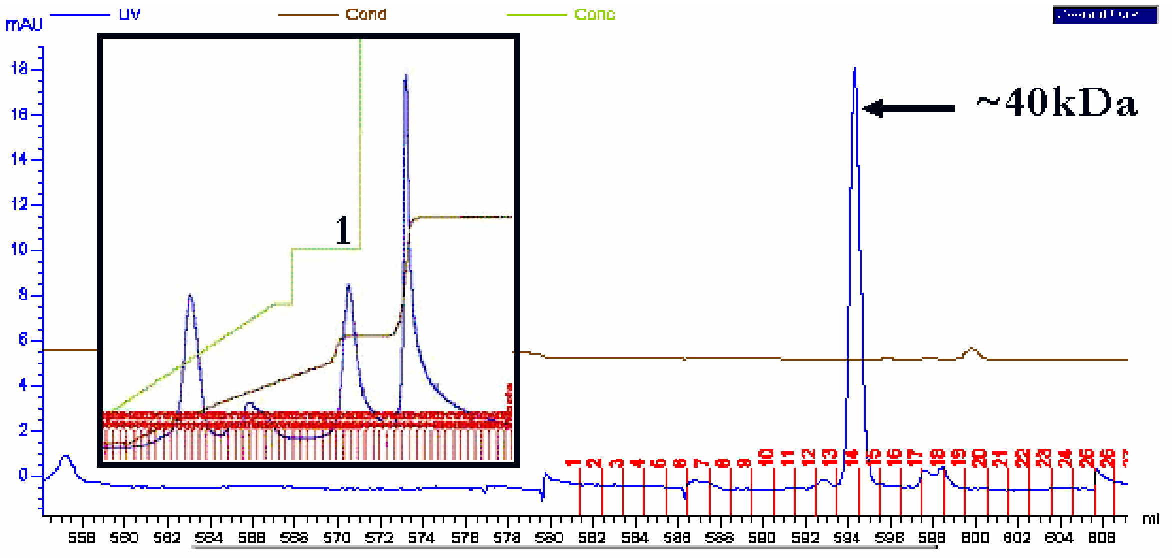


Abbildung 10: analytische Gelfiltration der Fraktion 1, die durch Anionen-Austausch-Chromatographie gewonnen wurde.

Analog dazu wurde Fraktion 2 aufgereinigt. Obwohl, wie im SDS-Gel der Abb. 14 zu erkennen, nur eine Proteinbande nach der Anionen-Austausch-Säule zu sehen war, spaltet das Proplasmepsin II in zwei unterschiedliche Fraktionen auf (Abb. 11). Dabei handelt es sich um die monomere (40 kDa) und dimere (80 kDa) Form von Proplasmepsin II, wie in den unveröffentlichten Ergebnissen von Roche bereits beschrieben.

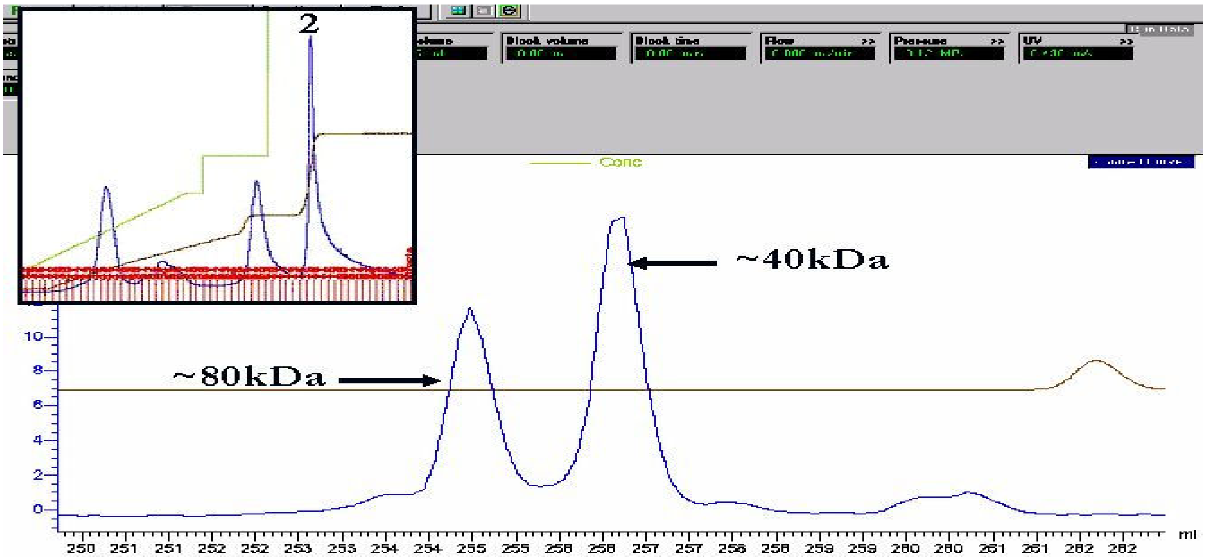


Abb. 11: analytische Gelfiltration der Fraktion 2, die durch Anionen-Austausch-Chromatographie gewonnen wurde.

Sowohl die monomere, als auch die dimere Form von Proplasmepsin II wurden durch pH-Erniedrigung autoaktiviert und ankonzentriert. Schließlich wurde mit Pepstatin A versucht zu kristallisieren. Wie erwartet, entstanden nach wenigen Tagen Kristalle von der monomeren Form. Aber im Gegensatz zu den Beschreibungen aus den unveröffentlichten Dokumenten bildeten sich zeitgleich auch analoge Kristalle aus einer Mutterlauge mit dem aktiven Dimer. Somit scheint ein Vorliegen der Plasmepsine in dimerer bzw. oligomerer Form, keinen Einfluss auf die Kristallisation zu nehmen.

3.3 References

1. Silva AM, Lee AY, Gulnik SV*, et al.* Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from Plasmodium falciparum. *Proc Natl Acad Sci U S A* 1996;**93**(19):10034-9.
2. Romeo S, Dell'Agli M, Parapini S*, et al.* Plasmepsin II inhibition and antiplasmodial activity of Primaquine-Statine 'double-drugs'. *Bioorg Med Chem Lett* 2004;**14**(11):2931-4.
3. Andrews KT, Fairlie DP, Madala PK*, et al.* Potencies of human immunodeficiency virus protease inhibitors in vitro against Plasmodium falciparum and in vivo against murine malaria. *Antimicrob Agents Chemother* 2006;**50**(2):639-48.
4. Hof F, Schutz A, Fah C*, et al.* Starving the malaria parasite: Inhibitors active against the aspartic proteases plasmepsins I, II, and IV. *Angewandte Chemie-International Edition* 2006;**45**(13):2138-2141.
5. Carcache DA, Hörtner SR, Bertogg A*, et al.* De Novo Design, Synthesis, and In Vitro Evalution of a New Class of Nonpeptidic Inhibitors of the Malarial Enzyme Plasmepsin II. *ChemBioChem* 2002;**3**(11):1137-1141.
6. Carcache DA, Hörtner SR, Bertogg A*, et al.* A New Class of Inhibitors for the Malarial Aspartic Protease Plasmepsin II Based on a Central 111-Azatricyclo[6.2.1.0 2,7]undeca-2,4,6-triene Scaffold. *Helvetica Chimica Acta* 2003;**86**:2192-2209.
7. Carcache DA, Hörtner SR, Seiler P*, et al.* Development of a New Class of Inhibitors for the Malarial Aspartic Protease Plasmepsin II Based on a Central 7-Azabicyclo[2.2.1]heptane Scaffold. *Helvetica Chimica Acta* 2003;**86**:2173-2191.
8. Asojo OA, Afonina E, Gulnik SV*, et al.* Structures of Ser205 mutant plasmepsin II from Plasmodium falciparum at 1.8 A in complex with the inhibitors rs367 and rs370. *Acta Crystallographica. Section D, Biological Crystallography* 2002;**58**(Pt 12):2001-8.
9. Prade L, Jones AF, Boss C*, et al.* X-ray Structure of Plasmepsin II Complexed with a Potent Achiral Inhibitor. *J. Biol. Chem.* 2005;**280**(25):23837-23843.
10. Asojo OA, Gulnik SV, Afonina E*, et al.* Novel uncomplexed and complexed structures of plasmepsin II, an aspartic protease from Plasmodium falciparum. *Journal of Molecular Biology* 2003;**327**(1):173-81.

4 Computer-aided Design and Synthesis of non-peptidic Plasmepsin II and IV inhibitors [1]

4.1 Abstract

Plasmepsin (Plm) II (EC-number: 3.4.23.39) and IV (EC-number: 3.4.23.B14) are aspartic proteases present in the food vacuole of the malaria parasite *Plasmodium falciparum* and are involved in host hemoglobin degradation. Based on our established efficient synthetic sequence, a series of inhibitors for Plm II and IV has been synthesized bearing a 2,3,4,7-tetrahydro-1*H*-azepine scaffold as the core structural element. During the computational design cycle, thorough investigations have been carried out in order to find a reasonable theoretical binding mode for Plm II and Plm IV, respectively. The conformation of Plm II in the crystal structure 1LF2 provides a good starting geometry for our virtual screening approach. In contrast, the solely available co-crystal structure for Plm IV (pdb-code: 1LS5) appears inappropriate for inhibitor design. Therefore, a homology model was constructed based on the Plm II 1LF2 structure. A combinatorial docking run using FlexX^c suggested compounds which, after synthesis, turned out to exhibit affinities in the submicromolar range. The observed structure-activity relationships of the synthesized compounds confirm the assumed binding mode for Plm II and IV. The best binding inhibitors designed for Plm II and IV are devoid of any inhibitory potency against human cathepsin D (EC 3.4.23.5).

---

[1] submitted to ChemMedChem, authored by T. Luksch, N.-S. Chan, S. Brass, C. A. Sotriffer, G. Klebe, W. E. Diederich, paper presently in revision
Computational design and Biological Evaluation were performed by T. Luksch

4.2 Introduction

The protozoan disease malaria is still one of the major inexorable infections worldwide, especially widespread throughout tropical regions. It is caused by parasites of the genus *Plasmodium* (*P. falciparum*, *P. vivax*, *P. malariae* and *P. ovale*) and transmitted by female *Anopheles* mosquitoes. The WHO estimates that approximately 350 to 500 million people are infected annually [1]. Infections by *P. falciparum*, the most severe form of malaria, are responsible or at least contribute to more than a million deaths per year. Especially children below an age of five years are affected [1]. The continuously increasing resistance of the vector towards insecticides as well as the emergence of multi-drug resistant parasites highly demands the development of new anti-infective substances favourably showing either a novel binding mode or a new mechanism of action, which is pivotal to combat this life-threatening disease.

During its life cycle, *P. falciparum* degrades the hemoglobin of the host cell to cover the energy and nutrition demands for its own growth and maturation. In this metabolic pathway, the following proteolytic enzymes are hitherto known to be involved: Aspartic proteases (plasmepsin (Plm) I [2, 3], II [4, 5], IV [6], and the closely related histoaspartic protease (HAP) [6, 7]), cysteine proteases (falcipain-1 [8, 9], -2 [10, 11], and -3 [12]), a metalloprotease (falcilysin) [13], and the recently discovered dipeptidyl aminopeptidase 1 (DPAP1) [14]. The degradation process appears to follow an ordered pathway [4, 13]. Although the precise sequence of events, especially whether a plasmepsin (Plm) or a falcipain catalyzes the initial degradation step [4, 15], is still under debate, it has been shown that malaria parasites, both in culture and in animal models, are eradicated by Plm inhibitors, thus providing a proof-of-concept that these proteases are viable drug targets [2, 16, 17]. Recent experiments with *P. falciparum* Plm knockout clones indicate the necessity of simultaneously inhibiting several of the food vacuole Plms in order to effectively tackle the parasite [18, 19]. Designing compounds that are able to inhibit several Plms in parallel seems to be inevitable not only with respect to an efficient eradication of the parasites, but also to impede the development of parasite resistance.

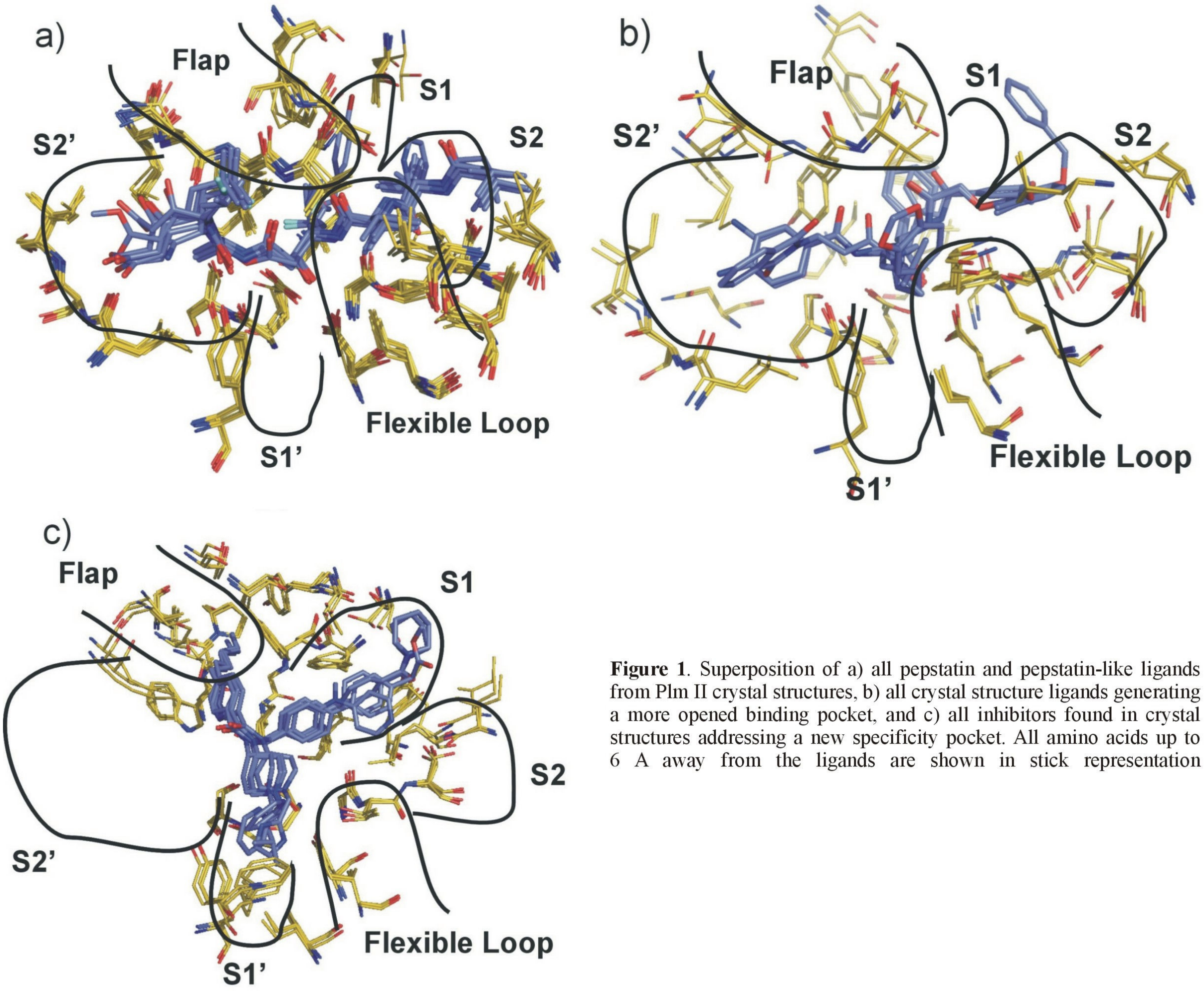


Figure 1. Superposition of a) all pepstatin and pepstatin-like ligands from Plm II crystal structures, b) all crystal structure ligands generating a more opened binding pocket, and c) all inhibitors found in crystal structures addressing a new specificity pocket. All amino acids up to 6 A away from the ligands are shown in stick representation

Most inhibitors synthesized so far to address Plm II are peptidomimetic transition state analogues [20] addressing via a hydroxyl or hydroxyl-like moiety the two aspartates of the active site thereby replacing the native substrate. These peptidomimetic inhibitors are highly active but often suffer from low bioavailability and, additionally, are difficult to synthesize. Several years ago, Boss et al. identified by high-throughput screening non-peptidic inhibitors for renin [21], a structurally related aspartic protease with a sequence identity of 35 % to Plm II. These inhibitors all feature a piperidine scaffold which addresses via its basic and most likely protonated secondary endocyclic amino functionality the two catalytic aspartates. Interestingly, some of these inhibitors also showed moderate activity against Plm I and II. Recently, substituted pyrrolidines have been described as micromolar inhibitors of the aspartic protease HIV-1 [22]. Diederich et al. identified a 7-azabicyclo[2.2.1]heptane scaffold as an adequate core structure for the development of Plm II and IV inhibitors [23, 24]. These results prompted us to investigate in greater detail the suitability of azacycles bearing a basic amino functionality as core element for the design and synthesis of Plm inhibitors.

At present, 14 Plm II inhibitor complex structures of *P. falciparum* have been deposited in the PDB, which can be classified in three groups exhibiting different ligand binding modes: Eight complexes contain either pepstatin A or a closely related pepstatin A analogue as inhibitors.

They exhibit as transition-state analogues an almost identical ligand binding mode (pdb-codes: 1SME, 1ME6, 1XDH, 1XE5, 1XE6, 1W6I, 1W6H, 1M43) (Fig. 1a). The flap lid shared in common by all aspartic proteases folds upon the bound inhibitor and displays the binding pocket in closed state. The second group comprises three complexes (1LF2, 1LEE, 1LF3), all sharing the binding pocket in a partially open conformation (Fig. 1b). In the third group, a new specificity pocket, the so-called flap pocket, is unlocked and addressed by the inhibitors (2BJU, 2IGY, 2IGX) (Fig. 1c) [25-28].

The crystallographically indicated adaptivity of the protein is further confirmed by molecular dynamics (MD) simulations [29-32]. They suggest that Plm II and most likely also Plm I and IV are highly flexible proteins adopting additional conformations not known so far but which could possibly be targeted by inhibitors [33].

Due to the observation that only three distinct binding modes have been discovered so far along with the fact that the target protein has been treated as rigid by most in silico studies, the application of automated docking procedures appears rather limited. Additionally, some cases have been reported in literature where docking failed to predict reasonable binding modes based on previously determined crystal structures [31, 32, 34]. The situation appears even worse for *P. falciparum* Plm IV, as for this enzyme only a single co-crystal structure in complex with pepstatin A is known (pdb-code: 1LS5) [27]. For Plm I, merely a homology model has been described so far. An in-depth search of the literature indicates that a possible explanation for the rather poorly described structural properties of these enzymes might originate from difficulties to grow sufficiently well diffracting crystals for structure determination of complexes with inhibitors not adopting a pepstatin-like binding mode. However, as our goal is the design and synthesis of promiscuous inhibitors addressing Plm II and IV equally well, we planned an azacyclic scaffold as suitable geometry to be docked into one of the currently available crystal structures.

4.3 Results and Discussion

4.3.1 In silico evaluation of the core structure

Preliminary modeling studies suggested that a 2,3,4,7-tetrahydro-1*H*-azepine scaffold could provide a promising starting point for the design of putative Plm inhibitors. As depicted in Figure 2, the azacycle could address via its basic nitrogen the two aspartates of the catalytic dyad.

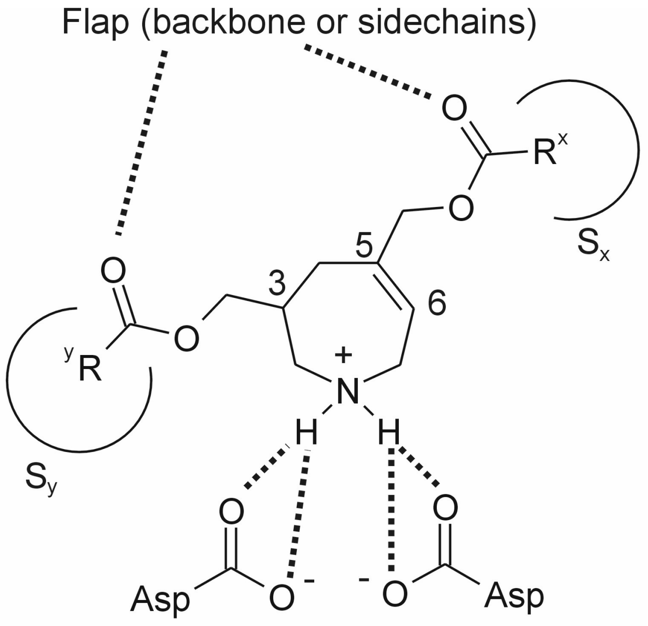


Figure 2. 3,5-Bis-carbonyloxymethyl-substituted azepine template and its putative hydrogen bond network within the active site of Plm.

Equipped with appropriate linker groups in position 3 and 5 respectively, the introduction of well-tailored substituents to explore and specifically address the different substrate recognition pockets is easily amenable. As a starting point we selected 3,5-bis-hydroxmethyl-2,3,4,7-tetrahydro-1*H*-azepine. The primary hydroxyl functionalities at this core structure can readily be esterified using standard coupling reactions. The carbonyl oxygens of both ester groups are thought to be well placed to establish hydrogen bonds to the flap region. Following our synthetic strategy, the azepine derivatives are obtained as racemic mixtures resulting from the evolving stereo center at position 3 of the ring system [35]. To obtain a deeper insight into the putative binding mode of the *R*- and the *S*-enantiomer, docking with FlexX [36] was performed: As reference geometries the protein conformers found in the crystal structures 1SME (pepstatin-like), 2BJU (flap pocket), and 1LF2 (opened binding pocket) were used. Both enantiomers were docked into the three protein conformers (Table 1).

For all reference protein conformers, the *R*-enantiomer obtained better scores with respect to the FlexX scoring function. Particularly, in the conformer displayed by the crystal structure 1LF2, the *R*-enantiomer received significantly better scoring compared to the other two Plm conformers. Closer investigation of the crystal structure 1LF2 with its bound parent inhibitor RS367 revealed that the hydroxyl group of the hydroxypropylamine isostere establishes hydrogen bonds to the catalytic dyad (Figure 3a). In addition, hydrogen bonds are formed to the flap residues V78 and S79.

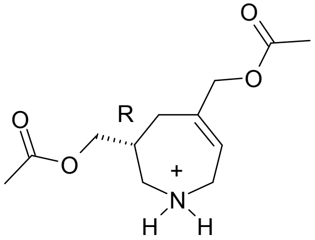
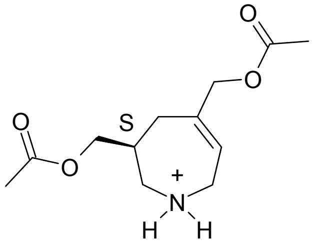
Azepine enantiomers		
PDB-code	Relative score	
1SME	-18.5	-17.8
2BJU	-16.1	-14.2
1LF2	-24.8	-18.2

Table 1. Scores of the FlexX^c scoring function generated for the *R*- (left) and *S*-enantiomer (right) as docked into three different Plm II crystal structure geometries (1SME, 2BJU and 1LF2, respectively).

A comparison of a docking pose of our azepine derivative (Figure 3b) with the crystal structure of the RS367 ligand suggests a similar hydrogen bond pattern: In our case, the quaternary amino group of the azepine moiety interacts via charge-assisted H-bonds with the two aspartates of the catalytic dyad. The carbonyl oxygen of the ester function at position 3 addresses the backbone NH of V78 in the flap whereas the ester carbonyl oxygen at position 5 forms a hydrogen bond to the hydroxyl group of S79 in the flap region. In this docking solution, the two terminal methyl groups are oriented towards the S1 and S2' pocket. Considering the placement of the methoxy bond vectors, a variety of different ester substituents can be added to properly address the corresponding subpockets.

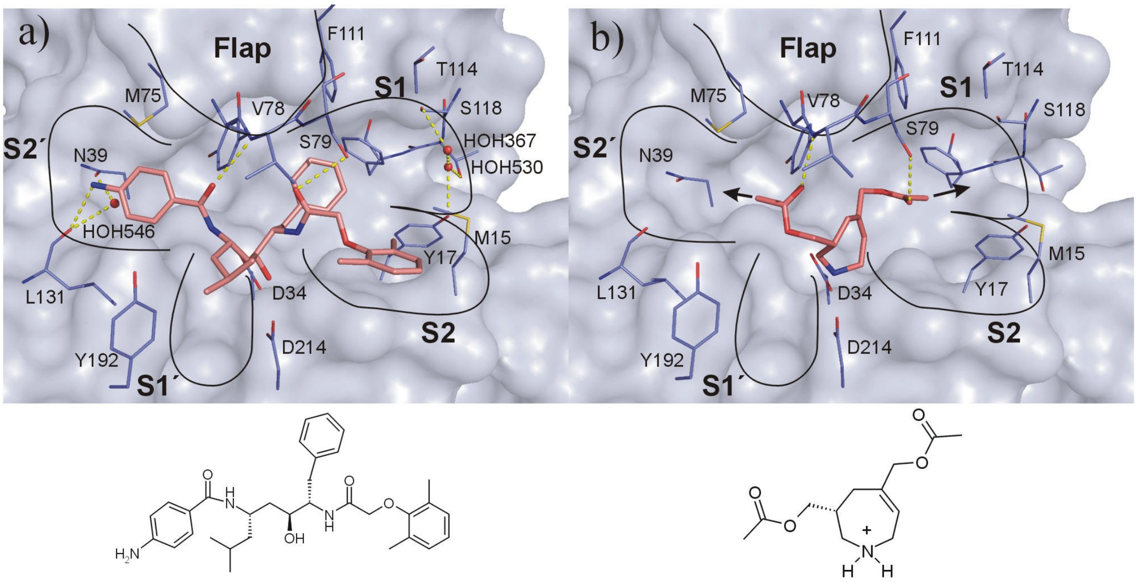


Figure 3. a) Crystal structure 1LF2 in complex with RS367. Hydrogen bonds to the flap amino acids V78 and S79 are formed. The catalytic dyad (D34 and D214) is addressed by a hydroxyl group. The S1, S2, S1' and S2' pockets are occupied by the inhibitor. b) Docking solution of the azepine core structure (*R*-enantiomer) in 1LF2. Hydrogen bonds to the flap as well as to the catalytic dyad are formed. The two substituents point towards the S1 and S2' pocket.

In order to verify the relevance of the considered ring conformation of the azepine moiety generated by FlexX, a search in the Cambridge Structural Database (CSD) [37] was performed with ConQuest [38]. A small molecule crystal structure could be identified also comprising a 2,3,4,7-tetrahydro-1*H*-azepine scaffold (refcode: CCDC-260740) [39]. The seven-membered ring in this crystal structure adopts, similarly to the docking solution, a boat conformation (Fig. 4).

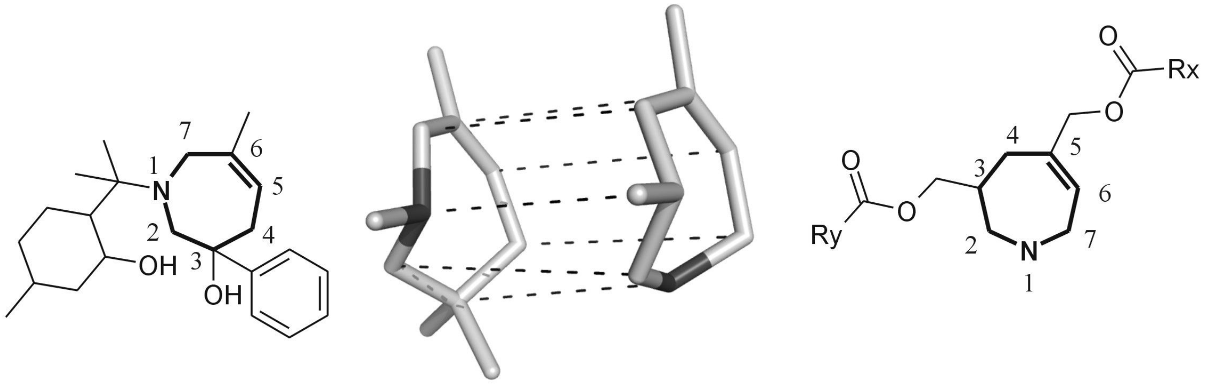


Figure 4. Pairwise matching of the considered ring conformation, generated by FlexX onto the ring atoms of a small molecule crystal structure (refcode: CCDC-260740). The rmsd of both rings is 0.27 Å.

A pairwise matching of both ring scaffolds was performed as depicted in Figure 4: Atoms of both core structures with corresponding geometries were aligned resulting in a low rmsd of 0.27 Å. Semi-empirical calculations based on the AM-1 method (MOPAC 5.1) [40] were performed in order to estimate the energy difference of both ring structures. The small energy difference of 0.8 kJ/mol indicates that a reasonable low-energy ring conformer has been computed by Corina [41] and used for docking with FlexX.

4.3.2 Docking starting geometry for Plm IV

For Plm IV of *P. falciparum*, only a single co-crystal structure with pepstatin A is available [27]. The binding mode of the inhibitor in this complex is very similar to that of pepstatin A in Plm II [25] also resulting in a binding-site conformer with closed flap lids. Based on the docking results of our core structure with the three distinct protein conformations which suggest preferred binding to the open conformer of Plm II, we propose that also in case of Plm IV a more open protein conformation is required to accommodate our azepine derivatives. Plm IV exhibits a high sequence identity of 69 % with Plm II. In the whole binding pocket, 24 residues out of 34 are conserved considering all residues up to 6 Å distant from the ligand RS367. Thus, a reasonable homology model of Plm IV could be generated in the more opened conformation based on the coordinates of the complex structure of Plm II (1LF2) as template (Figure 5).

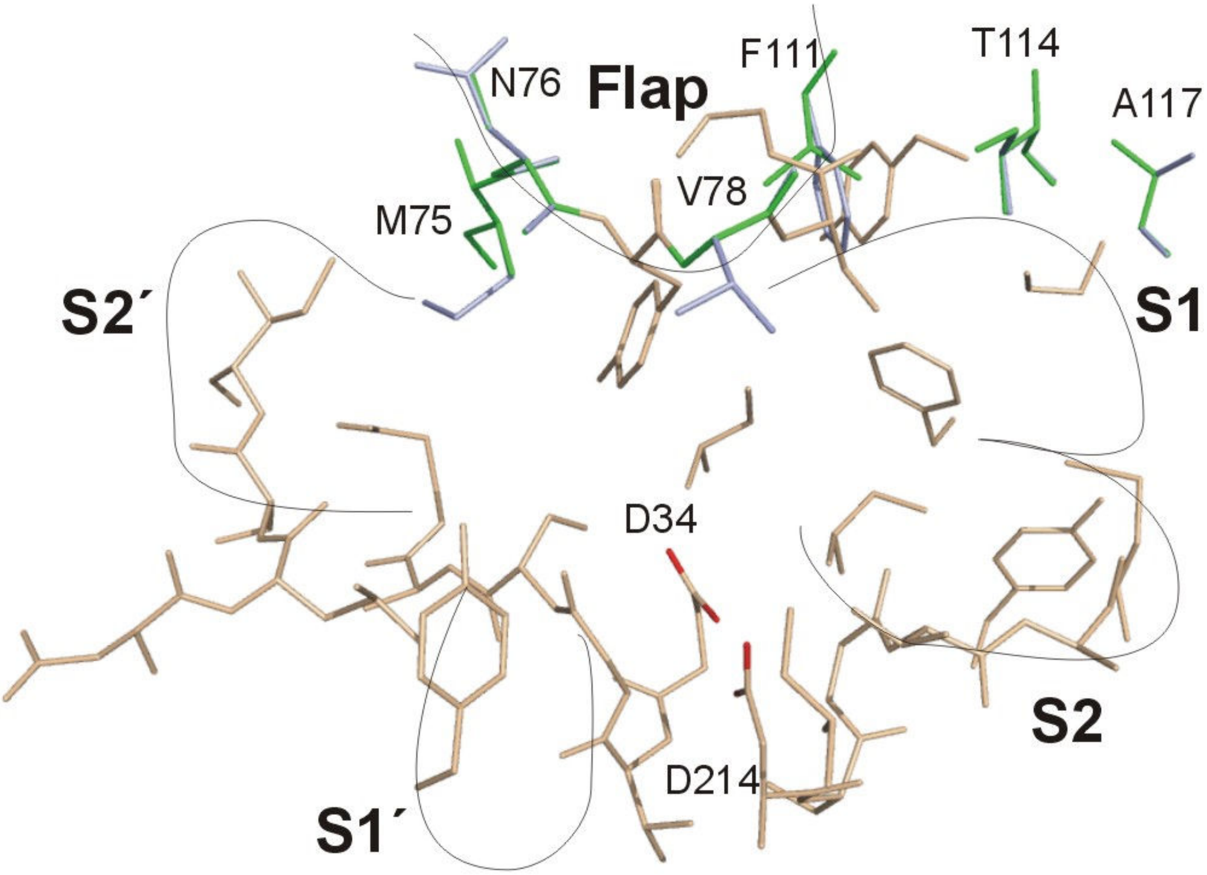
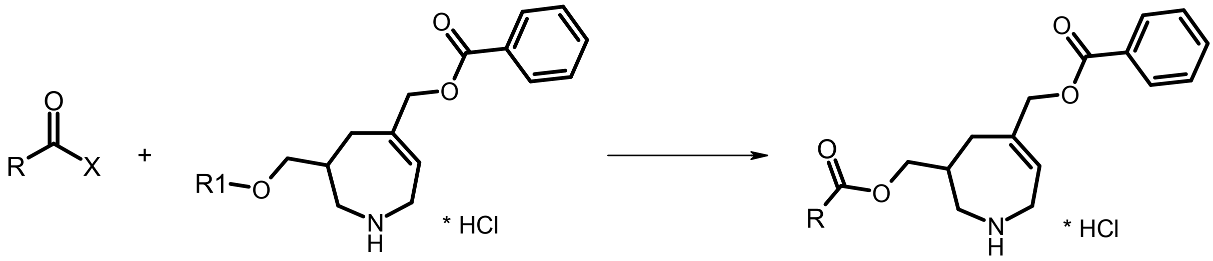


Figure 5. Homology model binding pocket of Plm IV based on the template 1LF2. Identical residues of Plm II and Plm IV are colored in beige, residues only present in Plm II are colored in blue, residues of Plm IV in green.

A superposition of the modeled Plm IV structure with the template 1LF2 reveals the differences (light blue and green) and correspondencies (beige) of the two binding pockets: In the S2' pocket of Plm II (light blue), N76 is replaced by a serine in Plm IV (green). Since the side-chains of N76 in Plm II and the corresponding serine in Plm IV are solvent-exposed, the shape of the binding pocket is not strongly altered. It has been shown that the replacement of M75 in Plm II by an isoleucine in Plm IV can have an influence on the accommodation of putative P2' substituents [42]. Located at the tip of the flap in Plm II, V78 corresponds to the less bulky G78 in Plm IV. In the S1 pocket, F111 is replaced by a leucine, T114 by an isoleucine, and A117 by a glycine, thus creating a slightly larger binding pocket. Apart from these residues, the binding pocket is highly conserved (identical residues for Plm II and IV in beige) so that it should be feasible to design inhibitors addressing both, Plm II and Plm IV, with almost identical binding modes.

4.3.3 Optimization of the P1 and P2' residues

Docking of the parent core structure suggested preferred binding of the *R*-enantiomer of our azepine to the open protein conformer (see above). Putative substitution vectors (Rx,Ry) (Figure 4) point towards the S2' and S1 pocket (Figure 2 and 3b). In the following, these pockets were thoroughly analyzed to identify putative substituents to optimally address both specificity pockets. To create some ideas of possible occupants for these pockets, we first performed a database search using *Cavbase* to retrieve similar subpockets present in other crystal structures [43, 44]. *Cavbase* is a method to describe and compare protein binding pockets in terms of shape and exposed physicochemical properties. As a matter of fact, most pockets exhibiting a pronounced physicochemical similarity to the Plm II recognition pockets also turned out to be members of the family of aspartic proteases. The molecular portions of ligands found to occupy similar subpockets were extracted and analyzed with respect to their suitability as building blocks for the synthesis of putative inhibitors. In approximately one third of all retrieved pockets exhibiting similarity to the Plm II S1 pocket, consistently a phenyl moiety was found (31 of 100 similar pockets). A similar search performed for the S2' pocket was less conclusive and a broad spectrum of structurally rather diverse molecular building blocks was suggested. This indicates a less pronounced specificity of this pocket for a single preferred occupant. As we planned to perform our design in a step-wise fashion, we decided to use at first a phenyl moiety as most suitable occupant to address the Plm II S1 pocket. The S2' pocket, as indicated by our *Cavbase* search, is less discriminative for a particular substituent. We therefore decided to perform a rather exhaustive screen for putative occupants using combinatorial docking with FlexX^c. From a synthetic point of view, we intended to apply ester bond formation. In consequence, we retrieved a diverse set of 2083 carboxylic acids from the Sigma-Aldrich catalogue. All considered acid fragments had a molecular weight below 250 Da. Based on the benzoyloxymethyl-azepine core structure, a library of 2083 esters was generated on the computer (Scheme 1):



Scheme 1. A library of 2083 esters was generated by linking the core with the fragments.

Subsequent to docking, a visual inspection of the 100 best-scored candidates was performed. The parent crystal structure 1LF2 with RS367 shows three water molecules mediating interactions to the bound ligand or to the protein. These water molecules present in the S1 as well as in the S2' pocket possess rather low B-factors. Similarly placed water molecules could potentially influence the binding modes of our docked library entries. HOH546 with a B factor below 10 Å^2, located in the S2' pocket, establishes hydrogen bonds to the N39 side-chain and to the carbonyl oxygen of L131. HOH367 (B-factor = 24 Å^2) and HOH530 (B-factor = 19 Å^2) both occupy the S1 pocket and mediate a hydrogen-bonding network between S118, Y17, and T119 (Figure 3a). As these water molecules were not considered during our initial combinatorial docking run, we added them in an additional energy minimization step in order to optimize the ligand geometries of the best scored candidate ligands in the presence of these water molecules. During minimization with the Moloc MAB force field [45], the protein and water atoms were kept rigid, whereas the ligand was assigned to full flexibility.

4.3.4 Analysis of the docked ligands

In almost all docking solutions the protonated amino functionality of the azepine core interacts with the two aspartates of the catalytic dyad as suggested by our initial docking runs (Figure 3b). As desired, the ester functionality in position 5 forms a hydrogen bond to the flap residue S79. Its phenyl ring accommodates the predominantly hydrophobic S1 pocket composed by M15, I32, F111, and F120 (Figure 6).

The carbonyl oxygen of the ester functionality in 3-position forms a hydrogen bond to the flap backbone NH of V78. As expected, in almost all docking solutions the attached R groups of the 2083 library members accommodate the S2' pocket. Among the best-scored solutions, preferentially aromatic ring moieties bearing hydroxyl or amino substituents at different ring positions were detected (Figure 6). Particularly, entries with an *o*-hydroxyphenyl substituent appeared to form hydrogen bonds to the backbone carbonyl oxygen of G36 and the hydroxyl group of Y192. Additional hydrogen bonds to the backbone carbonyl oxygen of L131 are suggested to be formed by a *m*-aminophenyl, *p*-aminophenyl or a *m*-hydroxyphenyl substituent. Furthermore, the backbone amide function of N76 could be addressed by a *m*-nitrophenyl moiety. Likewise, the side-chain of N39 could possibly interact with a *p*-amino-substituted phenyl ring. In summary, five amino acids (G36, Y192, N39, N76, L131) were identified in the S2' pocket that might be involved in hydrogen bonding to putative donor or acceptor functionalities of potential inhibitors.

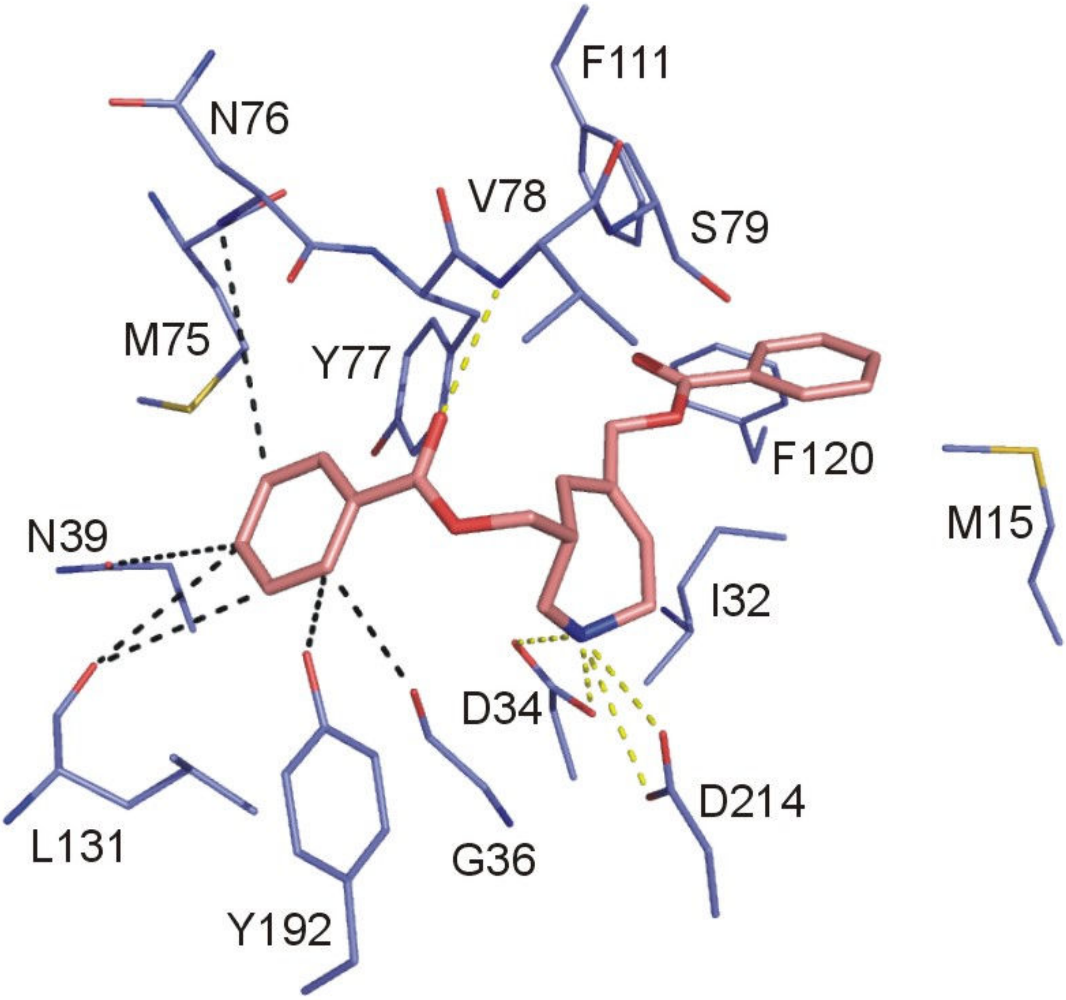


Figure 6. Substituents attached to the phenyl moiety such as hydroxyl or amino groups in different positions will probably lead to further hydrogen bond formation in the S2' pocket (indicated by black dotted lines). Hydrogen bonds are colored as yellow dotted lines.

A similar combinatorial docking followed by a subsequent energy minimization step was performed for the Plm IV homology model using the same library of candidate ligands as for Plm II. The docking results showed very similar binding poses and preferences: The phenyl moiety at position 5 accommodates in an alike fashion the S1 pocket. For the P2' substituent, aromatic ring systems equipped with hydroxyl or amino groups at different ring positions are also highly scored for this isozyme, particularly as the conserved residues G36, Y192, N39, S76, L131 are similarly involved in hydrogen bonding.

4.3.5 Synthesis and biological testing of the first design series

Based on the design hypothesis derived from our first cycle, additionally considering the synthetic feasibility of the proposed molecules, inhibitors **1** - **3** were prepared according to the synthesis scheme described below (Schemes 3, 4, and ref [35]). The compounds were tested for activity against Plm II, Plm IV, and the related human aspartic protease cathepsin D (Table 2). Compound **1** showed a K_i value of 4.7 µM for Plm II and 7.2 µM for Plm IV. In agreement with our modeling results, the importance of the amino group is demonstrated by the 60-fold affinity difference between **1** and **2**. The latter lacks a similar H-bond donor group.

In addition, the synthetic precursor of **1**, compound **3** bearing a *p*-nitrobenzoyloxymethyl substituent in position 3 of the azepine ring, was tested for Plm inhibition. However, in contrast to our prediction from docking, its K_i values for Plm II and Plm IV are unexpectedly low (Plm II K_i=29.0 µM; Plm IV K_i=38.6 µM). This compound received a worse scoring compared to the unsubstituted **2**. Assuming a similar binding pose of **3** compared to **1** (Figure 8a) and **2**, the *p*-nitrobenzoyloxymethyl moiety should experience repulsive interactions of its nitro group with the side-chain amide oxygen of N39 in the S2’ pocket.

To seek for a possible explanation of this unexpected finding, molecular dynamics (MD) simulations of the uncomplexed protein were performed. Along the trajectory, pronounced side-chain movements of N39 in the S2’ pocket could be observed. In addition to the conformation found in the crystal structure, a further low energy conformer seems to be significantly populated (Figure 7).

In this second conformation, the Nχ2 angle is rotated by 180°. Based on this alternative conformer, a reasonable binding geometry of **3** can be suggested (Figure 8b). The azepine moiety and the P1 substituent of **3** both remain in a similar fashion bound to the protein. However, the *p*-nitro group of the P2’ substituent replaces HOH546 and forms a direct hydrogen bond to the terminal carboxamide NH of the rotated N39 residue. This binding pose would be in agreement with the enhanced binding affinity of the nitro derivative compared to the unsubstituted derivative **2**.

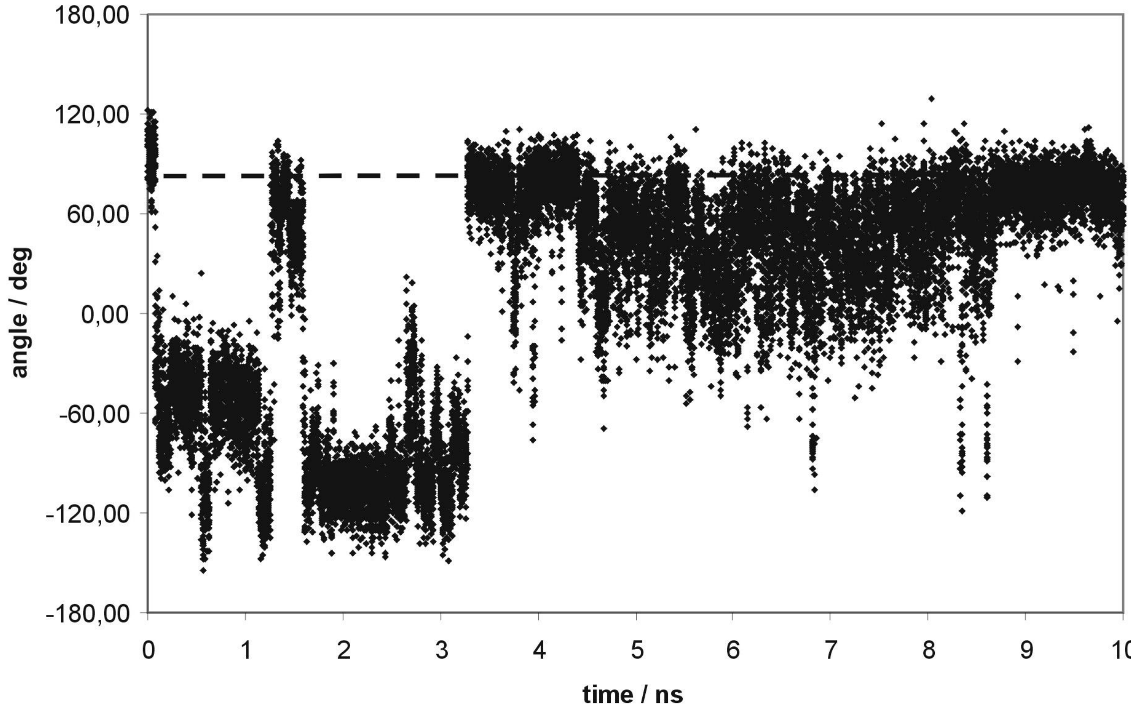
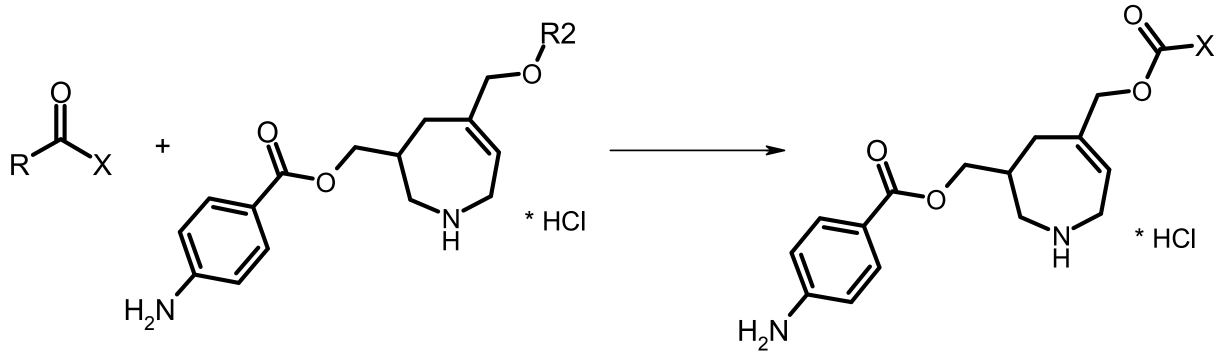


Figure 7. MD simulation of the fluctuation of the Nχ2 angle in asparagine 39 (N39) of Plm II, with the respective value from the crystal structure shown as reference; for details, see Experimental Section.

A second design cycle was performed now optimizing the substituent for the S1 pocket keeping the previously detected P2' substituents unchanged. Again, ester bond formation was anticipated for synthetic reasons. A similar combinatorial docking run using FlexX^c was carried out following a subsequent energy minimization step (Scheme 2).



Scheme 2. A library of 2083 esters was generated by linking the core with the fragments.

Table 2. Experimental Inhibitory Activities of the Compounds Studied

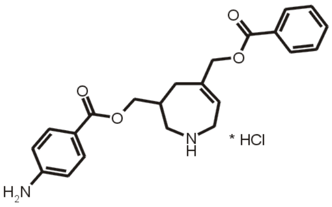
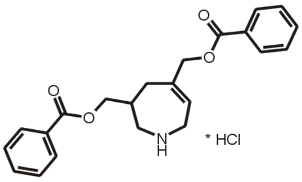
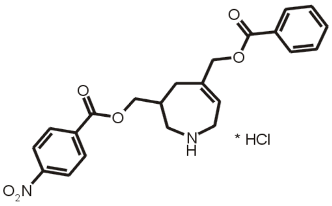
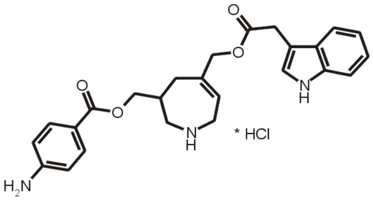
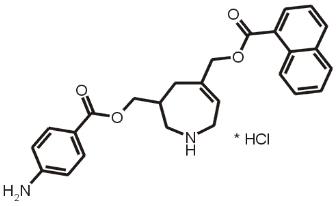
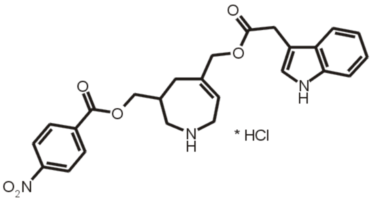
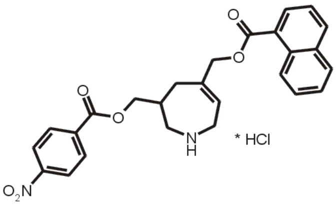
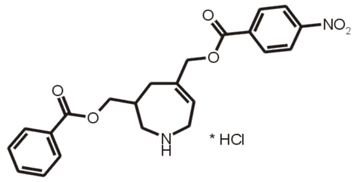
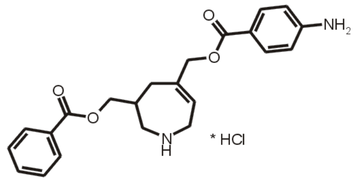
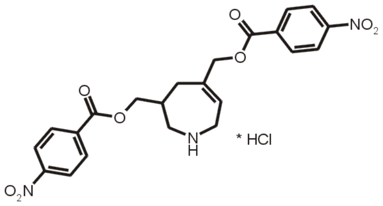
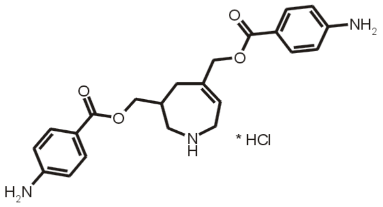
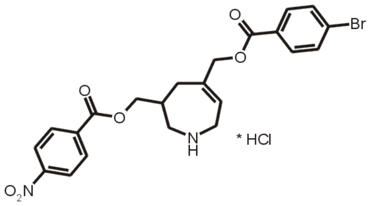
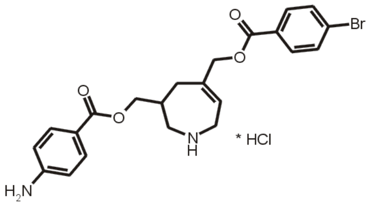
Compound	Structure	Enzyme K_i (µM)		
		Plm II	Plm IV	Cat D
1		4.7	7.2	260
2		303.3	211.0	>>320
3		28.8	38.6	348
4		1.0	1.5	262
5		2.4	15.9	>100
6		1.3	3.9	51
7		4.7	43.4	>100

Table 2. Experimental Inhibitory Activities of the Compounds Studied

Compound	Structure	Enzyme K_i (μM)		
		Plm II	Plm IV	Cat D
8	 NO2 O O O O N H * HCl	28.8	29.0	285
9	 NH2 O O O O N H * HCl	7.6	9.9	255
10	 NO2 O O O O N H * HCl O2N	10.9	22.1	321
11	 NH2 O O O O N H * HCl H2N	0.8	3.2	142
12	 Br O O O O N H * HCl O2N	49.0	68.8	178
13	 Br O O O O N H * HCl H2N	0.4	5.7	165

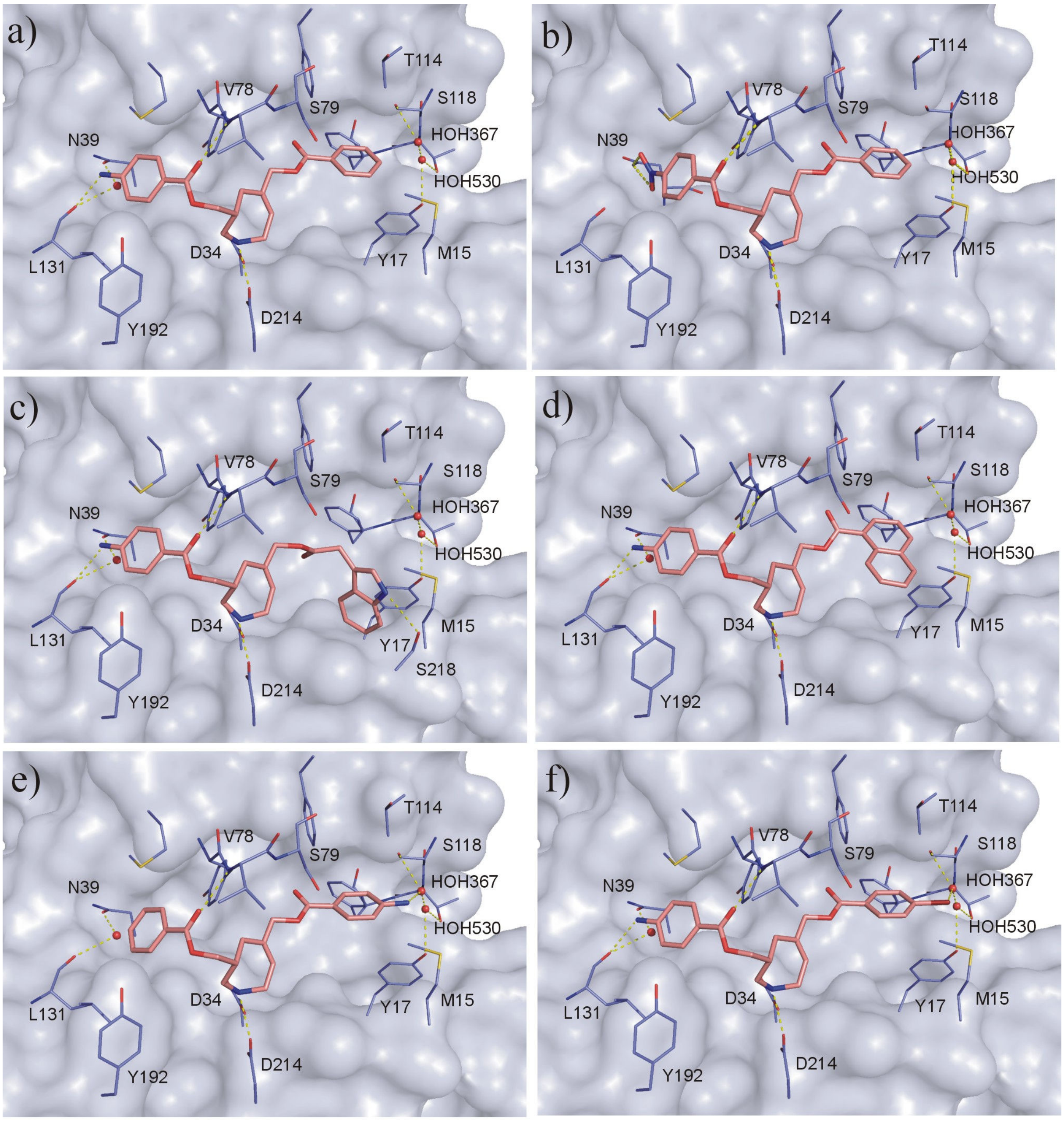

a)
T114
V78
S118
S79
HOH367
N39
HOH530
D34
M15
Y17
L131
D214
Y192
b)
T114
V78
S118
S79
HOH367
N39
HOH530
D34
M15
Y17
L131
D214
Y192
c)
T114
V78
S118
S79
HOH367
N39
HOH530
D34
M15
Y17
L131
S218
D214
Y192
d)
T114
V78
S118
S79
HOH367
N39
HOH530
D34
M15
Y17
L131
D214
Y192
e)
T114
V78
S118
S79
HOH367
N39
HOH530
D34
M15
Y17
L131
D214
Y192
f)
T114
V78
S118
S79
HOH367
N39
HOH530
D34
M15
Y17
L131
D214
Y192

Figure 8. Minimized docking solutions of **a) 1 b) 3 c) 6 d) 7 e) 9 f) 13** in 1LF2.

An extended set of P1 substituents with either multiple rotatable bonds or rather peptidic character received high scoring either for Plm II and Plm IV. To avoid these undesired solutions, we applied a filter and only allowed substituents without amide bonds and a limited number of rotatable bonds. Among the remaining most promising candidates, a 1*H*-indolyl-3-methyl substituent could be detected. It potentially forms a hydrogen bond to S218 (Figure 8c). Sterically more demanding hydrophobic groups such as a naphthyl moiety followed on subsequent ranks (Figure 8d). In consequence (Scheme 3), the *p*-amino derivatives **4** and **5** were synthesized along with their nitro-substituted precursors **6** and **7**. Compared to the benzyl derivative **1**, both, the 3-indolyl acetic acid substituted inhibitor **4** as well as the naphthyl derivate **5**, reveal improved binding towards Plm II. With respect to Plm IV, only the indolyl-type inhibitor **4** achieves better binding. The naphthyl derivative **5** drops slightly compared to the benzyl compound **1**. For Plm II, both nitro derivatives **6** and **7** experience improved binding compared to **3**. The same holds with respect to Plm IV binding of **6** whereas **7**, in agreement with the amino series, shows nearly unchanged affinity.

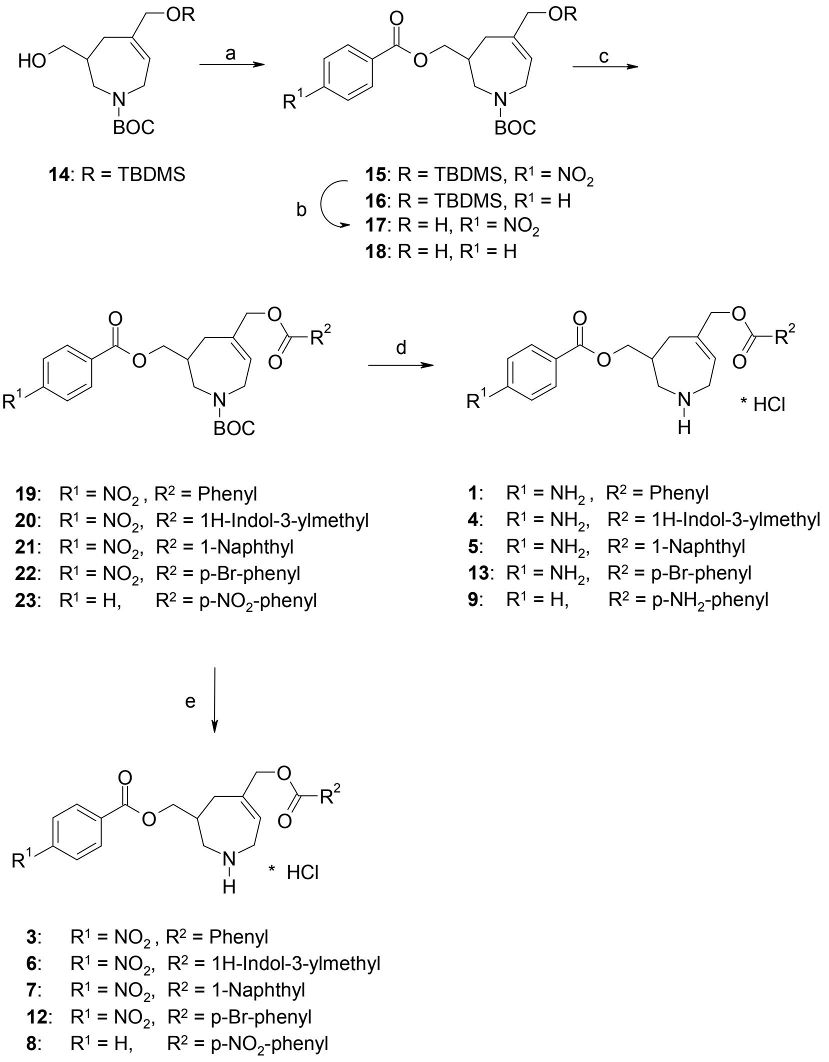
Based on these results we entered into a third design cycle. The modeled binding pose of **1** or **7** suggests introduction of a polar substituent either at *para*-position of the phenyl ring or 2-position of the naphthyl moiety. Such substituents could possibly interact with the neighboring water molecules present in the crystal structure with RS367. To test this hypothesis, we synthesized **8** and **9** featuring a *p*-nitro or *p*-amino moiety at this position (Figure 8e). Indeed, both show improved affinities compared to **2**, suggesting that an amino group is better suited than a nitro group. The indicated trend is further confirmed by **10** and **11**. Interestingly enough, also a polarizable bromine substituent at this position is beneficial for binding as indicated by the binding data of **12** and **13** (Figure 8f).

Overall, for Plm IV similar trends can be observed as for Plm II, strongly supporting our initial hypothesis that Plm IV is also capable of adopting a conformation with an open binding pocket as observed for Plm II (1LF2).

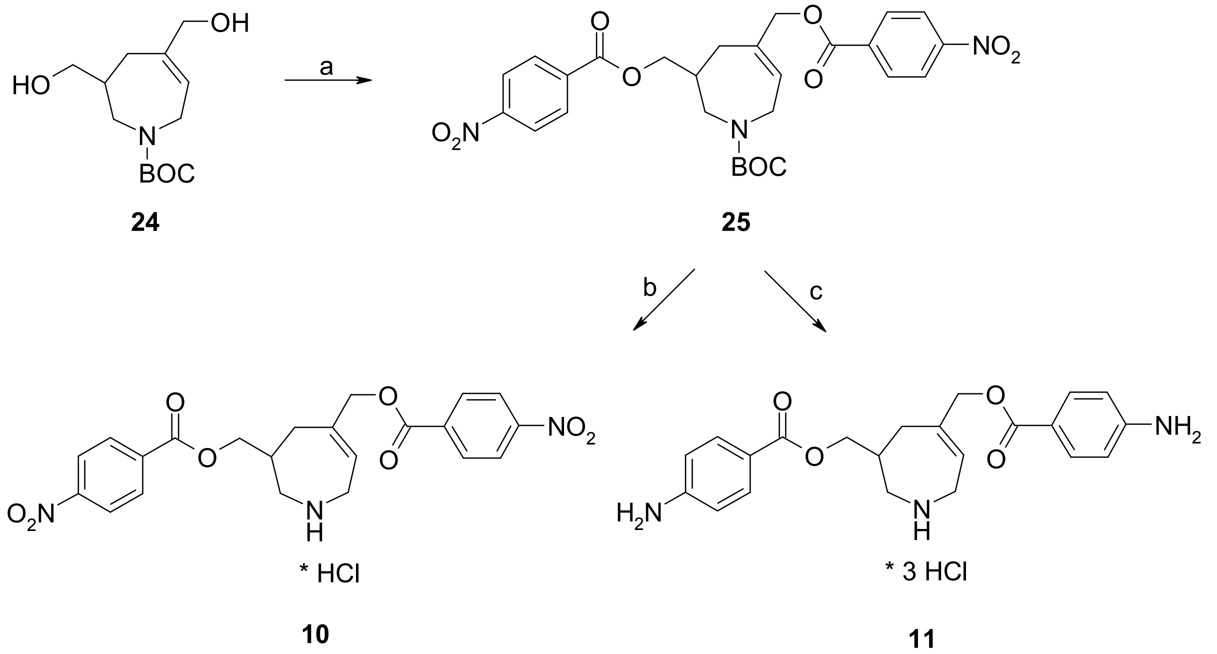
The affinities measured for Cathepsin D are all in a high micromolar range indicating satisfactory selectivity towards Plm II and IV.

4.4 Chemical Synthesis

N-BOC-2,3,4,7-tetrahydroazepine **14** bearing two anchoring groups, the unprotected hydroxymethyl function in position 3 as well as the TBDMS-protected one in position 5, is easily accessible via a straightforward convergent strategy in 7 steps in 21% overall yield [35]. Condensation of the primary hydroxyl group with either benzoyl chloride or *p*-NO_2-benzoic acid applying standard procedures easily renders the ester derivatives **15** and **16**. Selective cleavage of the TBDMS group furnishes the corresponding hydroxymethyl derivatives **17** and **18** respectively. These were subsequently condensed with a set of preliminary chosen aromatic carboxylic acids to give rise to the bis-ester derivatives **19-23**. In the final synthetic step leading to the first series of inhibitors **1**, **4, 5**, **9**, and **13**, the NO_2-groups of the substituents in **19-23** were reduced to the corresponding amino functionalities utilizing tin-(II)-chloride in ethyl acetate at elevated temperature. Concomitantly with this reaction step, the *N*-BOC protecting groups were also removed and by treatment with anhydrous HCl in diethyl ether the corresponding hydrochlorides **1, 4, 5**, **9**, and **13** were obtained. For a second series of inhibitors, the BOC-protecting group of nitro compounds **19-23** was directly cleaved under non-aqueous conditions affording inhibitors **3, 6-8**, and **12** as their hydrochloride salts in good overall yield. Starting from the BOC-protected bis-diol **24**, which can be obtained by deprotection of TBDMS-protected azepine **14**, the bis-*p*-nitrophenylester **25** is easily accessible in a one-step procedure. This precursor **25** provides direct and easy access to further inhibitors either by simple cleavage of the BOC protecting group giving rise to inhibitor **10** or by reduction of both nitro groups and subsequent BOC-deprotection rendering inhibitor **11** as its tri-hydrochloride.


a
c
b
d
e
* HCl
14: R = TBDMS
15: R = TBDMS, R¹ = NO2
16: R = TBDMS, R¹ = H
17: R = H, R¹ = NO2
18: R = H, R¹ = H
19: R¹ = NO2, R² = Phenyl
20: R¹ = NO2, R² = 1H-Indol-3-ylmethyl
21: R¹ = NO2, R² = 1-Naphthyl
22: R¹ = NO2, R² = p-Br-phenyl
23: R¹ = H, R² = p-NO2-phenyl
1: R¹ = NH2, R² = Phenyl
4: R¹ = NH2, R² = 1H-Indol-3-ylmethyl
5: R¹ = NH2, R² = 1-Naphthyl
13: R¹ = NH2, R² = p-Br-phenyl
9: R¹ = H, R² = p-NH2-phenyl
3: R¹ = NO2, R² = Phenyl
6: R¹ = NO2, R² = 1H-Indol-3-ylmethyl
7: R¹ = NO2, R² = 1-Naphthyl
12: R¹ = NO2, R² = p-Br-phenyl
8: R¹ = H, R² = p-NO2-phenyl

Scheme 3. Synthetic pathway towards inhibitors **1**, **3-9**, **12-13**.
Reagents and conditions: (a) for **15**: *p*-NO_2-benzoic acid, THF/DCM: 1/1, DIC, DMAP, 12 h, 71%; for **16**: Benzoyl chloride, DCM, TEA, DMAP, 18 h, 72%; (b) 1% HCl_{aq} in THF, rt; for **17**: 30 min, 77%, for **18**: 30 min, 90%; (c) for **19**: Benzoyl chloride, DCM, TEA, DMAP, 16 h, 87%; for **20**: 3-indole acetic acid, THF/DCM: 1/1, DIC, DMAP, 14 h, 84%; for **21**: 1-Naphthoic acid, THF/DCM: 1/1, DIC, DMAP, 12 h, 44%; for **22**: *p*-Br-benzoic acid, THF/DCM: 1/1, DIC, DMAP, 14 h, 58%; for **23**: *p*-NO_2-benzoic acid, THF/DCM: 1/1, DIC, DMAP, 18 h, 65%; (d) (i) $SnCl_2$ in EtOAc, 80 °C, 3h (ii) 2M HCl in diethyl ether, rt, 24 h; for **1**: 48%; for **4**: 52%; for **5**: 73%; for **13**: 47%; for **9**: 70%; (e) 2M HCl in diethyl ether, rt, 24 h; for **3**: 91%; for **6**: 85%; for **7**: 91%; for **12**: 94%; for **8**: 88%.



Scheme 4: Synthetic pathway towards inhibitors **10** and **11**
Reagents and conditions: (a) *p*-NO_2-benzoic acid, THF/DCM: 1/1, DIC, DMAP, 13 h, 69%; (b) 2M HCl in diethyl ether, rt, 24 h, 96%; (c) (i) $SnCl_2$ in EtOAc, 80 °C, 3h (ii) 2M HCl in diethyl ether, rt, 24 h; 59%.

4.5 Summary and Conclusion

In summary, we have developed novel non-peptidic Plm II and Plm IV inhibitors featuring a 2,3,4,7-tetrahydro-1*H*-azepine scaffold as core element. Equipped with suitable side-chains to address two of the enzyme's specificity pockets, these inhibitors show activity up to the nanomolar range. Derived from a combined subpocket search and a combinatorial docking approach, first inhibitors with affinities in the low micromolar range could be identified. They were subsequently synthesized following our previously established synthetic route, which is mainly based on commercially available, inexpensive starting materials. Structural modifications of the initial lead structure in a consecutive design cycle indeed led to inhibitors now exhibiting affinities in the submicromolar range. Overall, 13 inhibitors were synthesized, the experimentally determined K_i values of which are generally in good agreement with the design hypothesis. The found structure-activity relationships thus support our predicted binding mode for Plm II as well as Plm IV. Noteworthy, as compound **4** inhibits Plm II and Plm IV equally well and displays selectivity over the human cathepsin D, this inhibitor represents our new lead structure.

Future work will be directed towards addressing an additional substrate recognition pocket to further increase affinity against Plm II and Plm IV retaining or even increasing the selectivity against cathepsin D.

4.6 Experimental Section

4.6.1 Assays and K_i determinations

Plm II activity assays were performed in 96-well microtiter plates with a Tecan Spectra Fluor spectrometer at excitation wavelength 360 nm and emission wavelength 465 nm. The fluorogenic substrate Arg-Gln-Phe*Nphe-Ile-Thr was purchased from Bachem (the asterisk denotes the cleavage site and Nphe is p-nitrophenylalanine). A 1 mM substrate stock solution in DMSO was prepared. The stock solution was diluted in a ratio 1 to 50 with an acetic acid/acetate buffer (0.1 M; pH 4.5). 180 µl of this solution were pipetted in each well. The Plm II concentration, measured by UV/Vis, was 10 nM in an acetic acid/acetate buffer (0.1 M; pH 4.5). For the assay 18 µl of the protein were mixed with 2 µl inhibitor (dissolved in DMSO) and were allowed to fully equilibrate with enzyme for 5 min. DMSO (2.3 % final concentration) was used to guarantee complete dissolution of the inhibitors. For the reaction the 20 µl of the enzyme inhibitor solution were added to the 180 µl substrate solution. The final enzyme concentration was 1 nM, and the final substrate concentration 18 µM. The catalyzed substrate hydrolysis was recorded as an increase in fluorescence intensity over a period of 3 min, during which the intensity increased linearly with time. IC_{50} values were taken from plots of V_i/V_0 as a function of inhibitor concentration, for which V_i and V_0 are the initial rates of reaction in presence and absence of the inhibitor, respectively. To convert IC_{50} values to K_i values, the following equation was applied:

$$K_i = [IC_{50} - (E_t/2)][1 + (S/K_m)]^{-1}$$

where E_t is the total enzyme concentration (1 nM), K_m the Michaelis-Menten constant (63 µM), and S the substrate concentration (18 µM).

Plm IV activity assays were essentially performed as described for Plm II. The enzyme concentration was 10 nM, K_m 28 µM and S 18 µM.

Cathepsin D activity assays were essentially performed as described for Plm II. The enzyme concentration was 1 nM, K_m 16 µM and S 9 µM.

4.6.2 Homology model of Plm IV

As template for homology modeling, the complex structure of Plm II (1LF2) was used, assuming that Plm IV can also adopt a more opened binding pocket. The sequence of Plm IV was initially mapped onto that of Plm II using the align option in MOE [46]. The sequences aligned with 69.3 % identity (using the BLOSUM62 substitution matrix [47]).

The homology modeling of Plm IV was carried out in MOE version 2006.08 [46]. 10 models were generated and each was slightly minimized. The protein report of MOE was checked for outliers in the final model.

4.6.3 Molecular Dynamics Simulation

The MD simulation and all setup steps were performed with the Amber 8.0 suite of programs [48] using the Amber 1999 force field. The Plm structure with PDB code 1LEE was used as starting point. The ligand and all crystallographic water molecules were removed and hydrogen atoms were added with PROTONATE. After estimating protonation states of all histidines and of the two catalytic aspartates (Asp34 and Asp214) at pH 5 with Poisson-Boltzmann pKa calculations, histidines were set to the protonated form, the two catalytic aspartates to the deprotonated form. The protein was initially subjected to 200 steps of minimization using a generalized Born solvation model. Subsequently, the system was solvated in a box of ~10300 TIP3P water molecules [49], and sodium ions were added to ensure neutrality. After 200 steps of minimization of the solvated system, the MD simulation was started by heating the solvent to 300 K over a period of 20 ps and cooling to 100 K over 5 ps, keeping the protein fixed. Then the entire system was brought to 300 K over a period of 25 ps and the simulation was carried on for 10 ns under NPT conditions, using a time step of 2 fs and PME [50] for evaluating the electrostatic interactions. Energy data were saved every 20 fs, protein coordinates every 0.5 ps. CARNAL was used for further analysis of the trajectory and VMD 1.8.2 [51] for visualization.

4.6.4 Computational Docking

Docking studies were performed with the docking programm FlexX (version 2.0.3) [52]. The crystal structures of Plm II with PDB codes 1LF2, 1SME and 2BJU were used as starting geometries. The ligand and all crystallographic water molecules were removed. For the active site determination a radius of 13 Å was selected, always using complete amino acids. A fragment dataset with 2083 carboxylic acids was compiled from the Sigma-Aldrich catalogue, all with a molecular weight below 250 g/mol. The core and the R-group instances were preprocessed with Sybyl [53] as follows: correct atom and bond types as well as formal charges were assigned, hydrogens were added, 3D structures were generated and energy minimized. All fragments were marked with an attachment point X. The core structure was marked with an attachment point R1. Then the core and the fragments were linked and the connecting X- and R-atoms were removed from the atom list: A library of 2083 esters was generated [36]. Subsequently, the compounds were docked by using the *placeseq* option. During placement, the FlexX scoring function was applied for the final selection of the best-scored docking solutions.

4.6.5 Comparison of protein subpockets

The crystal structure of Plm II with PDB code 1LEE was used as starting geometry. All amino acids 6 Å around the ligand were used for the determination of the binding pocket. Then the complete pocket was further divided into subpockets. The S1 pocket (residues I14, M15, Y17, L32, D34, Y77, S79, F111, T114, S118, F120, I123, D214, S215, G216, S218) and S2' pocket (residues S37, A38, N39, M75, N76, Y77, V78, L131, S132, I133) were further analyzed. *Cavbase* [43], a database developed from the database RELIBASE [54], was used to compare the protein subpockets of Plm II with the subpockets of all stored binding pockets. All fragments which bind to similar subpockets were extracted.

4.7 References

1. WHO. *The World Health Report*, 2004.
2. Francis SE, Gluzman IY, Oksman A*, et al.* Molecular characterization and inhibition of a Plasmodium falciparum aspartic hemoglobinase. *Embo J* 1994;**13**(2):306-17.
3. Goldberg DE, Slater AF, Beavis R*, et al.* Hemoglobin degradation in the human malaria pathogen Plasmodium falciparum: a catabolic pathway initiated by a specific aspartic protease. *J Exp Med* 1991;**173**(4):961-9.
4. Gluzman IY, Francis SE, Oksman A*, et al.* Order and specificity of the Plasmodium falciparum hemoglobin degradation pathway. *J Clin Invest* 1994;**93**(4):1602-8.
5. Dame JB, Reddy GR, Yowell CA*, et al.* Sequence, expression and modeled structure of an aspartic proteinase from the human malaria parasite Plasmodium falciparum. *Mol Biochem Parasitol* 1994;**64**(2):177-90.
6. Humphreys MJ, Moon RP, Klinder A*, et al.* The aspartic proteinase from the rodent parasite Plasmodium berghei as a potential model for plasmepsins from the human malaria parasite, Plasmodium falciparum. *FEBS Lett* 1999;**463**(1-2):43-48.
7. Banerjee R, Liu J, Beatty W*, et al.* Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. *Proceedings of the National Academy of Sciences of the United States of America* 2002;**99**(2):990-5.
8. Salas F, Fichmann J, Lee GK*, et al.* Functional Expression of Falcipain, a Plasmodium-Falciparum Cysteine Proteinase, Supports Its Role as a Malarial Hemoglobinase. *Infect Immun* 1995;**63**(6):2120-2125.
9. Sijwali PS, Kato K, Seydel KB*, et al.* Plasmodium falciparum cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites. *Proc Natl Acad Sci USA* 2004;**101**(23):8721-8726.
10. Shenai BR, Sijwali PS, Singh A*, et al.* Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of Plasmodium falciparum. *J Biol Chem* 2000;**275**(37):29000-29010.
11. Sijwali PS, Rosenthal PJ. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by Plasmodium falciparum. *Proc Natl Acad Sci USA* 2004;**101**(13):4384-4389.
12. Sijwali PS, Shenai BR, Gut J*, et al.* Expression and characterization of the Plasmodium falciparum haemoglobinase falcipain-3. *Biochem J* 2001;**360**(Pt 2):481-9.
13. Eggleson K, Duffin KL, Goldberg DE. Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite Plasmodium falciparum. *Journal of Biological Chemistry* 1999;**274**(45):32411-32417.
14. Klemba M, Gluzman I, Goldberg DE. A Plasmodium falciparum dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation. *J Biol Chem* 2004;**279**(41):43000-7.
15. Coombs GH, Goldberg DE, Klemba M*, et al.* Aspartic proteases of Plasmodium falciparum and other parasitic protozoa as drug targets. *Trends Parasitol* 2001;**17**(11):532-537.
16. Moon RP, Tyas L, Certa U*, et al.* Expression and characterization of plasmepsin I from Plasmodium falciparum. *European Journal of Biochemistry* 1997;**244**(2):552-560.
17. Boss C, Richard-Bildstein S, Weller T*, et al.* Inhibitors of the Plasmodium falciparum parasite aspartic protease plasmepsin II as potential antimalarial agents. *Curr Med Chem* 2003;**10**(11):883-907.

18. Omara-Opyene AL, Moura PA, Sulsona CR*, et al.* Genetic Disruption of the Plasmodium falciparum Digestive Vacuole Plasmepsins Demonstrates Their Functional Redundancy. *J Biol Chem* 2004;**279**(52):54088-54096.
19. Liu J, Gluzman IY, Drew ME*, et al.* The Role of Plasmodium falciparum Food Vacuole Plasmepsins. *J Biol Chem* 2005;**280**(2):1432-1437.
20. Ersmark K, Samuelsson B, Hallberg A. Plasmepsins as potential targets for new antimalarial therapy. *Medicinal Research Reviews* 2006;**26**(5):626-666.
21. Oefner C, Binggeli A, Breu V*, et al.* Renin inhibition by substituted piperidines: a novel paradigm for the inhibition of monomeric aspartic proteinases? *Chemistry and Biology* 1999;**6**(3):127-131.
22. Specker E, Böttcher J, Lilie H*, et al.* An old target revisited: Two new privileged skeletons and an unexpected binding mode for HIV-protease inhibitors. *Angewandte Chemie-International Edition* 2005;**44**(20):3140-3144.
23. Carcache DA, Hortner SR, Bertogg A*, et al.* 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):1137-41.
24. Hof F, Schutz A, Fäh C*, et al.* Starving the malaria parasite: Inhibitors active against the aspartic proteases plasmepsins I, II, and IV. *Angewandte Chemie-International Edition* 2006;**45**(13):2138-2141.
25. Silva AM, Lee AY, Gulnik SV*, et al.* Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from Plasmodium falciparum. *Proc Natl Acad Sci USA* 1996;**93**(19):10034-10039.
26. Asojo OA, Afonina E, Gulnik SV*, et al.* Structures of Ser205 mutant plasmepsin II from Plasmodium falciparum at 1.8 A in complex with the inhibitors rs367 and rs370. *Acta Crystallogr D* 2002;**58**(Pt 12):2001-8.
27. Asojo OA, Gulnik SV, Afonina E*, et al.* Novel uncomplexed and complexed structures of plasmepsin II, an aspartic protease from Plasmodium falciparum. *J Mol Biol* 2003;**327**(1):173-81.
28. Prade L, Jones AF, Boss C*, et al.* X-ray Structure of Plasmepsin II Complexed with a Potent Achiral Inhibitor. *J Biol Chem* 2005;**280**(25):23837-23843.
29. Ersmark K, Feierberg I, Bjelic S*, et al.* Potent inhibitors of the Plasmodium falciparum enzymes plasmepsin I and II devoid of cathepsin D inhibitory activity. *J Med Chem* 2004;**47**(1):110-22.
30. Ersmark K, Feierberg I, Bjelic S*, et al.* C2-symmetric inhibitors of Plasmodium falciparum plasmepsin II: synthesis and theoretical predictions. *Bioorg Med Chem* 2003;**11**(17):3723-33.
31. Weik S, Luksch T, Evers A*, et al.* The potential of P1 site alterations in peptidomimetic protease inhibitors as suggested by virtual screening and explored by the use of C-C-coupling reagents. *Chemmedchem* 2006;**1**(4):445-457.
32. Ersmark K, Nervall M, Hamelink E*, et al.* Synthesis of malarial plasmepsin inhibitors and prediction of binding modes by molecular dynamics simulations. *J Med Chem* 2005;**48**(19):6090-6106.
33. Bursavich MG, Rich DH. Designing Non-Peptide Peptidomimetics in the 21st Century: Inhibitors Targeting Conformational Ensembles. *Journal of Medicinal Chemistry* 2002;**45**(3):541-558.
34. Carcache DA, Hoertner SR, Seiler P*, et al.* Development of a new class of inhibitors for the malarial aspartic protease plasmepsin II based on a central 7-azabicyclo[2.2.1]heptane scaffold. *Helvetica Chimica Acta* 2003;**86**(6):2173-2191.
35. Brass S, Chan NS, Gerlach C*, et al.* Synthesis of 2,3,4,7-tetrahydro-1H-azepines as privileged ligand scaffolds for the design of aspartic protease inhibitors via a ring-closing metathesis approach. *Journal of Organometallic Chemistry* 2006;**691**(24-25):5406-5422.

36. Rarey M, Lengauer T. A recursive algorithm for efficient combinatorial library docking. *Perspectives in Drug Discovery and Design* 2000;**20**(1):63-81.
37. Allen FH. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr B* 2002;**B58**:380-388.
38. Bruno IJ, Cole JC, Edgington PR*, et al.* New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Crystallographica Section B-Structural Science* 2002;**58**:389-397.
39. Pedrosa R, Andres C, Gutierrez-Loriente A*, et al.* Sequential diastereoselective addition of allylic and homoallylic grignard reagents to 2-acyl-perhydro-1,3-benzoxazines and ring-closing metathesis: an asymmetric route to azepin-3-ol and azocin-3-ol derivatives. *European Journal of Organic Chemistry* 2005(12):2449-2458.
40. Dewar MJS, Zoebisch EG, Healy EF*, et al.* The development and use of quantum mechanical molecular models. 76. AMI: a new general purpose quantum mechanical molecular model. *J Am Chem Soc* 1985;**107**:3902-3909.
41. Sadowski J, Gasteiger J, Klebe G. Comparison of Automatic 3-Dimensional Model Builders Using 639 X-Ray Structures. *J Chem Inf Comput Sci* 1994;**34**(4):1000-1008.
42. Gutierrez-de-Teran H, Nervall M, Ersmark K*, et al.* Inhibitor binding to the Plasmepsin IV aspartic protease from Plasmodium falciparum. *Biochemistry (Mosc)* 2006;**45**(35):10529-10541.
43. Schmitt S, Kuhn D, Klebe G. A new method to detect related function among proteins independent of sequence and fold homology. *J Mol Biol* 2002;**323**(2):387-406.
44. Kuhn D, Weskamp N, Schmitt S*, et al.* From the similarity analysis of protein cavities to the functional classification of protein families using Cavbase. *J Mol Biol* 2006;**359**(4):1023-1044.
45. Gerber PR, Muller K. Mab, a Generally Applicable Molecular-Force Field for Structure Modeling in Medicinal Chemistry. *J Comput Aided Mol Des* 1995;**9**(3):251-268.
46. Group CC. Molecular Operating Environment (MOE). 2006.08 ed. Montreal, Canada.
47. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* 1992;**89**(22):10915-9.
48. Case DA, Cheatham TE, Darden T*, et al.* The Amber biomolecular simulation programs. *Journal of Computational Chemistry* 2005;**26**(16):1668-1688.
49. Jorgensen WL, Chandrasekhar J, Madura JD*, et al.* Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* 1983;**79**(2):926-935.
50. Darden T, York D, Pederson L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* 1993;**98**(12):10089-10092.
51. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph* 1996;**14**(1):33-&.
52. Rarey M, Kramer B, Lengauer T*, et al.* A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 1996;**261**(3):470-489.
53. Tripos A. St. Louis, MO, U.S.A, 2004.
54. Hendlich M, Bergner A, Gunther J*, et al.* Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. *J Mol Biol* 2003;**326**(2):607-620.

5 The Potential of P1 site alterations in peptidomimetic protease inhibitors as suggested by virtual screening and explored by the use of C-C-Coupling reagents [1]

5.1 Abstract

A synthetic concept is presented that allows the construction of peptide isostere libraries through polymer-supported C-acylation reactions. A phosphorane linker reagent is used as a carbanion equivalent; by employing MSNT as a coupling reagent, the C-acylation can be conducted without racemization. Diastereoselective reduction was effected with l-selectride. The reagent linker allows the preparation of a norstatine library with full variation of the isosteric positions including the P1 side chain that addresses the protease S1 pocket. Therefore, the concept was employed to investigate the P1 site specificity of peptide isostere inhibitors systematically. The S1 pocket of several aspartic proteases including plasmepsin II and cathepsin D was modeled and docked with ~500 amino acid side chains. Inspired by this virtual screen, a P1 site mutation library was designed, synthesized, and screened against three aspartic proteases (plasmepsin II, HIV protease, and cathepsin D). The potency of norstatine inhibitors was found to depend strongly on the P1 substituent. Large, hydrophobic residues such as biphenyl, 4-bromophenyl, and 4-nitrophenyl enhanced the inhibitory activity (IC50) by up to 70-fold against plasmepsin II. In addition, P1 variation introduced significant selectivity, as up to 9-fold greater activity was found against plasmepsin II relative to human cathepsin D. The active P1 site residues did not fit into the crystal structure; however, molecular dynamics simulation suggested a possible alternative binding mode.

---

[1] published in ChemMedChem 2006; 1 (4):445-57, authored by S. Weik, T. Luksch, A. Evers, J. Böttcher, C. A. Sotriffer, A. Hasilik, H. G. Löffler, G. Klebe, J. Rademann
Contributions of T. Luksch concern the binding mode analysis and biological evaluation for Plm II and hCatD.

5.2 Introduction

Proteases constitute one of the few protein classes widely established as putative drug targets and are involved in pathological states ranging from viral and parasitic infections, through blood pressure, coagulation, and neurodegenerative diseases such as Alzheimer's, to osteoporosis and cancer. Therefore, protease inhibition by small organic compounds has been studied extensively, and protease inhibitors have found broad application as admitted drugs. A plethora of knowledge has been compiled on the design and synthesis of these molecules [1–3].
Most reversible protease inhibitors bind to the active site of the enzyme by mimicking the tetrahedral intermediate formed during the catalyzed amide bond cleavage. A noncleavable molecular motif addresses the interaction pattern exposed to the active site; in the considered case of aspartic proteases, peptide isosteres that contain a secondary alcohol or amine have often been used successfully as such 'transition-state mimics'. In these inhibitory structures, the mimicking core addresses the interaction pattern at the catalytic center, notably the interaction with the proteolytic machinery of the enzyme. With its surrogate of the P1 side chain of the natural substrate, it accommodates the S1 binding pocket. Beyond the P1 site, the mimicking core is flanked by extended recognition elements at both ends targeting the binding pockets in either the N-terminal (S2, S3, S4, …) or C-terminal direction (S1', S2', S3', …).

It has been known for a long time that the P1 site is of primordial importance to establish selectivity and affinity in substrate recognition by a particular protease; accordingly many proteases have been categorized with respect to their P1 site specificity. Thus, it can be assumed that also for the molecular skeleton of protease inhibitors, the nature of the P1 site is a determinant for selectivity and affinity. The work presented herein demonstrates a novel, synthesis-based approach to investigate the P1 site contribution in peptidomimetic inhibitors.

In principle, the contribution of individual structural elements to the overall binding affinity of an extended inhibitor structure can be investigated by the systematic variation of the respective elements. This approach assumes that the contributions of the individual elements to the total binding properties are additive. Over the last decade, the systematic variation of side chains within a series of protease inhibitors has been facilitated considerably by the introduction of combinatorial concepts both in solution and on solid phase [4–9]. However, upon closer inspection, the combinatorial screen has mainly been limited to the flanking regions of the mimicking core. Accordingly, most structure–activity relationship studies of protease inhibitors focus exclusively on the elaborate variation of the C- and N-terminal sites, that is, the P2, P3, … sites in the N-terminal direction and the P1', P2', … sites in the C-terminal direction of the native substrate. In contrast, P1 mutations remain largely unexplored and have been studied only in special cases [10–13]. As a result, elaborate and exhaustive P1 site mutations have not yet been performed to investigate their influence on the affinity or specificity of protease inhibitors. Therefore, we considered the development of novel synthetic strategies leading to efficient P1 mutations as highly attractive goal.

The main reason for limited P1 site variation in peptidomimetics has been the difficult access to diversely substituted isostere building blocks. Classically, peptide isosteres are constructed by multistep synthesis in solution. Often the isosteric core is assembled by a C-C-coupling step that requires harsh conditions as applied for strongly basic carbanion chemistry. For an efficient combinatorial variation of the mimicking core, however, it is necessary to construct an isostere through a C-C-coupling reaction on the polymer support along with the introduction of a side-chain functionality and with simultaneous control of the stereochemistry. For better accessibility, the isosteric cores should be constructed solely from commercially available chiral building blocks such as protected amino acids. All reaction conditions have to be smooth and fully compatible with standard peptide chemistry, thereby excluding the use of harsh reagents such as strong bases, strong acids, etc. To our knowledge, general combinatorial approaches to construct and vary mimicking cores including the P1 substituent purely based on commercially available chiral building blocks (amino acids) using mild reaction conditions have not been reported so far.

5.3 Results and Discussion

C-C-Coupling reactions are the key step in the synthesis of most peptide isosteres. Therefore, the implementation of such reactions in solid-phase synthesis protocols is a prerequisite for the flexible variation of peptide isosteres. α-Hydroxy-β-amino acids that are found in norstatines, for example, were selected as a model for the construction of peptide isosteres on solid phase. These isosteres have been reported as efficient inhibitors for several different classes of hydrolases, including aspartic proteases [4–6, 14] and metalloproteases [15, 16] depending on their stereochemistry. Norstatines can be derived directly or indirectly from α-amino aldehydes through the addition of various C1 nucleophiles. Cyanide anions [17, 18], isocyanide-based multicomponent reactions [19–22], and formaldehyde hydrazones [23, 24] have been reported as C nucleophiles in these umpolung reactions; additionally, vinyl Grignard and lithium organic compounds [25–29] have been used as masked C1 anion equivalents. In these reactions, the stereocontrol in positions 2 and 3 is problematic due to the facile racemization of the amino aldehyde building block and the limited stereospecificity of the nucleophilic attack.

As an alternative, C-C couplings that lead to isosteric building blocks can be furnished by a C acylation reaction. This strategy mimics the chemistry of the lipid and polyketide pathways in which acylations of thioesters are employed. The C-acylation strategy has a number of distinct advantages over the addition strategy. The preparation of peptide aldehydes is avoided and the chiral pool of amino acids can be used directly. Following the racemization-free acylation of a 1,1-dipole synthon representing a masked carbonyl group, α-keto derivatives are obtained which can be reduced stereospecifically. Thus, the stereochemistry of the norstatine products is controlled separately for both stereocenters.

In the group of Prof. Dr. Rademann the synthesis routine was established [30] in order to construct isostere libraries with variation of the P1 position while we designed a norstatine library by virtual screening.

5.3.1 Virtual screening of the plasmepsin II S1 site and design of a norstatine library

Plasmepsin II, an aspartic protease responsible for the degradation of human hemoglobin in malarial infections, was selected as the first inhibitor target addressed by a P1 mutation library of norstatines. To develop optimally suited P1 occupancy, a virtual screen of putative side-chain decorations for the S1 pocket in plasmepsin II was performed. Starting from the published enzyme–inhibitor complex of plasmepsin II and EH-58 (**11**) containing a hydroxyethylamine isostere (PDB code: 1LF3, Figure 1) [31], docking studies were carried out by applying the FlexX docking software with a set of potential norstatine inhibitors **10**, which were structurally derived from the published statine inhibitor **12** [32] (Figure 2). Similar positions and orientations of the P1 residue for the hydroxyethylamine, statine (=hydroxyethyl) and norstatine (=hydroxymethyl) isosteres were obtained. For the purpose of further ligand optimization, the position of the core skeleton was spatially fixed, and the P1 side chain was replaced systematically by 500 amino acid residues referenced in the Sigma–Aldrich and ACD databases. The resulting docking solutions were assessed by a scoring function. Some general conclusions could be derived from these screens: apart from a phenyl residue, which usually addresses S1 in the natural substrate hemoglobin, phenyl moieties substituted by polar residues such as hydroxy or methoxy groups are also predicted to be strong binders (Figure 1b). For an indolyl or substituted indolyl group, hydrogen bonds to a serine side chain in S1 were predicted. Finally, sterically more demanding hydrophobic residues such as naphthyl also received a high score (Figure 1c).

Thus, for synthesis, a small library of 17 norstatine inhibitors solely differing at P1 was assembled based on the indicated preferences found by virtual screening. Some P1 side chains were selected in accordance with the top-scoring virtual screening examples (**10d, e, l, n, o, p**). Further side chains were selected to assess and complement the predictions by spatially and/or electronically modulated hydrophobic side chains not directly suggested in the computational approach. Upon first glance, both configurations of the norstatine building blocks could bind plasmepsin II; accordingly, we decided to synthesize and test controlled mixtures of the corresponding diastereomers.

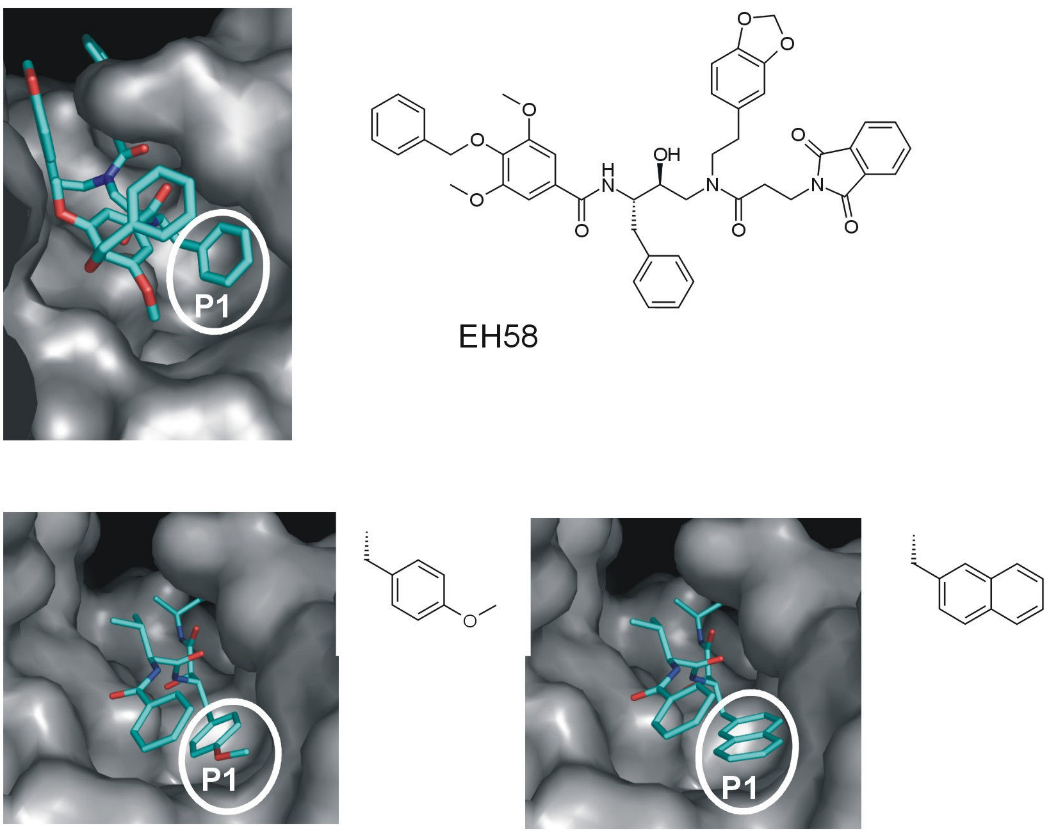
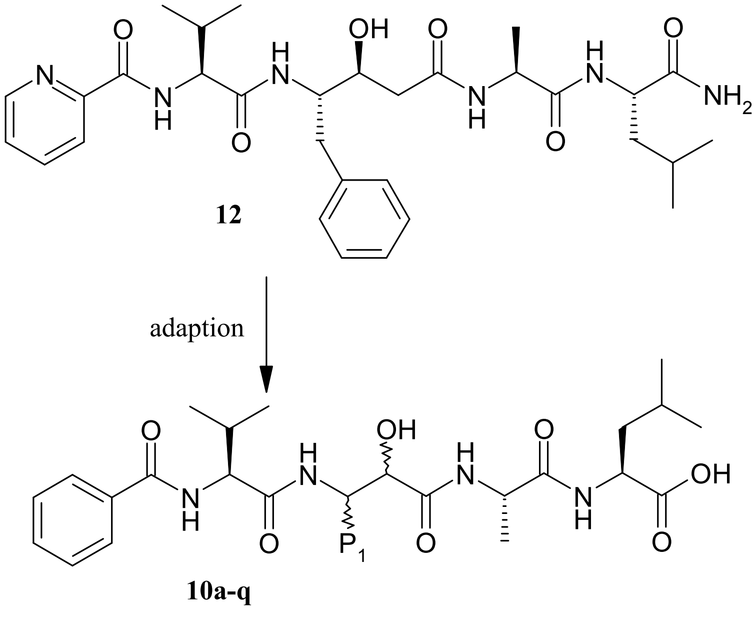


Figure 1. a) Structure and position of inhibitor EH-58 complexed with plasmepsin II [31]. The phenyl side chain (P1) was systematically exchanged in the designed norstatine analogues by other residues in a docking screen; particularly, **b)** a 4-methoxy phenyl group and **c)** a 2-naphtyl moiety were highly scored.


12

adaption

10a–q

norstatine scaffold

Figure 2. A virtual compound library **10a–q** was designed based on the docking of the norstatines structurally derived from the published statine inhibitor **12**.

5.3.2 Screening of inhibitors 10a-q against plasmepsin II, cathepsin D and HIV protease

Following the optimized reaction conditions, the synthesis of all designed inhibitors was pursued [30]. All synthesized compounds **10a–q** were assayed against different aspartic proteases. The main focus was directed toward the inhibition of plasmepsin II and to a comparison with the predicted favorable P1 side chains. Due to its structural similarity with plasmepsin II (35% sequence similarity) human cathepsin D, a lysosomal protease responsible for the degradation of proteins to small peptides, was also considered with particular respect to the selectivity potential of the prepared library. Finally, the inhibition of HIV protease was also determined. Both plasmepsin and HIV protease had been expressed, purified, and activated from the respective pro-enzyme forms as described elsewhere [33, 34]. Cathepsin D was used as the two chain mature enzyme purified from human placenta [35]. Compounds **10a–p** were tested as diastereomeric mixtures to allow an overall estimate of their potential.

The lead compound **10a**, which comprises a central phenyl norstatine, was attributed a relative activity of 1; as expected, it displays a lower inhibitory activity (IC_{50}) than that of the natural product pepstatin A. The relative activities of **10b–q** show the contribution of the different P1 side chains to the total inhibition (Table 2). In contrast to the predictions suggested by computer screening, hydrophilic phenyl substituents exhibiting additional hydrogen-bonding potential (in compounds **10o**, **10p**, and **10q**) proved to be less potent as plasmepsin II inhibitors. Similarly, a 1-naphthyl group (in **10d**) or indolyl residue (in **10l**) did not show potent inhibition. However, three norstatines (**10g**, **10j**, and **10k**) with large hydrophobic substituents at the para position of the phenyl ring depart significantly in affinity compared with all other members of the series. Clearly, with respect to all three proteases, the biphenyl residue (in **10g**) represents the most potent side chain. Whereas most of the less active compounds showed higher affinity toward cathepsin D than toward plasmepsin II, **10g**, **10j**, and **10k** displayed some selectivity in favor of plasmepsin II.

5.3.3 Molecular dynamics simulations indicate adaptibility of the S1 pocket

Virtual screening failed to predict ligands with extended P1 residues such as **10g**, **10j**, and **10k** as potent inhibitors for plasmepsin II. Considering the size of the S1 pocket as observed in the crystal structure, it appears to be too confined to accommodate residues of this size. To assess whether this pocket might adapt flexibly to the shape of a bound ligand, molecular dynamics (MD) simulations were performed. As a measure of structural stability, the root mean square deviation (rmsd) from the starting structure as a function of simulation time was calculated based on a superposition of all Cα atoms. All rmsd values remained below 1.90 Å, indicating that no significant structural instability was given.

By visual inspection of the simulation trajectory, methionine 15 (M15) was identified as rather flexible amino acid, whereas further residues of the S1 pocket show only minor changes. Quantitative analysis of molecular motion indicates that the $\chi1$ angle of M15 is not responsible for the large side-chain fluctuation; it only shows standard deviations from average (160°) of ± 24°. In contrast, the $\chi2$ angle exhibits a high standard deviation of 54° from the mean value of 184°. Besides the conformation found in the crystal structure, two further conformations, in particular, are highly populated and seem to be energetically favorable (Figure 3). Similar observations are made for the $\chi3$ angle: a high standard deviation of 53° from the average (182°), caused by two highly populated alternative conformations.

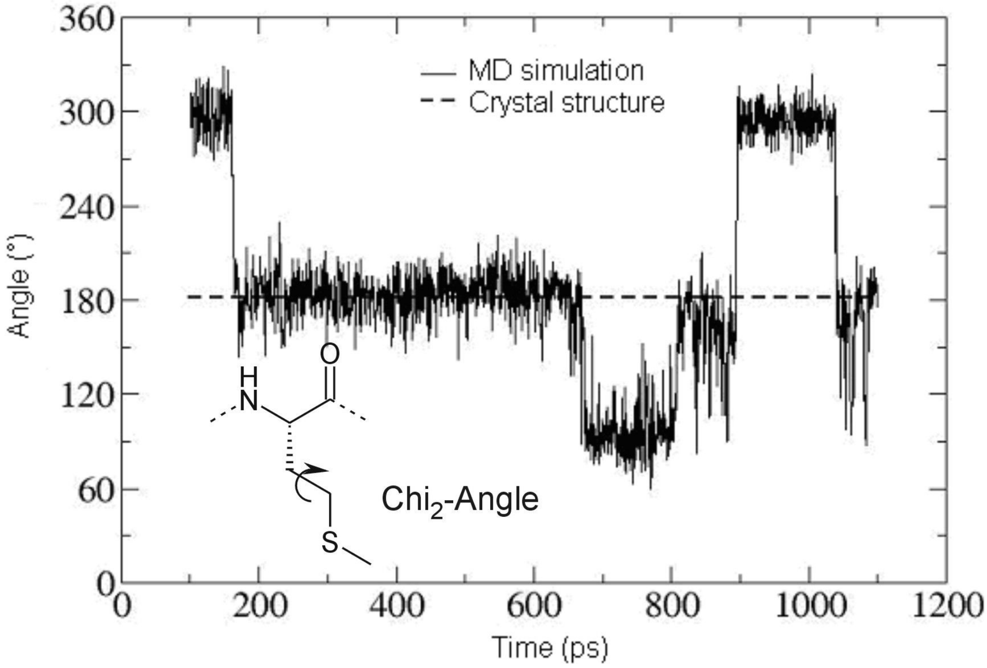


Figure 3. MD simulation (—) of the fluctuation of the $\chi2$ angle in methionine 15 (M15) of plasmepsin II, with the respective value from the crystal structure (---) shown as reference; for details, see the Experimental Section.

In our initial docking runs, the derivative with the largest P1 side chain was the naphthyl derivative **10d**. For the present analysis, the binding mode of this ligand was extended to a biphenyl substituent. If the conformation of plasmepsin II in the crystal structure (PDB code: 1LF3) is considered, the biphenyl group would be too large and would overlap spatially with the protein (Figure 4a). This explains why we failed to predict such derivatives as promising inhibitor candidates. However, as indicated by the MD simulations, M15 shows pronounced flexibility and can possibly adopt different conformations. In some of the frames collected along the trajectory, a significantly expanded S1 binding pocket is suggested. Taking these protein conformers as a reference, **10g**, **10j**, and **10k** can be accommodated into this site (Figure 4b), thus explaining the enhanced binding affinity of these derivatives with large hydrophobic P1 substituents. This hypothesis is strongly supported by a recently published crystal structure that shows the suggested, conformationally expanded S1 binding pocket in plasmepsin II [36].

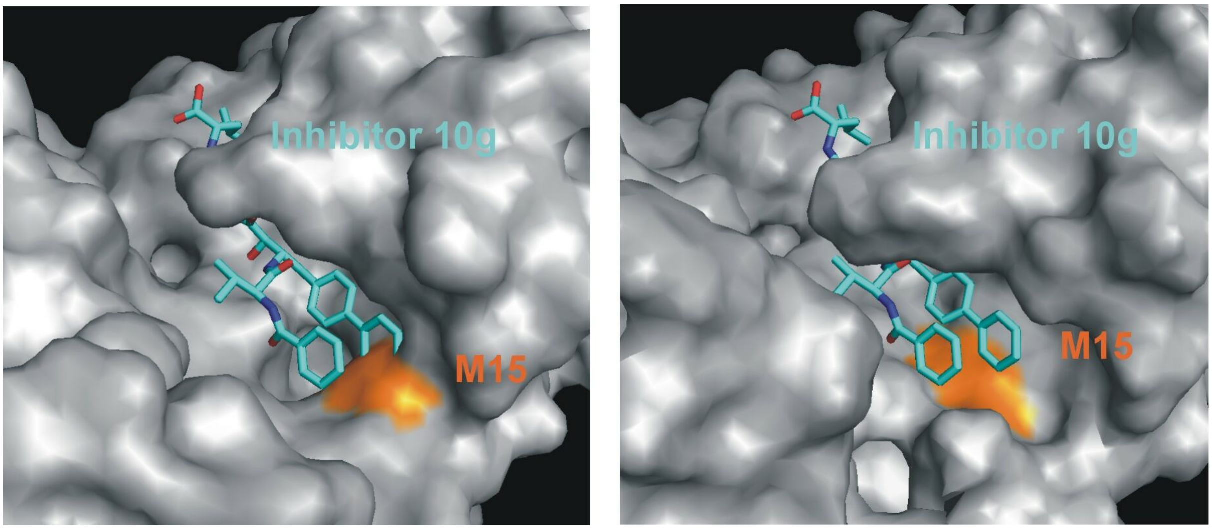
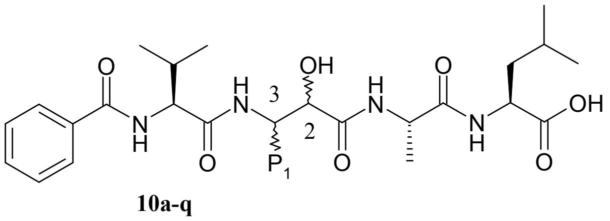
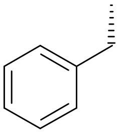
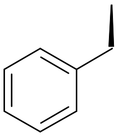
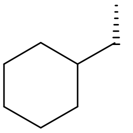
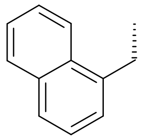
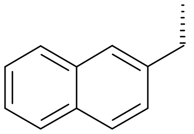
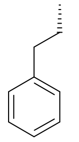
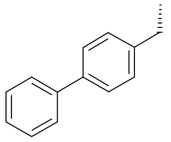
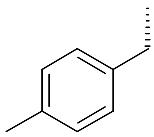
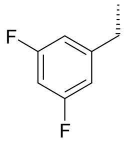
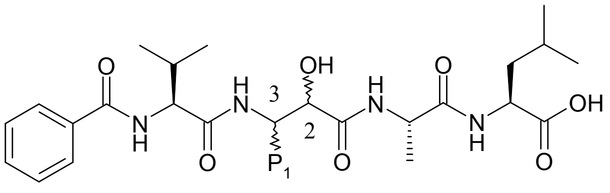


Figure 4. a) Based on the binding mode used in the initial docking study (see Figure 1), the P1 substituent was extended to a biphenyl group as found for the most potent inhibitor **10g**. The protein conformer derived from the crystal structure does not provide sufficient space to accommodate this sterically demanding substituent. **b)** A modified protein conformer with the M15 side chain (orange surface) populating an alternative conformation as indicated by MD simulation (Figure 3) can favorably host the extended biphenyl ligand.

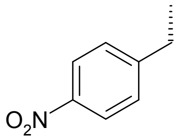
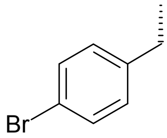
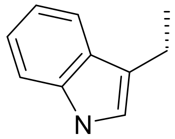
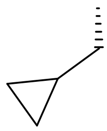
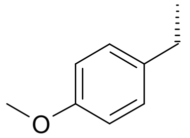
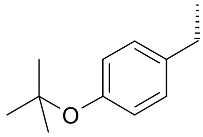
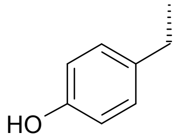
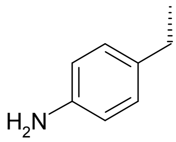
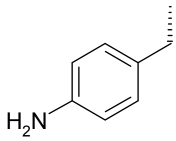


10a-q

		Plm II		Cat D		HIV protease	
Compound	Structure	IC$_{50}$ µM	rel. Inhib..	IC$_{50}$ µM	rel. Inhib	IC$_{50}$ µM	rel. Inhib
10a		346	1	163	1	271	1
10b		>350	<1	228	<1	>300	<1
10c		>350	<1	25	6.5	285	<1
10d		300	1.2	34	4.8	ND	<1
10e		110	3.1	32	5.1	100	2.7
10f		>350	<1	93	1.8	>300	<1
10g		5	69	7	23	14	19
10h		255	1.4	88	1.9	ND	<1
10i		>350	<1	43	3.8	100	2.7



10a-q

Compound	Structure	Plm II		Cat D		HIV protease	
		IC$_{50}$ µM	rel. Inhib..	IC$_{50}$ µM	rel. Inhib	IC$_{50}$ µM	rel. Inhib
10j		25	14	92	1.8	ND	<1
10k		11	32	95	1.7	15	18
10l		>350	<1	180	<1	>300	<1
10m		>350	<1	>350	<1	128	2.1
10n		120	2.9	66	2.5	90	3
10o		>350	<1	185	<1	322	<1
10p		>350	<1	>350	<1	583	<1
10qa (2S, 3S)		>350	<1	164	1	ND	<1
10qb (2R, 3S)		>350	<1	93	1.8	ND	<1

5.4 Conclusions

Virtual screening suggested that P1 site variation, a phenomenon that has been hardly investigated in greater detail until now, might in fact play a prominent role in increasing the affinity and specificity of peptidomimetic inhibitors. FlexX-mediated docking of 500 amino acid side chains was conducted in the active site of the aspartic protease plasmepsin II. Best binding was predicted for inhibitors containing large hydrophobic side chains such as naphthyl groups or large aromatic side chains with additional H-bond donor or acceptor sites. Inspired by the predictions of the virtual screen, a norstatine library was designed to address the S1 site with natural and unnatural amino acid side chains.

The synthesis of the norstatine library was carried out. The norstatines were subjected to inhibition assays against three prominent aspartic proteases: the validated malaria target plasmepsin II, HIV protease, and the homologous human protease cathepsin D. The potency of our norstatine inhibitors was found to depend strongly on the P1 substituent. In the case of plasmepsin II, a large, hydrophobic residue such as biphenyl enhanced the inhibitory activity (IC_{50}) by up to 70-fold. A similar effect was observed for 4-bromophenyl (30-fold) and 4-nitrophenyl (14-fold). In contrast, hosting additional hydrogen-bond acceptor or donor sites within the S1 pocket did not result in increased inhibitory activity. The enhancing effect was restricted to 4-substituted aromatic residues; both naphthyl and 3,5-disubstituted phenyl groups did not experience significant increase in the inhibitory activity, even though they received high scores in the virtual screen. In addition, P1 variation introduced significant selectivity enhancement: up to 9-fold discrimination advantage was observed for plasmepsin II relative to human cathepsin D. Interestingly, the extended P1 side chain could not be accommodated in the S1 pocket as observed in a crystal structure.

However, MD simulations suggest a conformational opening of this pocket which results in a significant increase in the active volume.

In summary, our study indicates the significance of P1 site substitution in peptidomimetic inhibitors against aspartic proteases. This effect has not been recognized so far and might be of considerable value for increasing the affinity and selectivity of protease inhibitors. After completion of our work, two independent groups were able to confirm our results impressively.

The research groups of Hallberg and Samuelsson reported a 50-fold inhibitory enhancement toward plasmepsin I of a statine (hydroxyethyl) transition-state isostere exhibiting a biphenyl moiety in the opened S1 pocket [37]. Furthermore, the research group at Actelion Pharmaceuticals communicated a crystal structure exhibiting a biphenyl moiety in the S1 pocket [36].

5.5 Experimental Section

5.5.1 Determination of inhibition constants toward plasmepsin II

Plasmepsin II activity assays were performed in 96-well microtiter plates with a Tecan Spectra Fluor spectrometer at excitation wavelength 360 nm and emission wavelength 465 nm. Hemoglobin labeled with the fluorogen 7-amino-4-methylcoumarin-3-acetic acid (AMCA) was used as substrate. For IC_{50} determinations, inhibitors were allowed to fully equilibrate with enzyme at 37°C for 5 min. The final substrate concentration was 2.6 µM. The plasmepsin II concentration, measured by UV/Vis, was 260 nM in an acetate/formate buffer (0.06 M, pH 4.65). DMSO (2.5%) was used to guarantee complete dissolution of the inhibitors. IC_{50} values were taken from plots of V_i/V_0 as a function of inhibitor concentration, for which V_i and V_0 are the initial rates of reaction in presence and absence of the inhibitor, respectively. Substrate hydrolysis was recorded as an increase in fluorescence intensity over a period of 3 min, during which the intensity increased linearly with time.

5.5.2 Determination of inhibition constants toward cathepsin D

Cathepsin D activity assays were essentially performed as described for plasmepsin II. The pH value of the buffer, however, was changed to 3.65 and the cathepsin D concentration was 380 nM. Mature two-chain-form human cathepsin D was used. It was purified from extracts of human placental lysosomes by pepstatin affinity and cation-exchange chromatography as described previously [35].

5.5.3 Virtual screening to select optimal occupancy of the S1 pocket of plasmepsin II

Docking studies were performed with the docking program FlexX (version 1.13) [38]. As starting geometry, the crystal structure of plasmepsin II with bound EH-58 (PDB code: 1LF3) was used. Based on the coordinates of EH-58, the core structure of the norstatine inhibitor was constructed with the P1 site unoccupied. A dataset of 500 unnatural amino acids was compiled from either the Sigma–Aldrich catalogue or the ACD. Subsequently, these residues were docked by using the *mapref* option in FlexX into the complex structure of plasmepsin II. The corresponding peptide backbone portion was considered as base fragment in *mapref*. During placement, the FlexX scoring function was applied for the final selection of the best-scored docking solutions; in addition, DrugScore was consulted [39].

5.5.4 Molecular dynamics simulation

The MD simulation and all setup steps were performed with the Amber 8.0 suite of programs [40], using the Amber 1999 force field. The plasmepsin structure with PDB code 1LEE was used as the starting point. The ligand and all crystallographic water molecules were removed, and hydrogen atoms were added with PROTONATE. After estimating protonation states of all histidines and of the two catalytic aspartates (Asp34 and Asp214) at pH 5 with Poisson–Boltzmann pKa calculations, histidine groups were set to the protonated form, and the two catalytic aspartates were set to the deprotonated form. The protein was initially subjected to 200 steps of minimization with a generalized Born solvation model. Subsequently, the system was solvated in a box of ~10300 TIP3P water molecules [41], and sodium ions were added to ensure neutrality. After 200 steps of minimization of the solvated system, the MD simulation was started by heating the solvent to 300 K over a period of 20 ps and cooling to 100 K over 5 ps, keeping the protein fixed. Then the entire system was brought to 300 K over a period of 25 ps, and the simulation was continued for 1 ns under NPT conditions using a time step of 2 fs and PME [42] for evaluating the electrostatic interactions. Energy data were saved every 20 fs, and protein coordinates were saved every 0.5 ps. CARNAL was used for further analysis of the trajectory and VMD 1.8.2 was used for visualization [53].

5.6 References

1. Leung D, Abbenante G, Fairlie DP. Protease inhibitors: current status and future prospects. Journal of Medicinal Chemistry 2000;43(3):305-41.
2. Babine RE, Bender SL. Molecular Recognition of Proteinminus signLigand Complexes: Applications to Drug Design. Chem Rev 1997;97(5):1359-1472.
3. Cooper JB. Aspartic proteases in Disease: A structural Perspective. Current Drug Targets 2002;3(2):155-173.
4. Wallace OB, Smith DW, Deshpande MS, et al. Inhibitors of Abeta production: solid-phase synthesis and SAR of alpha-hydroxycarbonyl derivatives. Bioorg Med Chem Lett 2003;13(6):1203-6.
5. Shuto D, Kasai S, Kimura T, et al. KMI-008, a novel beta-secretase inhibitor containing a hydroxymethylcarbonyl isostere as a transition-state mimic: design and synthesis of substrate-based octapeptides. Bioorg Med Chem Lett 2003;13(24):4273-6.
6. Mimoto T, Hattori N, Takaku H, et al. Structure-activity relationship of orally potent tripeptide-based HIV protease inhibitors containing hydroxymethylcarbonyl isostere. Chem Pharm Bull (Tokyo) 2000;48(9):1310-26.
7. Haque TS, Skillman AG, Lee CE, et al. Potent, low-molecular-weight non-peptide inhibitors of malarial aspartyl protease plasmepsin II. Journal of Medicinal Chemistry 1999;42(8):1428-40.
8. Lee A, Huang L, Ellman AE. General solid-phase method for the preparation of mechanism-based cysteine protease inhibitors. Journal of the American Chemical Society 1999;121:9907-9914.
9. Noteberg D, Hamelink E, Hulten J, et al. Design and synthesis of plasmepsin I and plasmepsin II inhibitors with activity in Plasmodium falciparum-infected cultured human erythrocytes. J Med Chem 2003;46(5):734-46.
10. Lee CE, Kick EK, Ellman JA. General solid-phase synthesis approach to prepare mechanism-based aspartyl protease inhibitor libraries. Identification of potent cathepsin D inhibitors. Journal of the American Chemical Society 1998;120:9735-9747.
11. Takashiro E, Hayakawa I, Nitta T, et al. Structure-activity relationship of HIV-1 protease inhibitors containing alpha-hydroxy-beta-amino acids. Detailed study of P1 site. Bioorg Med Chem 1999;7(9):2063-72.
12. Nishizawa R, Saino T, Takita T, et al. Synthesis and structure-activity relationships of bestatin analogues, inhibitors of aminopeptidase B. J Med Chem 1977;20(4):510-5.
13. Iizuka K, Kamijo T, Harada H, et al. Synthesis and structure-activity relationships of human renin inhibitors designed from angiotensinogen transition state. Chem Pharm Bull (Tokyo) 1990;38(9):2487-93.
14. Iizuka K, Kamijo T, Harada H, et al. Orally potent human renin inhibitors derived from angiotensinogen transition state: design, synthesis, and mode of interaction. J Med Chem 1990;33(10):2707-14.
15. Thunnissen MM, Nordlund P, Haeggstrom JZ. Crystal structure of human leukotriene A(4) hydrolase, a bifunctional enzyme in inflammation. Nat Struct Biol 2001;8(2):131-5.
16. Yuan W, Munoz B, Wong CH, et al. Development of selective tight-binding inhibitors of leukotriene A4 hydrolase. J Med Chem 1993;36(2):211-20.

17. Peet NP, Burkhart JP, Angelastro MR, et al. Synthesis of peptidyl fluoromethyl ketones and peptidyl alpha-keto esters as inhibitors of porcine pancreatic elastase, human neutrophil elastase, and rat and human neutrophil cathepsin G. J Med Chem 1990;33(1):394-407.
18. H. Harada, A. Tsubaki, T. Kamijo, K. Iizuka, Y. Kiso, Chem. Pharm. Bull. 1989, 37, 2570 –2572.
19. L. Banfi, G. Guanti, R. Riva, A. Basso, E. Calcagno, Tetrahedron Lett. 2002, 43, 4067 –4069.
20. Banfi L, Basso A, Guanti G, et al. Passerini reaction--amine deprotection--acyl migration (PADAM): A convenient strategy for the solid-phase preparation of peptidomimetic compounds. Mol Divers 2003;6(3-4):227-35.
21. A. Basso, L. Banfi, R. Riva, P. Piaggio, G. Guanti, Tetrahedron Lett. 2003, 44, 2367 –2370.
22. Semple JE, Owens TD, Nguyen K, et al. New synthetic technology for efficient construction of alpha-hydroxy-beta-amino amides via the Passerini reaction. Org Lett 2000;2(18):2769-72.
23. R. FernUndez, J. M. Lassaletta, Synlett 2000, 1228 – 1240.
24. D. Enders, M. Bolkenius, J. VUzquez, J. M. Lassaletta, R. FernUndez, J. Prakt. Chem. 1998, 340, 281–285.
25. S. Kourtal, J. Paris, Lett. Pept. Sci. 1996, 3, 73– 78.
26. G. Veeresha, A. Datta, Tetrahedron Lett. 1997, 38, 5223 – 5224.
27. J. P. Burkhart, N. P. Peet, P. Bey, Tetrahedron Lett. 1990, 31, 1385– 1388.
28. S. Hormuth, H. U. Reissig, D. Dorsch, Liebigs Ann. Chem. 1994, 121–127.
29. J. Deng, Y. Hamada, T. Shioiri, S. Matsunaga, N. Fusetani, Angew. Chem. 1994, 106, 1811 – 1813; Angew. Chem. Int. Ed. Engl. 1994, 33, 1729 – 1731.
30. Weik S, Luksch T, Evers A, et al. The potential of P1 site alterations in peptidomimetic protease inhibitors as suggested by virtual screening and explored by the use of C-C-coupling reagents. Chemmedchem 2006;1(4):445-457.
31. Asojo OA, Gulnik SV, Afonina E, et al. Novel uncomplexed and complexed structures of plasmepsin II, an aspartic protease from Plasmodium falciparum. Journal of Molecular Biology 2003;327(1):173-81.
32. Silva AM, Lee AY, Gulnik SV, et al. Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from Plasmodium falciparum. Proc Natl Acad Sci U S A 1996;93(19):10034-9.
33. Taylor A, Brown DP, Kadam S, et al. 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):205-10.
34. Hill J, Tyas L, Phylip LH, et al. High level expression and characterization of Plasmepsin II, an aspartic proteinase from Plasmodium falciparum. FEBS Letters 1994;352(2):155-8.
35. Hasilik A, Neufeld EF. Biosynthesis of lysosomal enzymes in fibroblasts. Synthesis as precursors of higher molecular weight. J Biol Chem 1980;255(10):4937-45.
36. Prade L, Jones AF, Boss C, et al. X-ray Structure of Plasmepsin II Complexed with a Potent Achiral Inhibitor. J. Biol. Chem. 2005;280(25):23837-23843.
37. Johansson PO, Lindberg J, Blackman MJ, et al. Design and synthesis of potent inhibitors of plasmepsin I and II: X-ray crystal structure of inhibitor in complex with plasmepsin II. J Med Chem 2005;48(13):4400-9.
38. Rarey M, Kramer B, Lengauer T, et al. A fast flexible docking method using an incremental construction algorithm. Journal of Molecular Biology 1996;261(3):470-489.

39. Gohlke H, Hendlich M, Klebe G. Knowledge-based scoring function to predict protein-ligand interactions. J Mol Biol 2000;295(2):337-56.

40. Case DA, Cheatham TE, Darden T, et al. The Amber biomolecular simulation programs. Journal of Computational Chemistry 2005;26(16):1668-1688.

41. Jorgensen WL, Chandrasekhar J, Madura JD, et al. Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics 1983;79(2):926-935.

42. Darden T, York D, Pederson L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. The Journal of Chemical Physics 1993;98(12):10089-10092.

43. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. Journal of Molecular Graphics 1996;14(1):33-&.

6 Pyrrolidine derivatives as Plasmepsin Inhibitors: Binding mode analysis assisted by Molecular Dynamics

6.1 Abstract

Plasmepsin (Plm) II (EC-number: 3.4.23.39) and IV (EC-number: 3.4.23.B14) are aspartic proteases present in the food vacuole of the malaria parasite *Plasmodium falciparum* which is involved in host hemoglobin degradation. A series of pyrrolidine derivatives, originally synthesized as HIV-1 protease inhibitors, was tested for Plm activity. Inhibitors in the nanomolar range were discovered for Plm II and IV. In order to explain structure-activity relationships detailed studies were carried out in finding putative binding modes. For pyrrolidine-diol-diester-derivatives and the substituted diamino pyrrolidine-diamide inhibitor resonable binding modes could be generated by using the Plm II conformer 1LF2. Modeling studies indicated, that the flap lid of available plasmepsin crystal structures is not opened enough to accommodate the pyrrolidinedimethylene diamines. Molecular dynamics were performed to analyze in more detail protein flexibility, leading to a ligand binding mode hypothesis for the pyrrolidinedimethylene diamines and to general implications for Plm II inhibitor design.

6.2 Introduction

Recently, a novel scaffold, a pyrrolidinedimethylene diamine, was designed and synthesized as general aspartic protease lead structure [1]. First pyrrolidinedimethylene diamine derivatives were decorated with side chains transferred from other ligands known as potent inhibitors of this family [2]. K_i values in the low μM range could be achieved for HIV-1 protease and Cathepsin D. The crystal structure could subsequently be determined with HIV-1 protease (pdb-code: 1XL2). In a second design cycle ten further compounds were synthesized (publication in preparation), altogether resulting in two series with 12 compounds.

In literature multiple examples are recorded with inhibitors designed for HIV-1 protease showing simultaneously powerful inhibition of Plm II [3-5] although the Plms are aspartic proteases with a pepsin-like fold, whereas HIV-1 protease belongs to the retroviral proteases. However, both proteins share similarities, that could be responsible for alike inhibition profiles towards the two protein families: Both proteins exhibit the catalytic dyad and the peptide recognition motif in common [1]. Furthermore, both aspartic proteases feature a flap (two in HIV-1 protease) that is involved in ligand binding. The flaps close upon substrate accommodation, thereby excluding solvent from the inner catalytic cavity of the enzyme [6].

Since the pyrrolidine derivatives were designed to address these family-wide features, inhibition of the Plms seems likely.

6.3 Results and Discussion

6.3.1 Enzyme Inhibition Data

12 compounds were tested for activity against Plm II, Plm IV and human aspartic protease Cathepsin D (Cat D). These compounds can be divided into three classes of molecules with different core structures (**1**, **2**, **3**) (Table 1). **1.1** - **1.9** belong to the first series of molecules, all consisting of a 3*S*,4*S*-pyrrolidine-diol-diester scaffold (**1**) that is symmetrically addressed with substituents differing in size and polarity.

A second molecule class is represented by **2**, which is composed by a 3*S*,4*S*-diamino pyrrolidine-diamide scaffold (**2**) and decorated with four benzyl moieties. Compounds **3.1** and **3.2** are racemic mixtures of pyrrolidinedimethylene diamines (**3**), both decorated with four side chains thought to interact with the subpockets of diverse aspartic proteases (Table 1).

3.1 and **3.2** achieve inhibition in the submicromolar range for the Plms. K_i values in the same range are measured for Cat D [2].

Well-diffracting crystals for HIV-1 protease in complex with one pyrrolidinedimethylene diamine were obtained [2]. These crystals were grown in the presence of a racemic inhibitor mixture [2]. Since the 3*R*,4*R*-pyrrolidinedimethylene diamine was found in the crystal structure (1XL2), it could be hypothesized that the *R*,*R*-enantiomer reveals stronger binding in comparison to the *S*,*S*-enantiomer. Based on these results, a second compound series was synthesized to inhibit HIV protease (**1.1** - **1.9**), following an enantioselective synthetic route (publication in preparation). Thereby, the scaffold was altered by reducing the linker length in order to optimize key interactions to the flap residues of HIV-1 protease. Furthermore, the number of decorating side chains was limited to two.

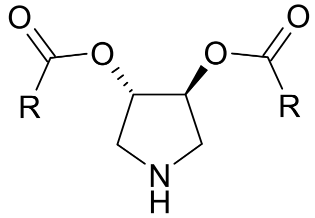
Compounds **1.1** - **1.9**, originally synthesized for HIV-1 protease were also tested against Plm II, IV and Cathepsin D.

While the pyrrolidinedimethylene diamines (**3**) decorated with four monocyclic aromatic ring systems exhibit K_i values in the submicromolar range, for the pyrrolidine-diol-diester derivatives equipped with two monocyclic aromatic rings, affinity data were only obtained in the single- and double-digit μM range, e.g. **1.3** showed a K_i value of 11 μM for Plm II, 7 μM for Plm IV and 53 μM for Cat D.

Within the series of the derivatives bearing **1** as core structural element, the compounds with 1-isobytyl-, 1-neopentyl-, and cyclohexyl-moieties were the weakest binders to the Plms with K_i values in the triple digit μM range (**1.1**, **1.2** and **1.5**). For Cat D no inhibition up to the 1 mM level was detectable. A gain in affinity was achieved by replacing the alkyl chains with mono cyclic aromatic ring systems. **1.3** is with a K_i of 11 μM towards Plm II about four fold more potent than **1.2** (K_i = 47 μM). Similar results were obtained for Plm IV (~ six fold increase in affinity). The removement of one methylene unit (**1.4**) in comparison to **1.3** has almost no effect on the binding affinity to Plm II, whereas a seven-fold decrease was detected for Plm IV. Furthermore, **1.6**, equipped with a *p*-methyl group, was not able to enhance affinity significantly. First the introduction of an acceptor group in ortho position of the benzyl moiety leads to a drop in affinity in the single-digit μM range (**1.7**). The pyrrolidine-diol-diester decorated with two 1-naphthyl moieties further increased the potency (**1.8**). The K_i value of **1.8** to Plm II is micromolar (K_i = 1μM), for Plm IV and Cat D slightly better K_i values in submicromolar range have been achieved (K_i = 0.8 μM and 0.5 μM respectively).

Table1. Experimental Inhibitory Activities of the Compounds Studied

1



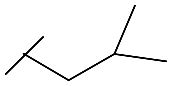
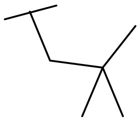
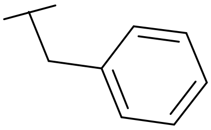
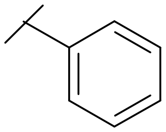
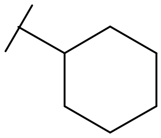
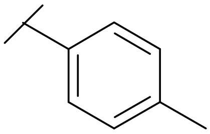
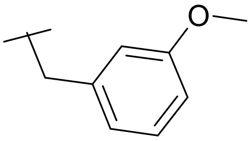
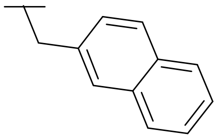
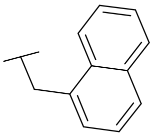
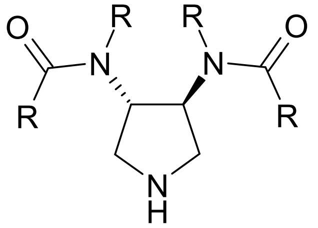
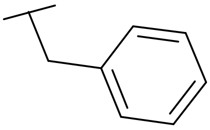
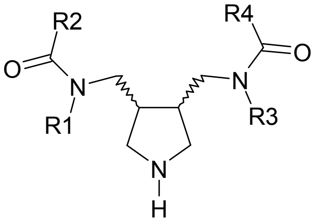
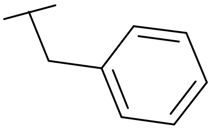
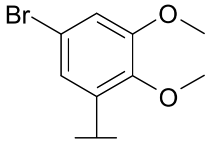
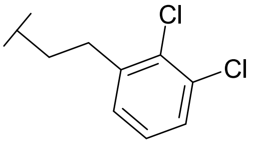
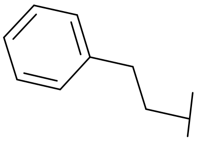
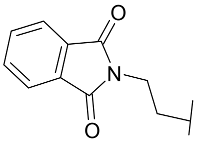
Compound	Structure	Enzyme K_i (µM)		
		Plm II	PlmIV	Cat D
1.1		51.2	168.3	>1000
1.2		47.2	41.2	>1000
1.3		11.4	7.4	53.4
1.4		19.0	53.8	>1000
1.5		163.1	133.9	>1000
1.6		15.3	10.5	11.0
1.7		2.6	5.4	19.9
1.8		1.0	0.8	0.5
1.9		0.1	0.089	0.374

Table1. Experimental Inhibitory Activities of the Compounds Studied

<table>
<tr><td>2</td><td colspan="4"></td></tr>
<tr><td rowspan="2">Compound</td><td rowspan="2">Structure</td><td colspan="3">Enzyme K_i (µM)</td></tr>
<tr><td>Plm II</td><td>PlmIV</td><td>Cat D</td></tr>
<tr><td>2</td><td></td><td>0.43</td><td>1.5</td><td>11.7</td></tr>
<tr><td>3</td><td colspan="4"></td></tr>
<tr><td colspan="2">R1
</td><td colspan="2">R2
</td><td>R3
</td></tr>
<tr><td rowspan="2">Compound</td><td rowspan="2">Structure</td><td colspan="3">Enzyme K_i (µM)</td></tr>
<tr><td>Plm II</td><td>Plm IV</td><td>Cat D</td></tr>
<tr><td>3.1</td><td>R4
</td><td>0.63</td><td>0.35</td><td>0.67</td></tr>
<tr><td>3.2</td><td>R4
</td><td>0.50</td><td>0.18</td><td>0.37</td></tr>
</table>

The most potent inhibitor in this series is **1.9** with a K_i value of 100 nM for Plm II and 89 nM for Plm IV. It shows a 1600-fold stronger binding to Plm II in comparison to **1.5**, the weakest binding inhibitor of all measured compounds.

Replacing the ester functionalities by amides and following a subsequent decoration with four benzyl moieties, an inhibitor with a K_i value of 430 nM for Plm II and 1.5 µM for Plm IV was obtained (**2**).

6.3.2 Computational binding mode generation

In order to hypothesize reasonable binding modes, previously determined Plm II crystal structures were consulted for their suitability to accommodate the core scaffolds studied in this contribution.

Currently 14 Plm II inhibitor complex structures of *P. falciparum* are available, which can be subdivided into three groups exhibiting different ligand binding modes [7].

The pepstatin A and related ligands are tightly embraced by the protein (pdb-codes: 1SME, 1ME6, 1XDH, 1XE5, 1XE6, 1W6I, 1W6H, 1M43). The hydroxyl group of the statine moiety makes hydrogen bonds to the catalytic dyad (Figure 1a). Further key interactions are formed by two inhibitor amide oxygens to the backbone nitrogen of V78 and to the sidechain of S79. Both amino acids belong to a *β*-hairpin structure, known as “flap”, closing up with virtually perpendicular orientation to the binding cleft. The flap is a flexible structural unit, present in all aspartic proteases, which interacts with substrates and inhibitors [8, 9].

The second group of ligands, found in co-crystal structures (pdb-codes 1LEE, 1LF2 and 1LF3) comprises a hydroxypropylamine scaffold and forms a similar pattern of conserved hydrogen bonds to the catalytic aspartates and the flap residues V78 and S79 [10] as detected in the group of the pepstain A like ligands (Figure 1b). In these complexes, the binding cavities remain more open compared to the pepstatin A complexes [11].

The third group of ligands, present in the co-crystal structures 2BJU, 2IGY, 2IGX induces significant differences in the protein conformation [12]. The *N*-substituted piperidine scaffold interacts with the catalytic aspartates via an interstitial water molecule. Moreover, no interactions to the flap residues are observed. Instead, major changes in the flap region occur and a new tunnel-shaped hydrophobic cavity opens up.

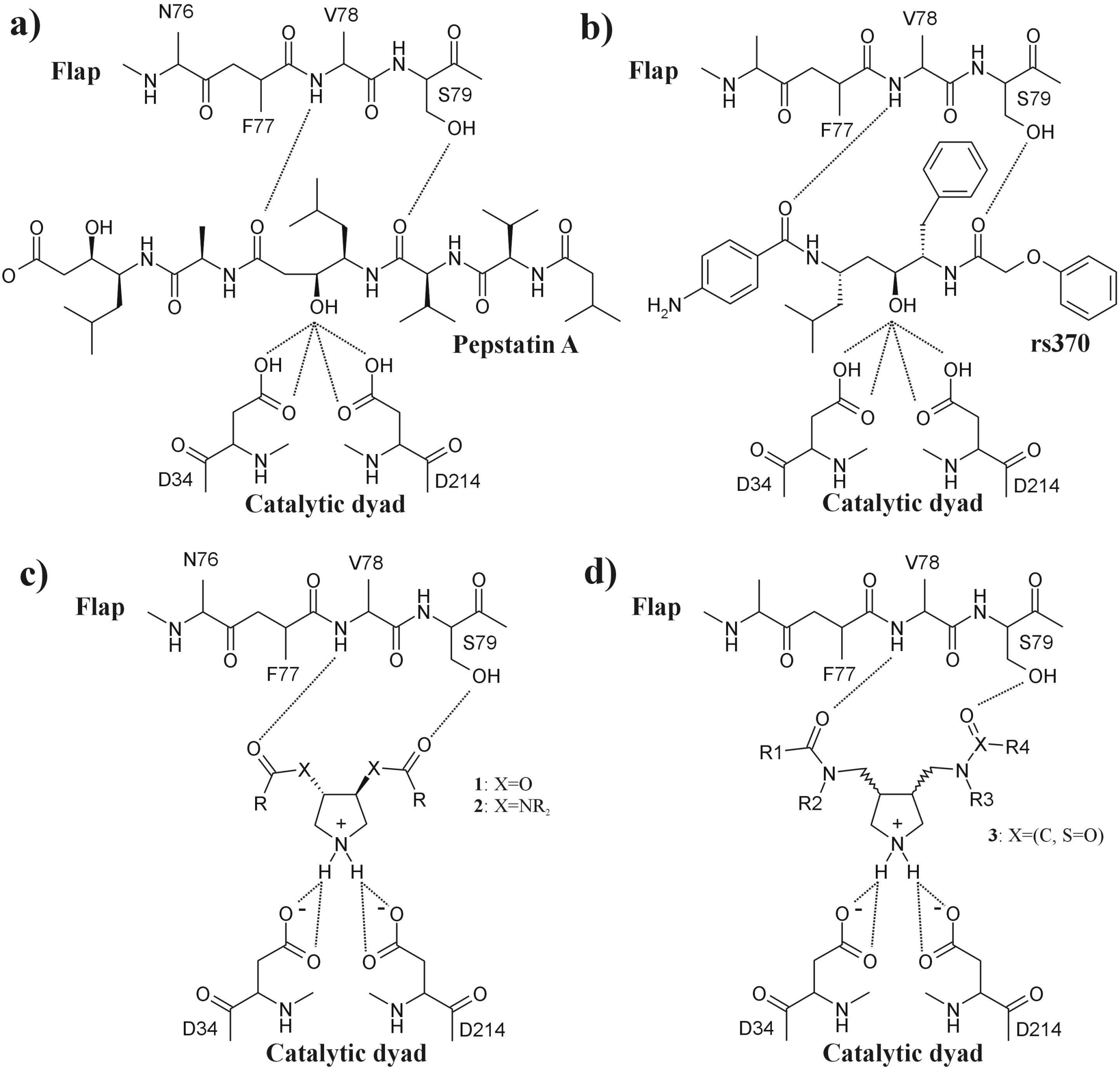


Figure1. Hydrogen bonds of **a)** pepstatin A and **b)** rs370 in the active site of Plm II. **c)** Putative hydrogen bonds of the 3*S*,4*S*-pyrrolidine-diol-diester scaffold **1** and 3*S*,4*S*-diamino pyrrolidine-diamide **2** to Plm II **d)** Putative hydrogen bonds of the pyrrolidinedimethylene diamine unit **3** to Plm II.

In the design of the derivatives with **1**, **2** and **3** as core structures, it was anticipated to directly address both, the aspartates of the catalytic dyad and the flap residues (Figure 1c, 1d). Plm II in the open conformation and in the conformation found in the complex structure with pepstatin A seemed therefore an interesting starting point for our theoretical considerations to generate appropriate binding modes.

For HIV-1 protease a co-crystal structure with **3** as scaffold was solved (pdb-code: 1XL2). The central pyrrolidine moiety is found at the pivotal position between both catalytic aspartate residues [2]. The ring adopts an envelope conformation and both ring substituents are found in axial position. Experimental pK_a values for unsubstituted pyrrolidines suggest the central nitrogen of the ligand to be protonated [13]. Furthermore, pK_a calculations of a pyrrolidine derivative in complex with HIV-1 protease suggest that both catalytic aspartates of the HIV-1 protease should be deprotonated upon ligand binding [14].

Taking these findings as a prerequesite for our model generation, we tried to manually place the ligand core structures of **1** - **3** into different conformers of Plm II either found in the complex with pepstatin A (1SME), or into the more open binding pocket (pdb-code: 1LF2), respectively. The ligand functionalities assigned to interact with the flap (Figure 1c, 1d) were rotated towards the backbone nitrogen of V78 and to the sidechain of S79.

For our three ligand series, no hydrogen bonds could be formed to the flap residues as found in the orientation observed in the close Plm II conformer with bound pepstatin A (1SME).

For **1** and **2** a manual placement into the binding pocket of the structure 1LF2 could be achieved. In order to optimize these initial ligand geometries, an energy minimization was performed. During this minimization with the Moloc MAB force field [15], the protein was kept rigid, whereas the inhibitor was assigned to full flexibility. For all minimized complexes, the pyrrolidine remains between the two catalytic aspartates making short H-bonds to both carboxylate groups. Furthermore, H-bonds are formed to the flap residues V78 and S79 (Figure 2). The side-chains of all minimized inhibitors point towards the S2 and S2' pocket.

As described above, within the series **1.1** - **1.9** affinity increases from 1-isobutyl < benzyl < *m*-methoxybenzyl < 1-naphtyl < 2-naphtyl (Table 1). The modeled binding modes support these findings (Figure 2). S2 and S2' are spacious pockets that are not fully occupied by small side chains such as a 1-isobutyl substituent (Figure 2a). The larger benzyl moiety can form better hydrophobic interactions (Figure 2b). Attaching meta methoxy groups to the benzyl moieties allows to establish hydrogen bonds to Y192 and S218 (Figure 2c). A further gain in affinity is observed for two 1-naphthyl substituents (Figure 2d). The inhibitor with two 2-naphthyl groups shows a ten-fold improved affinity in comparison to the 1-naphthyl substituted diester, leading to a K_i value of 100 nM for Plm II. This can be explained to the more favorable burial of the ligand side chains in the S2 and S2' pockets (Figure 2e).

A similar binding mode can be proposed for **2**, bearing the 3*S*,4*S*-diamino pyrrolidine-diamide scaffold: Both aspartates as well as the flap residues V78 and S79 are addressed by **2** (Figure 2f). The four side chains of this inhibitor are placed in the four subpockets (S1, S2, S1' and S2') making hydrophobic interactions.

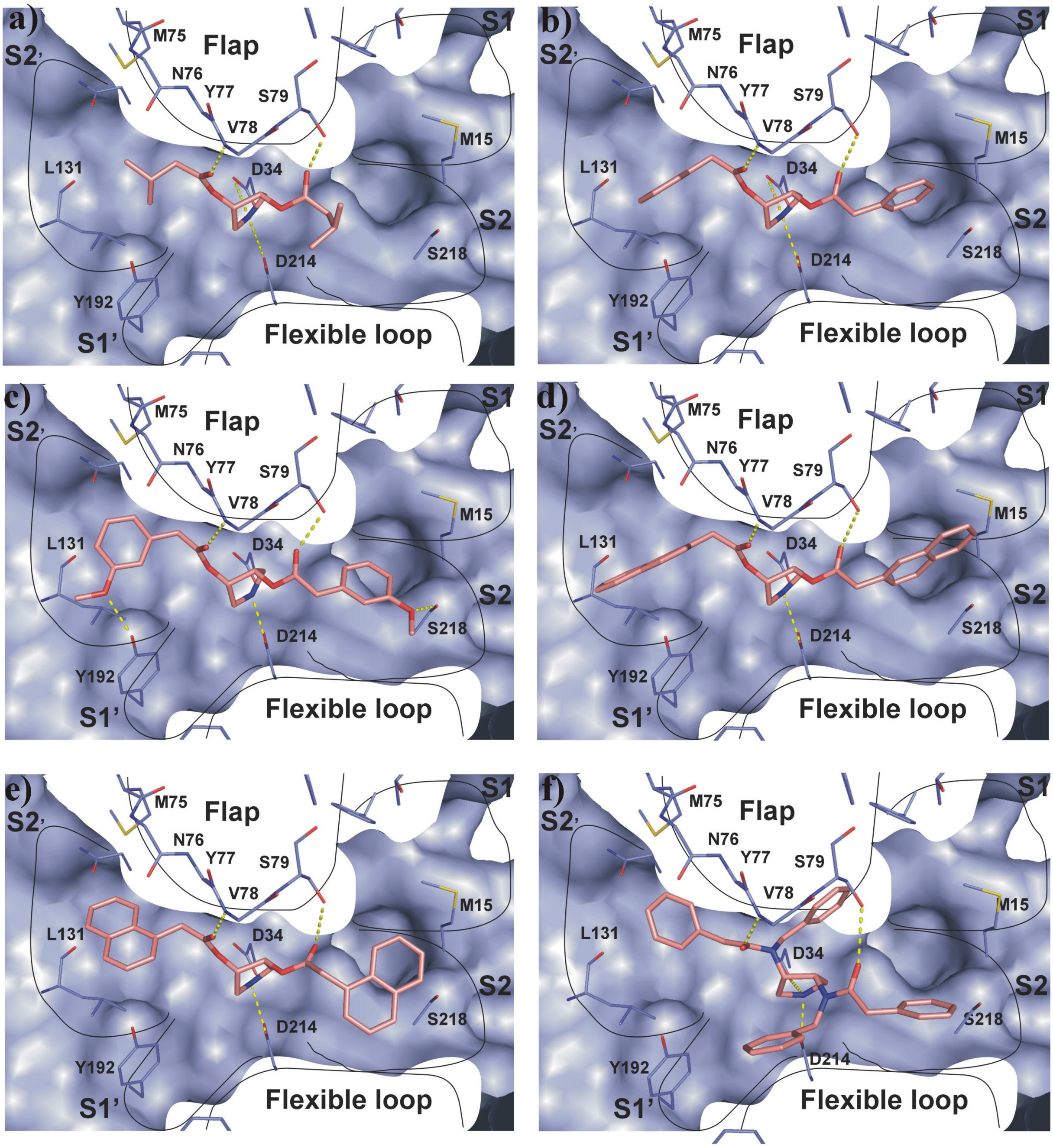


Figure 2. Modeled binding modes of **a) 1.1 b) 1.3 c) 1.7 d) 1.8 e) 1.9** and **f) 2** in the binding pocket of Plm II (1LF2).

The compounds of ligand series **1** and **2** were synthesized enantiopure, exhibiting *S*,*S*-configuration at both stereogenic centers. The derivatives of series **3** have been prepared and tested as racemic mixtures. Considering the crystal structure with HIV protease, only the 3*R*,4*R*-enantiomer showed binding. Subsequent separation and individual testing of the enantiomers confirmed higher potency of the *R*,*R*-enantiomer.

In analogy to the results observed for HIV protease also for Plm II better binding of the *R*,*R*-enantiomer of **3** is anticipated. Docking of these stereoisomers of **3.1** - **3.2** to the Plm II conformer observed in the 1LF2 crystal structure have been attempted. However, no reasonable binding mode could be generated that accomplish all interactions to the flap and catalytic dyad. Obviously, the flap lid would have to open further in order to allow H-bond formation simultaneously with the two aspartates and residues V78 and S79. Nevertheless, as potent binding is experimentally observed, it is in question whether Plm II can adopt to host **3.1** to **3.2** in an extended binding pocket. In order to obtain some insights into the dynamic properties of the enzyme molecular dynamics simulations have been performed.

6.3.3 Molecular Dynamics Simulations

The available structural studies underline that Plm II is a highly flexible protein, e.g. two molecules are found in the asymmetric unit of the Plm II pepstatin A co-crystal structure (1SME) that do not display the same conformation. This indicates intrinsic interdomain flexibility [9]. Furthermore, comparison of uncomplexed Plm II (1LF4) with Plm II in complex with EH58 (1LF3) reveals movements of residues in the binding pocket that give space to accommodate bulky groups of the inhibitor or to form hydrogen bonds [11].

Up to now only three distinct binding modes have been discovered for the different inhibitor classes. However, structural studies indicate that Plm II allows for significant conformational flexibility and could possibly adopt in a way to accommodate derivatives **3.1** - **3.2** exhibiting the extended pyrrolidine core structure.

To further sample conformational space of the protein an MD simulation over a period of several nanoseconds was performed. In order to explore the adaptive properties of the protein. Recently, we presented already MD simulations of uncomplexed Plm II. The obtained results helped to rationalize unexpected binding in two examples [7, 16]. Encouraged by these results, we extended simulation time from 1 ns to 10 ns and closely monitored further conformational transitions along the trajectory.

In the remainder we will focus exclusively on the binding-site region as being the area of highest interest in the context of structure-based drug design, particularly with respect to the flexibility issue. Overall, 32 residues were selected as part of the binding pocket. Mobility in this region is quantitatively analyzed by a mutual comparison of individual frames along the trajectory in terms of a 2D-RMSD plot. For this purpose, 500 snapshots were extracted at regular 20 ps intervals along the entire 10 ns trajectory. It is immediately apparent that three main conformational substates can be distinguished (Figure 3). The first is sampled for 4.3 ns at the beginning of the trajectory named "conformation$_{all}$ 1" (C$_{all}$1), while the second is called "conformation$_{all}$ 2" (C$_{all}$2). The conformational family C$_{all}$2 is then maintained for about 4.4 ns, before a transition to C$_{all}$3 occurs. Conformations within the same substate show RMSD

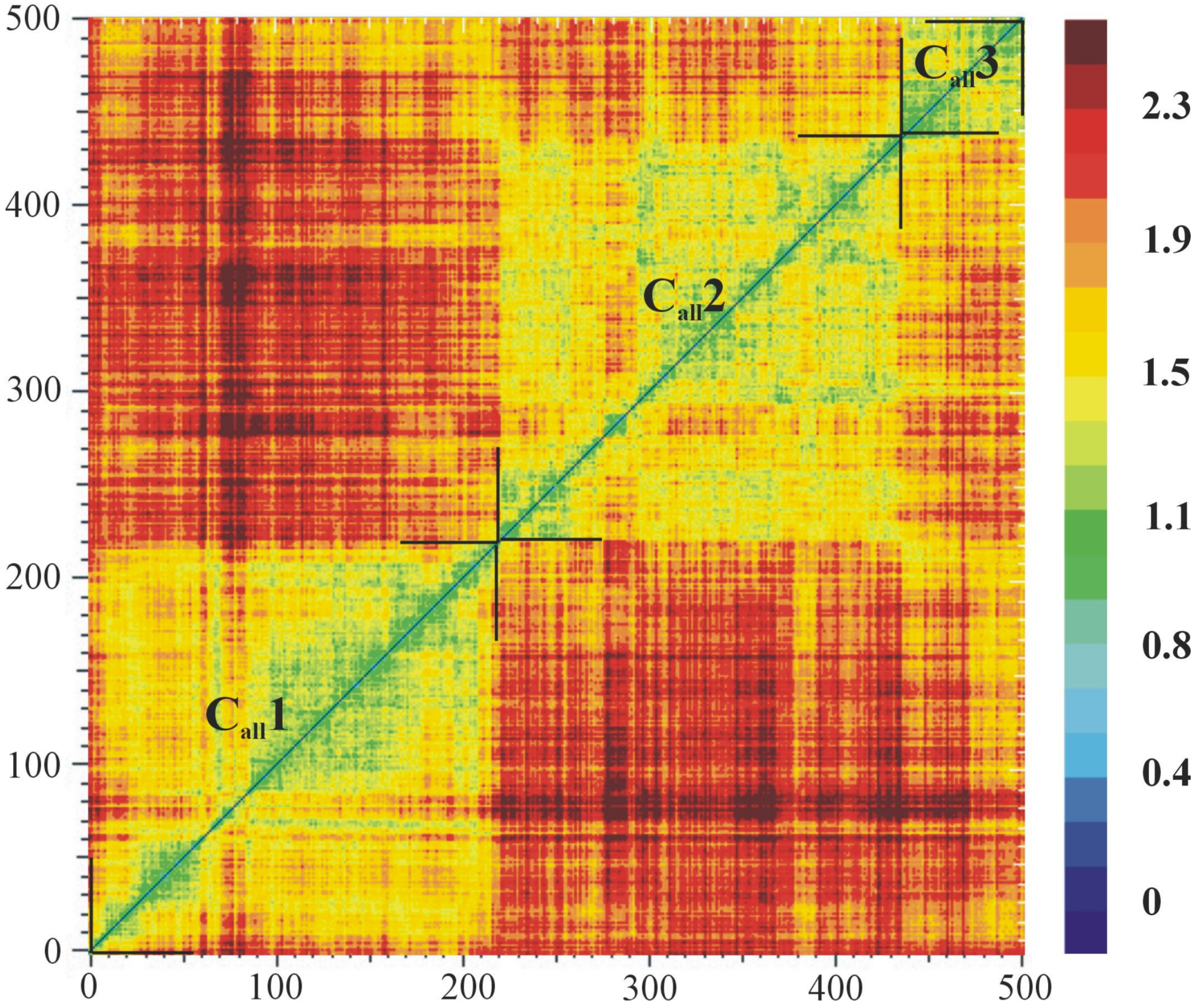


Figure 3. 2D-RMSD plot using snapshots extracted every 20 ps. The mutual RMSD between two snapshots is represented by the color code displayed on the right. All values are given in Å.

values below 1.4 Å, while conformations of different substates show RMSD values up to 2.3 Å.

To detect the molecular changes giving rise to this conformational flip, we further divided the binding pocket into individual subpockets S1, S2, S1' and S2'. Similar 2D-RMSD plots were generated (Figure 4).

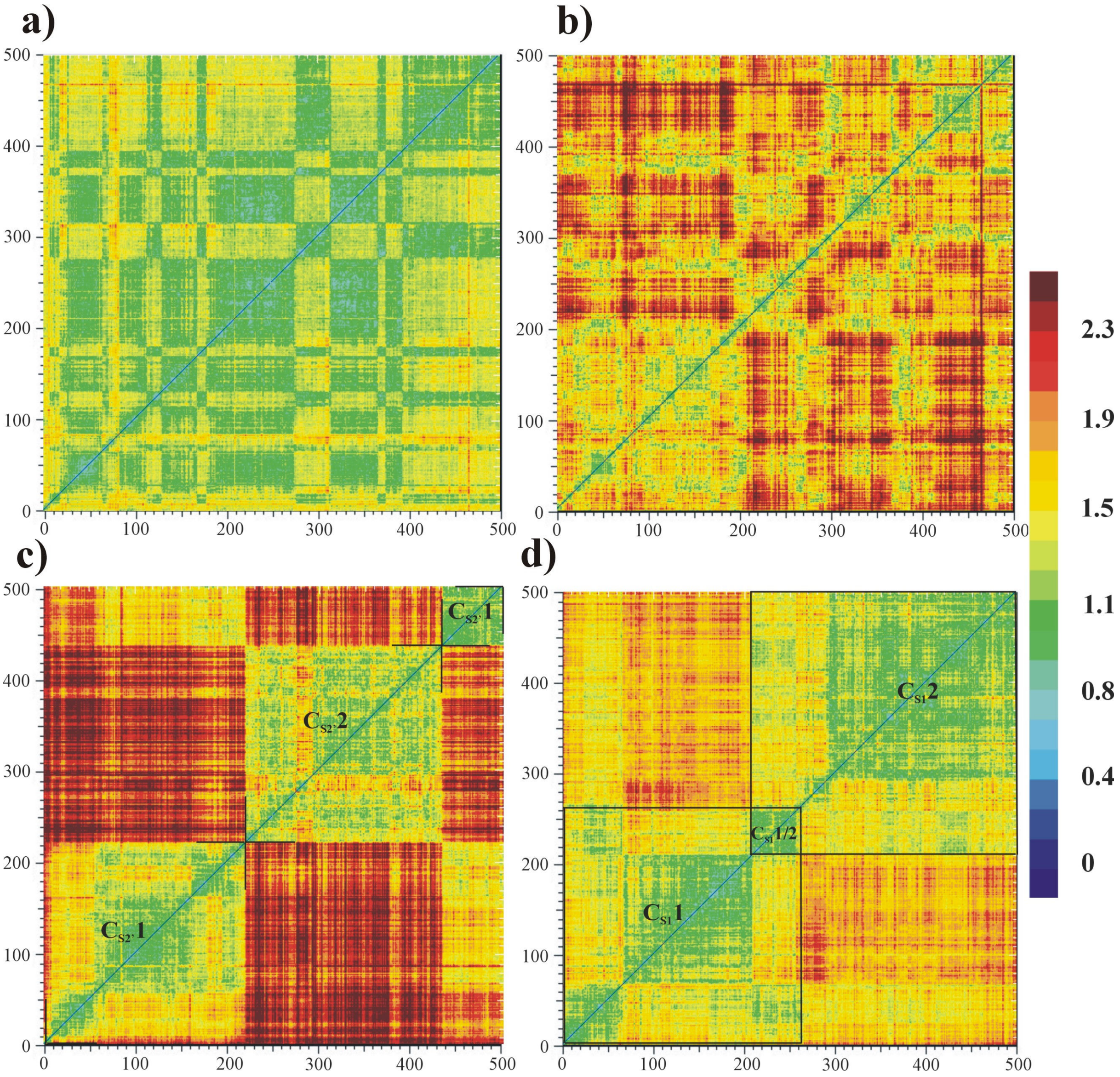


Figure 4. 2D-RMSD plot using snapshots extracted every 20 ps. The mutual RMSD between two snapshots is represented by a relative color code (s. right). All values are given in Å. Only residues of the S2 pocket are considered in **a)** of the S1' pocket in **b)** of the S2' pocket in **c)** of the S1 pocket in **d)**.

The S2 pocket (T35, S215, T217, S218, A219, T221 and I290) remains nearly unchanged and represents essentially only one conformation along the entire trajectory (Figure 4a). The RMSD deviations of these residues are considerably low with movements of 1.3 Å on average.

Similar properties are suggested for the S1' pocket (Q194, N210, I212, L292, F294, I300), however here the overall deviations among the snapshots are significantly higher (Figure 4b) and exceed to 4.0 Å (red areas). Except F294 (average RMSD 2.8 ± 0.8 Å) and L292 (average RMSD 2.4 ± 0.6 Å) all residues in this pocket are conformationally rather restricted. For F294 maximal RMSD values of 5.8 Å were detected. These results are consistent with the observations by Ersmark et al. [4, 17]. They describe a slight expansion of the S1' pocket by side chain rotations of F294 and L292 that are necessary for the accommodation of bulky P1' groups [4, 17].

Two distinct major states can be discerned from the 2D plot of the S2' pocket (G36 - N39, M75 - V78, L131 - I133, Y192) (Figure 4c). The distribution of clusters highly resembles the plot found for the entire binding pocket, indicating that the overall parent conformations $C_{all}1$, $C_{all}2$ and $C_{all}3$ are mainly determined by changes in the S2' pocket. Interestingly, a sudden and clear-cut transition from the first to the second state (called $C_{S2'}1$/$C_{S2'}2$) is observed after 4.3 ns. The second conformational family is then maintained for about 4.4 ns ($C_{S2'}2$), before a new one step transition back to geometries closely related to $C_{S2'}1$ occurs. A comparison of the 2D RMSD plots (Figure 4) indicates, that a transition of the conformational states within the S2' pocket does not affect transitions of the other subpockets. Therefore, the movements of the particular subpockets are independent to each other and not concerted.

The amino acids N76 - V78 are known to interact with P2' substituents. Their mobility is pronounced as they are involved in closing and opening of the flap. These movements are of particular interest to discover a putative binding-competent conformer to accommodate ligands with **3** as scaffold. This will be discussed in more detail in the next section. Furthermore, residues in this pocket close to the catalytic dyad (G36 - A38) show little structural mobility, whereas large fluctuations are observed for N39, M75, I133 and Y192 (Table 2). For N39, especially side chain movements ($N\chi_1$ and $N\chi_2$) are pronounced and can be important for ligand binding. Upon 180° rotation of the terminal amide function, hydrogen acceptor and hydrogen donor functionalities are mutually interchanged [7]. Similarly, for Y192 no particular state is pronouncly populated. This residue fluctuates between different states. Maximal deviations of 5.2 Å are observed. This information is of interest, since the hydroxyl group of Y192 could possibly be addressed by ligands in order to form hydrogen bonds, as observed for all pepstatin A like ligands.

Residue	Average RMSD [Å]	Standard deviation [Å]	Maximum deviation [Å]	Minimum deviation [Å]
G36	1.5	0.3	3.1	0.5
S37	1.5	0.4	2.7	0.3
A38	1.3	0.3	2.5	0.4
N39	2.0	0.3	3.4	0.3
M75	2.2	0.3	3.5	1.1
I133	3.6	1.3	8.1	0.4
Y192	2.2	0.6	5.2	0.3

Table 2. RMSD-, standard deviation-, maximum deviation- and minimum deviation- measurements for residues of the S2' pocket. Measurements were carried out including all side chain atoms.

An extensive flexibility of the S2’ pocket has also been described by Ersmark et al., mainly based on side chain movements of M75 [17]. Analysis of the Mχ_2 angle along the trajectory results in analogous observations (Figure 5). The χ_2 angle exhibits a high standard deviation of 45° from the mean value of 118°. Besides the conformation found in the crystal structure, in particular two further conformations are highly populated and seem to be energetically favorable. Similar observations are made for the χ_3 angle: a high standard deviation of 65° from the average (54°), caused by two highly populated alternative conformations. The M75 side chain points in the crystal structure into the S2’ pocket. Upon rotation it oscillates intermediately towards the interior of the enzyme, thereby opening and closing a pocket.

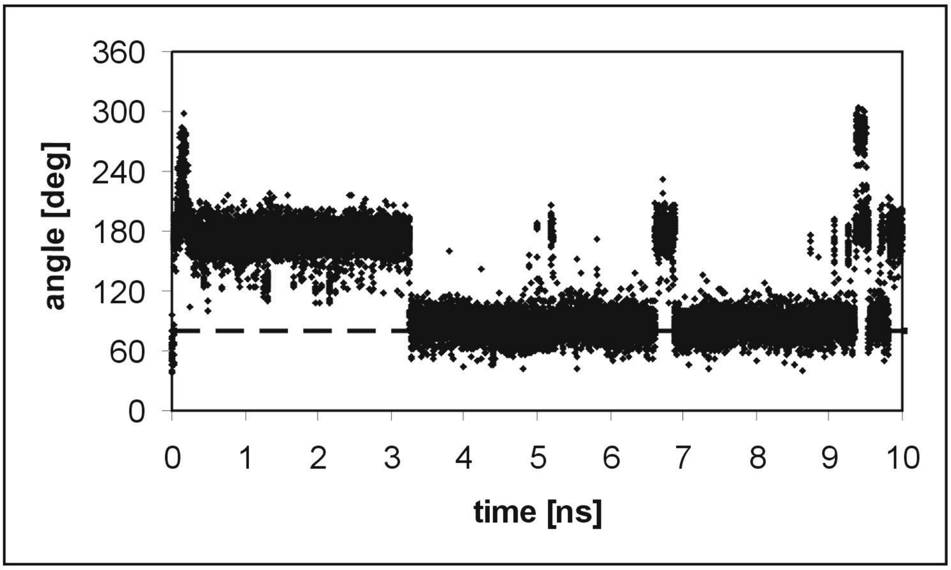


Figure 5. MD simulation of the fluctuation of the Mχ2 angle in M75 of Plm II, with the respective value from the crystal structure shown as reference.

However, I133 plays the key role. The RMSD plot (Figure 6) shows an elevated average RMSD value of 3.6 Å with very large standard deviation (1.3 Å) and a maximal deviation of 8.1 Å. For the first 4.3 ns the average RMSD is below 3 Å with low standard deviations. For the next 4.4 ns the RMSD rises to 5 Å with high standard deviations. The last 1.3 ns the values return back to values below 3 Å. These time steps are identical to those observed in the overall 2D RMSD plot (Figure 3) indicating that this residue is responsible for the distinct conformational states in the S2’ pocket.

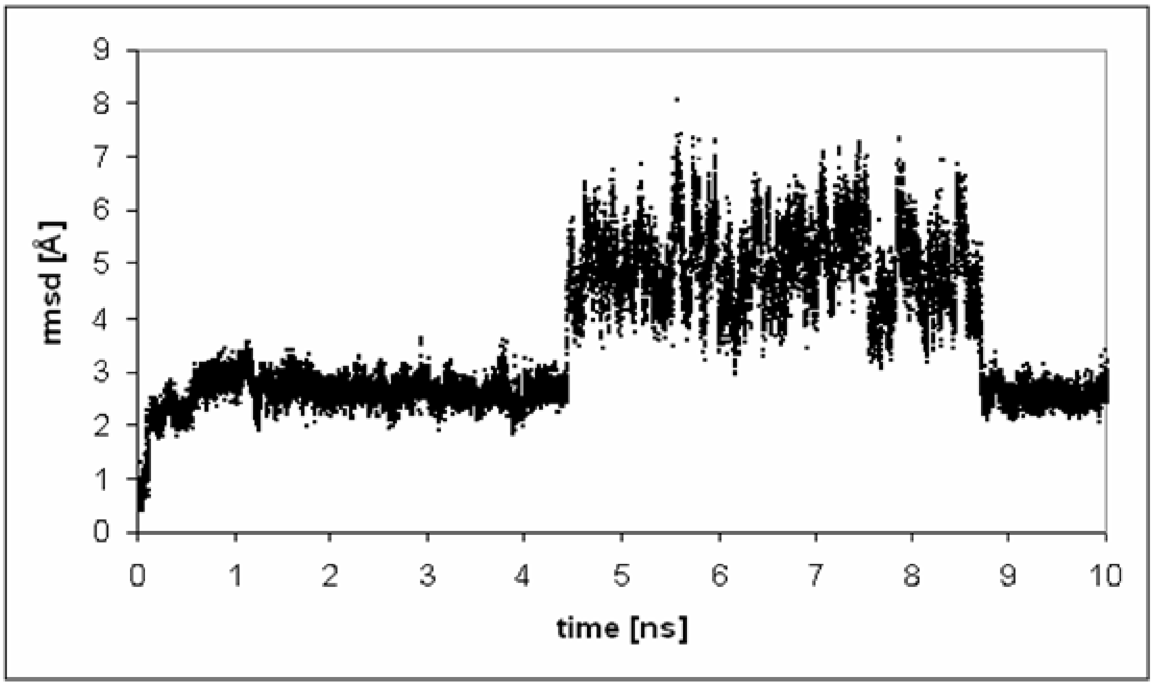


Figure 6. All atom RMS deviation of I133 for the MD simulation trajectory.

RMSD fluctuations should not be overinterpreted and have to be complemented by qualitative visual inspections. A comparison of frames extracted from each cluster provides insight into the conformational states sampled during the simulations (Figure 7). From the 2D-RMSD plot for the S2' pocket (Figure 4c) four MD-generated snapshots will be compared, showing all residues of this pocket (G36 - N39, M75 - V78, L131 - I133, Y192). The snapshots are representatives from three different clusters: The first frame shown in blue ($1_{S2'}$) belongs to the cluster in the lower left of Fig. 7d ($C_{S2'}1$), the second and third frame ($2_{S2'}$ and $3_{S2'}$) (salmon and green) originate from the cluster in the center ($C_{S2'}2$) and the fourth frame ($4_{S2'}$) (grey) is extracted from the cluster in the upper right ($C_{S2'}1$).

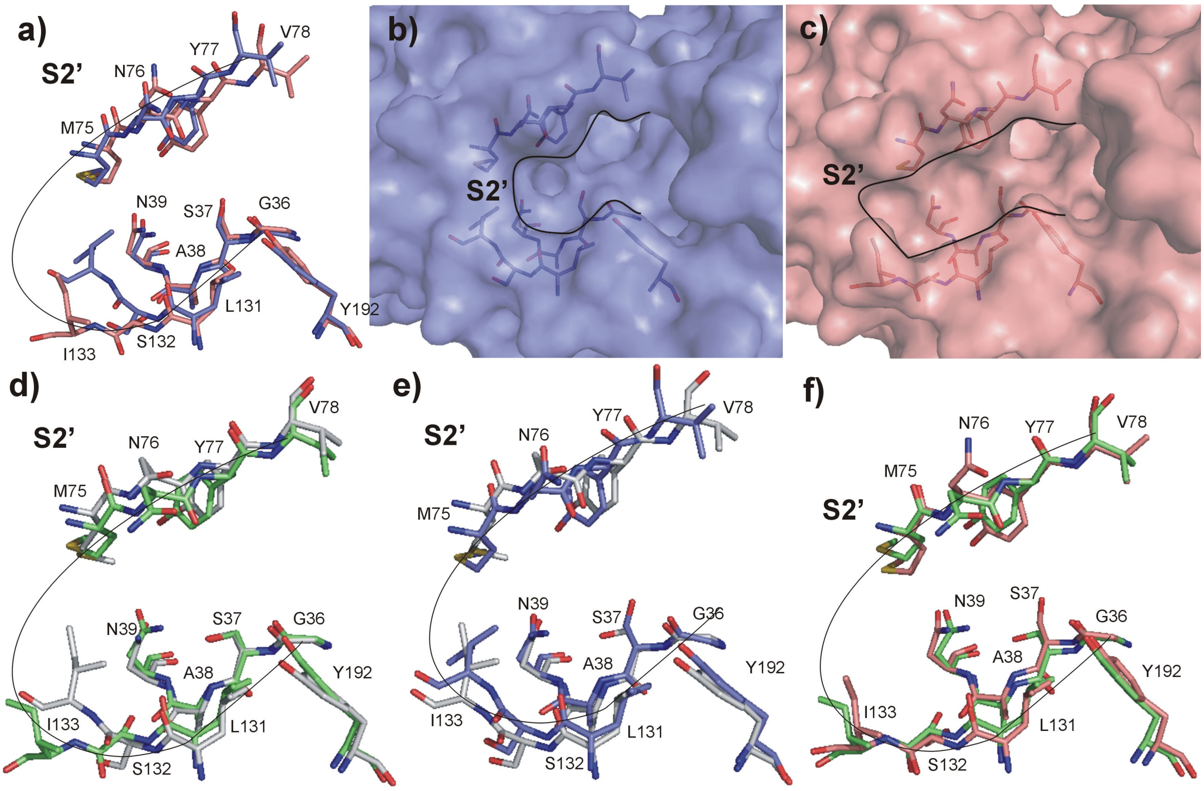

a)
S2'
N76
Y77
V78
M75
N39
S37
G36
A38
L131
Y192
S132
I133
b)
S2'
c)
S2'
d)
S2'
N76
Y77
V78
M75
N39
S37
G36
A38
Y192
I133
L131
S132
e)
S2'
N76
Y77
V78
M75
N39
S37
G36
A38
Y192
I133
L131
S132
f)
N76
Y77
V78
M75
S2'
N39
S37
G36
A38
Y192
I133
L131
S132

Figure 7. Fluctuation of the S2' pocket. **a)** Superposition of the frames $1_{S2'}$ and $2_{S2'}$ extracted from cluster $C_{S2'1}$ and $C_{S2'2}$ (Fig. 4c) **b)** Surface representation of $1_{S2'}$ **c)** Surface representation of $2_{S2'}$ **d)** Superposition of the frames $2_{S2'}$ and $4_{S2'}$ **e)** Superposition of the frames $1_{S2'}$ and $4_{S2'}$ **f)** Superposition of the frames $2_{S2'}$ and $3_{S2'}$.

The superposition of the first and second frame ($1_{S2'}/2_{S2'}$) shows that except for I133 no significant conformational variations are given (Figure 7a). Mobility of I133 is not restricted to the side-chain level, but includes the backbone as well. The C^{α}-atom is displaced by 4.2 Å and the side-chain rotates out of the S2' pocket. As a comparison of the surface representations of both frames shows, I133 swings out of space and opens up a wider pocket (Figure 7b, c). In consequence, accommodation of ligands with more bulky groups should be possible. The transition of frame $2_{S2'}$ to frame $4_{S2'}$ is illustrated in Figure 7d. As in Fig. 7a, the S2' flanking residues remain at their positions, except I133. A superposition (Figure 7e) of $1_{S2'}$, a representative of cluster 1 ($C_{S2'}1$, Fig. 4c lower left) with $4_{S2'}$ from cluster 3 ($C_{S2'}1$, Fig. 4c upper right) clearly shows that I133 adopts the same orientation in both cases and indeed both frames represent the same state. Figure 7f shows that residues found in two snapshots extracted from the same cluster occur with closely related geometry (both frames belong to cluster $C_{S2'}2$, Fig. 4c).

Since Ersmark et al. collected data only up to 2 ns and in our simulation the first conformational change occurs after 4.3 ns it agrees to our results, that as long as the system remains in one conformational state no significant displacement of the backbone portion are observed during the MD simulations [17].

The 2D-RMSD plot for the S1 pocket (M15, I32, S79, F111, T114, F120, I123) is shown in Figure 4d. In contrast to the 2D-RMSD plots for the S2' and the whole binding pocket, the borderlines between two conformational states are less clear-cut. The first conformational family is maintained for 5.2 ns ($C_{S1}1$), the second for a time period of 5.8 ns ($C_{S1}2$). Interestingly, between both clusters one sub-cluster is observed ($C_{S1}1/2$), exhibiting similarities with both. While F120, I32 and I123 appear rather rigid during the MD simulations, M15, S79, F111 and T114 are amino acids with elevated flexibility (Table 3).

Residue	Average RMSD [Å]	Standard deviation [Å]	Maximum deviation [Å]	Minimum deviation [Å]
M15	2.1	0.5	3.8	0.7
I32	1.1	0.2	2.2	0.7
S79	2.4	0.7	5.1	0.6
F111	1.7	0.6	3.8	0.3
T114	1.9	0.5	4.3	0.4
F120	1.0	0.3	2.4	0.2
I123	0.8	0.2	2.1	0.2

Table 3. RMSD-, standard deviation-, maximal deviation- and minimal deviation-measurements for residues of the S1 pocket. Measurements were carried out including all side chain atoms.

In most of the available crystal structures the hydroxyl group of S79 in the flap region is involved in the formation of crucial hydrogen bonds to a bound ligand. S79 experiences an average RMSD of 2.4 Å but with a high standard deviation of 0.7 Å leading to maximal values of about 5 Å. Similarly, the orientation of the hydroxyl group, described by the $S\chi 1$ angle is involved in this pronounced flexibility. This observation might be of interest in finding a putative binding mode for the pyrrolidinedimethylene diamine derivatives. For T114 no preferred state is highly populated, the residue fluctuates over multiple states. M15 and F111 are responsible for the occurrence of distinct states of the S1 pocket, as illustrated in Figure 8a-c. Neglecting both crucial residues for the generation of the 2D-RMSD plot results essentially in one preferred conformer of the S1 pocket described by the remaining residues along the trajectory (Figure 8a). Discarting either one of the flexible residues from the analysis produces 2D-RMSD plots with several distinct states exhibiting clearly defined borderlines (Figure 8b, 8c).

The flexibility of M15 has recently been addressed by us [16]. Fluctuations in the $M\chi 1$ and $M\chi 2$ angles result in an expansion of the S1 pocket. In consequence, the pocket successfully accommodates norstatine derivatives with large hydrophobic side chains.

In order to analyze the conformations of F111 in more detail, representative frames extracted from the $C_{S1\_M}1$ and $C_{S1\_M}2$ cluster (Figure 8b) were selected and visually inspected. The superposition of both frames shows that movements of S79, M15 and F111 cause the main structural differences in the S1 pocket (Fig. 8d). But responsible for the occurence of the distinct substates in Fig. 8b is only the rotation around the Fχ2 angle of F111.

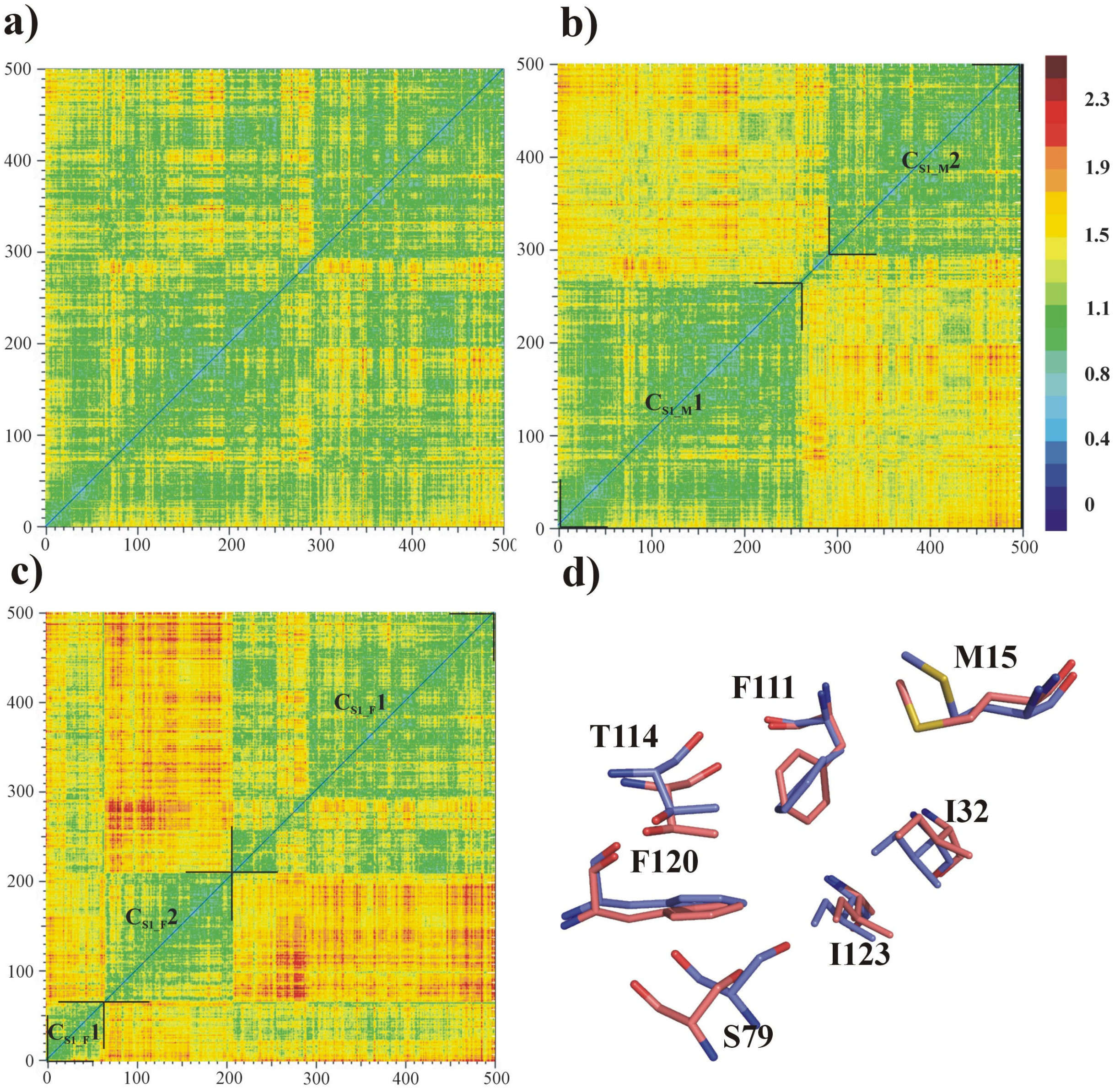


Figure 8. 2D-RMSD plot considering residues of the S1 pocket **a)** neglecting M15 and F111 **b)** discarting M15 only. Two main conformational substates can be distinguished that originate from fluctuations of F111. **c)** discarting F111 only. The observed conformational substates originate from fluctuations of M15. **d)** A frame taken from the cluster $C_{S1\_M}1$ (red) is superimposed with a representative one from cluster $C_{S1\_M}2$. The Mχ2 angle of F111 remains for the first five ns nearly unchanged but then shows a rotation about 90°. This conformation is remained for the last four ns. S79 is as well an amino acid with elevated flexibility and no preferred state is highly populated. The residue fluctuates over multiple states.

6.3.4 Putative binding mode generation for pyrrolidinedimethylene diamines

As indicated above, a conformational opening of the flap would be required in order to establish key interactions to the pyrrolidinedimethylene diamines. The involved residues N76 - S79 at the tip of the flap, indeed show pronounced flexibility during the simulations (Table 4).

Residue	Average RMSD [Å]	Standard deviation [Å]	Maximum deviation [Å]	Minimum deviation [Å]
N76	3.0	0.6	5.7	1.2
Y77	1.7	0.4	3.3	0.4
V78	2.7	0.8	5.6	0.5
S79	2.4	0.7	5.1	0.6

Table 4. RMSD, standard deviation-, maximal- and minimal-deviation measured for residues comprising the flap. Deviations are based on the average of experienced fluctuations by all side chain atoms.

During molecular dynamics the backbone atoms of the flap moves in total by 5 Å compared to the crystal structure. Snapshots along the trajectory that show a shift of 2 Å would be suitable to reasonably accommodate **3.2** (Figure 9). This is well in the range of the observed average RMSDs. Hydrogen bonds to the catalytic dyad and both flap residues could be established. The phthalimide substituent has been placed into the S2' pocket and hydrogen bond formation to Y192 is suggested. Its benzyl moiety would orient into the S1' pocket, the 2,3-dimethoxy-5-bromphenyl substituent towards the S2 pocket and the dichlorophenyl ring would occupy the S1 pocket.

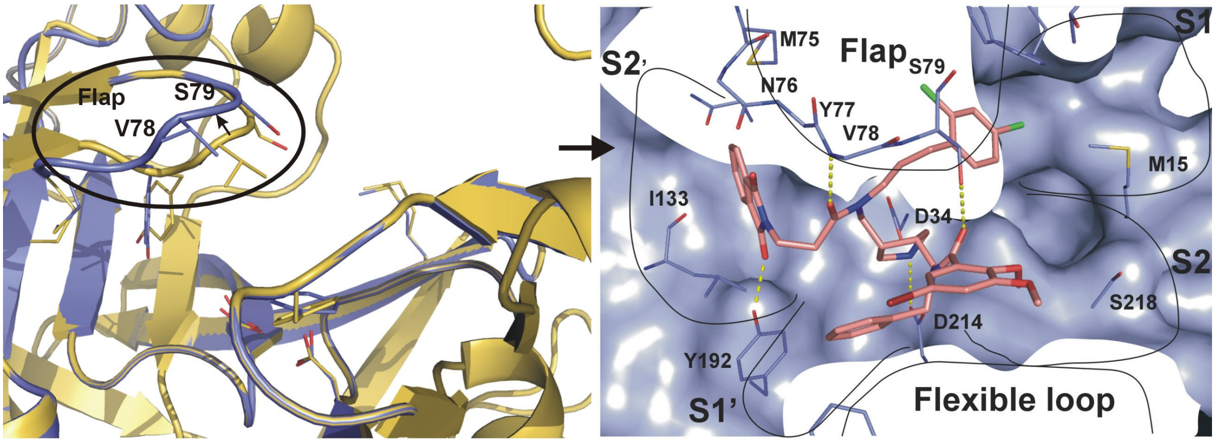


Figure 9. A conformational opening of the flap is required for the establishment of key interactions between ligand **3** and protein (s. left). The pyrrolidine scaffold of **3** was placed manually between the catalytic aspartates of 1LF2, keeping the ring geometry as observed in the crystal structure of the ligand with HIV-1 protease (1XL2). The amide oxygens were rotated towards V78 and S79 (s. right). Geometry was optimized with MOLOC, assigning full flexibility to the ligand. Apart from the flap residues N76 - S79, the protein was defined as stationary. Hydrogen bonds to D34, D214, V78, S79 and Y192 were formed (dotted yellow line).

6.3.5 Relevance and implications for inhibitor design

Plms are obviously difficult targets for structure-based ligand design due to their high structural flexibilities. In consequence, routine drug design protocols based on one single rigid protein conformer of the binding site are not sufficient to successfully predict correct binding modes. Likely, the originally picked binding pocket would only be suitable to accommodate a certain fraction of the ligands tested by virtual screening. As long as complete flexibility of the protein structure cannot be handled simultaneously or in an adaptive manner during the virtual screening campaign, the currently favoured strategy uses a limited set of representative protein conformers in order to capture all possible alternatives in a comprehensive way. This set of parent conformers may either result from multiple experimental determinations or as individual frames from MD simulations.

With respect to experimental crystal structures, undoubtedly the available data displays relevant conformations. They served as a reference in our design in generating reasonable binding modes for derivatives with **1** and **2** as core structure. However, since the presently determined crystal structures represent only three major conformations, it seems unlikely that they cover sufficiently the broad spectrum of potentially binding-competent conformers. This fact became evident for inhibitors with **3** as scaffold.

In this study, highly flexible residues could be identified by MD simulations via analysis of residue mobility and distinct conformers are apparent. Based on the results of the MD simulations, a putative binding mode for the pyrrolidinedimethylene diamine derivatives (**3.1** - **3.2**) can be suggested. Furthermore, simulations of the uncomplexed state helped to understand surprising affinity data in two cases. Interestingly, an opening of the S2’ pocket could be observed, giving space to accommodate bulky ligand side chains.

6.4 Summary and conclusion

In summary, we have discovered novel non-peptidic Plm II and IV inhibitors, featuring a pyrrolidine scaffold as core element. These compounds, originally designed for HIV-1 protease, show activity up to the nanomolar range. Plausible binding modes could be predicted for the pyrrolidine-diol-diester derivatives (**1.1 - 1.9**) and the diamino pyrrolidine-diamide (**2**), based on the crystal structure 1LF2 as a starting point. Putative binding modes for Plm II are in agreement with structure-activity relationships. Unfortunately, with the available crystallographic information it was not possible to predict a reasonable binding mode for the pyrrolidinedimethylene diamine derivatives (**3.1** and **3.2**). None of the three major binding site conformers seen in the presently available crystal structures allow accommodation. However, the structural studies indicate, that Plm II is a highly flexible protein. To tackle this problem in more detail, MD simulations of the uncomplexed state were carried out and transitions along the trajectory could be monitored. Supported by the MD simulations a putative binding mode for the inhibitors with a pyrrolidinedimethylene diamine scaffold (**3**) could be suggested and provides further insights into binding pocket adaptions that possibly have implications for Plm II inhibitor design.

6.5 Materials and Methods

6.5.1 Assays and K_i determinations

The substrate used for the Plm II assay (Bachem) is a synthetic peptide (Arg-Gln-Phe*Nphe-Ile-Thr). The asterisk denotes the cleavage site and Nphe is p-nitrophenylalanine. The Michaelis-Menten constant (K_m) for Plm II is 63 µM. The assays were performed with a Tecan Spectra Fluor spectrometer at excitation wavelength 360 nm and emission wavelength 465 nm. A volume of 180 µl assay buffer (0.1 M acetic acid/ sodium acetate pH 4.5, 20 µM substrate, 2 % DMSO) was added to 96-well plates. For the assay 18 µl of enzyme solution (0.1 M acetic acid/ sodium acetate buffer, pH 4.5) were mixed with 2 µl inhibitor solution (dissolved in DMSO). After a 5-min incubation the enzyme-inhibitor solution was added to the 180 µl substrate solution (final enzyme concentration = 1 nM). Substrate hydrolysis was recorded as an increase in fluorescence intensity over a period of 3 min.

IC_{50} values were converted to K_i values by applying the following equation:

$K_i = [IC_{50} - (E_t/2)][1 + (S/K_m)]^{-1}$

where E_t is the total enzyme concentration (1 nM), K_m the Michaelis-Menten constant (63 µM), and S the substrate concentration (18 µM).

Plm IV activity assays were similarly performed as described for Plm II. The enzyme concentration was 10 nM, K_m 28 µM and S 18 µM.

Cathepsin D activity assays were essentially performed as described for Plm II. The enzyme concentration was 1 nM; K_m =16 µM and S = 9 µM.

6.5.2 Molecular Dynamics Simulation

The MD simulation and all setup steps were performed with the Amber 8.0 suite of programs [18] using the Amber 1999 force field. The plasmepsin structure with PDB code 1LEE was used as starting point. After removal of the ligand and all crystallographic water molecules, hydrogens were added with PROTONATE. Protonation states were calculated with Poisson Boltzmann for all histidines and the aspartates of the catalytic dyad (D34 and D214) at pH 5: Both catalytic aspartates were assumed to be protonated and all histidines were set to the protonated form. Subsequently, 200 steps of minimization were applied to the protein using a generalized Born solvation model. After addition of sodium counterions to ensure neutrality the system was solvated in a box of ~10300 TIP3P water molecules [19]. The solvated system was subjected to 200 steps of minimization.

The MD simulation was then started by heating the solvent to 300 K over a period of 20 ps and cooling to 100 K over a period of 5 ps, keeping the protein fixed.

After this, the entire system was gradually brought to 300 K over a period of 25 ps. The simulation was carried out for 10 ns under constant temperature and pressure (NPT) applying periodic boundary conditions. A time step of 2 fs and PME (Particle-Mesh-Ewald) [20] for evaluating the electrostatic interactions were used. Energy data were saved every 20 fs, protein coordinates every 0.5 ps. CARNAL was used for further analysis of the trajectory and VMD 1.8.2 for visualization.

The C^{α}-atom RMS fluctuation with respect to the reference crystal was on average 2.1 ± 0.3 Å across the trajectory, indicating a high stability of the overall fold (Figure 10). Considering all atoms the trajectory-averaged RMSD value is 2.7 ± 0.3, indicating more pronounced motions once the side chains are included to the analysis.

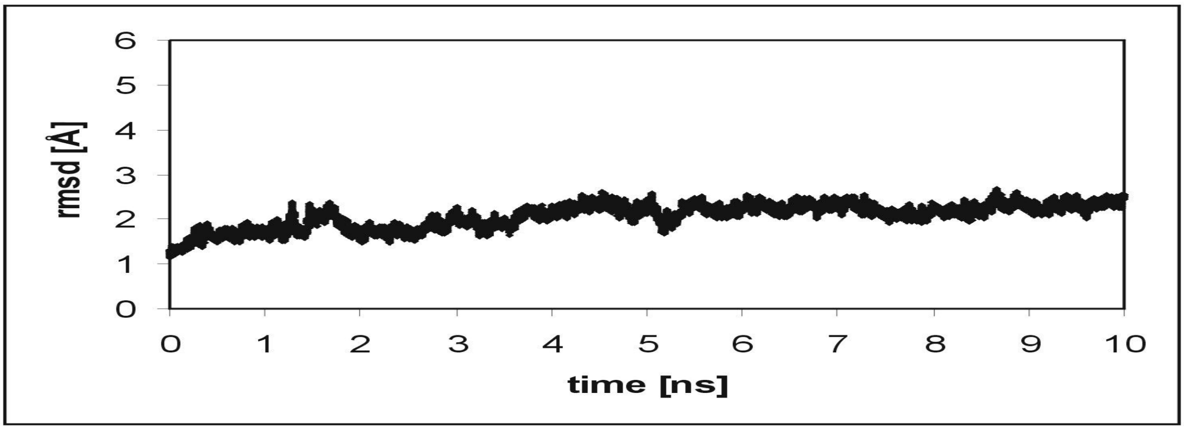


Figure 10. Cα-RMS deviation for the MD simulation trajectory.

6.6 References

1. Specker E, Böttcher J, Lilie H, et al. An old target revisited: two new privileged skeletons and an unexpected binding mode for HIV-protease inhibitors. Angew Chem Int Ed Engl 2005;44(20):3140-4.
2. Specker E, Böttcher J, Brass S, et al. 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.
3. Nezami A, Luque I, Kimura T, et al. Identification and Characterisation of Allophenylnorstatine-Based Inhibitors of Plasmepsin II, an Antimalarial Target. Biochemistry 2002;41:2273-2280.
4. Ersmark K, Feierberg I, Bjelic S, et al. C2-symmetric inhibitors of Plasmodium falciparum plasmepsin II: synthesis and theoretical predictions. Bioorg Med Chem 2003;11(17):3723-33.
5. Andrews KT, Fairlie DP, Madala PK, et al. Potencies of human immunodeficiency virus protease inhibitors in vitro against Plasmodium falciparum and in vivo against murine malaria. Antimicrob Agents Chemother 2006;50(2):639-48.
6. Bjelic S, Nervall M, Gutierrez-de-Teran H, et al. Computational inhibitor design against malaria plasmepsins. Cell Mol Life Sci 2007.
7. Luksch T, Chan N-S, Brass S, et al. Design and Synthesis of nonpeptidic antimalarial Plasmepsin II and IV inhibitors of Plasmodium falciparum. ChemMedChem submitted 2007.
8. Gustchina A, Weber IT. Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. FEBS Lett 1990;269(1):269-72.
9. Silva AM, Lee AY, Gulnik SV, et al. Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from Plasmodium falciparum. Proc Natl Acad Sci U S A 1996;93(19):10034-9.
10. Asojo OA, Afonina E, Gulnik SV, et al. Structures of Ser205 mutant plasmepsin II from Plasmodium falciparum at 1.8 A in complex with the inhibitors rs367 and rs370. Acta Crystallographica. Section D, Biological Crystallography 2002;58(Pt 12):2001-8.
11. Asojo OA, Gulnik SV, Afonina E, et al. Novel uncomplexed and complexed structures of plasmepsin II, an aspartic protease from Plasmodium falciparum. Journal of Molecular Biology 2003;327(1):173-81.
12. Prade L, Jones AF, Boss C, et al. X-ray structure of plasmepsin II complexed with a potent achiral inhibitor. J Biol Chem 2005;280(25):23837-43.
13. Cabani S, Conti G, Lepori L. Thermodynamic study on aqueous dilute solutions of organic compounds. Trans. Faraday Soc. 1971;67:1933-1942.
14. Czodrowski P, Sotriffer CA, Klebe G. Atypical protonation states in the active site of HIV-1 protease: a computational study. J Chem Inf Model 2007;47(4):1590-8.
15. Gerber PR, Muller K. Mab, a Generally Applicable Molecular-Force Field for Structure Modeling in Medicinal Chemistry. Journal of Computer-Aided Molecular Design 1995;9(3):251-268.
16. Weik S, Luksch T, Evers A, et al. The potential of P1 site alterations in peptidomimetic protease inhibitors as suggested by virtual screening and explored by the use of C-C-coupling reagents. ChemMedChem 2006;1(4):445-57.
17. Ersmark K, Nervall M, Hamelink E, et al. Synthesis of malarial plasmepsin inhibitors and prediction of binding modes by molecular dynamics simulations. J Med Chem 2005;48(19):6090-106.

18. Case DA, Cheatham TE, 3rd, Darden T, et al. The Amber biomolecular simulation programs. J Comput Chem 2005;26(16):1668-88.

19. Jorgensen W.L. JC, Jeffry D. Madura, Roger W. Impey and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics 1983;79(2):926-935.

20. Darden T, D. York, and L. Pedersen. Panicle mesh Ewald-an N.Log(N) method for Ewald sums in large systems. J. Chem. Phys. 1993;98:10089-10092.

7 Lead identification for the antimalarial target Plm by using Ftrees

7.1 Abstract

Plasmepsin (Plm) II (EC-number: 3.4.23.39) and IV (EC-number: 3.4.23.B14 Plm IV) are aspartic proteases present in the food vacuole of the malaria parasite *Plasmodium falciparum* which is involved in host hemoglobin degradation. In order to find new plasmepsin inhibitors with novel lead structures, we performed a feature tree (Ftree) search based on a Plm inhibitor, recently developed in our group, bearing a tetrahydro-azepine as core structural element. Ftree is a fast method to detect molecular similarity in terms of physicochemical properties between small organic compounds. In order to retrieve candidate molecules for testing with similar binding properties but exhibiting a molecular skeleton deviating from the query reference we applied two different search strategies, leading to new inhibitors with K_i values in the micromolar range. Novel scaffolds could be discovered. For the most promising one, based on a central thiophen moiety, a series of compounds was synthesized. By means of the Gewald reaction modifications of the P1, P1' and P2' substituents were easy accessible and within short time the inhibitor scaffold could be optimized to nanomolar range.

7.2 Introduction

Recently we have developed novel non-peptidic Plm II and Plm IV inhibibitors featuring a 2,3,4,7-tetrahydro-1*H*-azepine scaffold as core element. Equipped with suitable side chains to address two of the enzyme's specificity pockets, first inhibitors showed activity up to the low micromolar range [1]. Further investigations were performed to improve inhibitor potency [1]. Simultaneously, based on the promising results obtained for the azepines, we intended a scaffold hopping approach to detect novel core structures.

7.3 Selection of the screening strategy

For the purpose to detect novel promising core structures, we applied Ftrees, a software tool capable to detect new but property-related scaffolds [2]. It is a time-efficient method that allows to screen huge small molecule databases for compounds with high molecular similarity with a given reference. Additionally, the calculated similarity values are straight forward to interpret by a simple mapping of molecular fragments. Instead of a linear molecular representation in terms of fingerprints or by calculating sophisticated three-dimensional structural alignment of molecules, *feature tree* is a two-dimensional description, reflecting a rough approximation of a compound's structural formula [2].

Feature trees are composed by nodes that are connected with edges. The nodes represent physicochemical and sterical properties of a molecule and the edges illustrate how these nodes are topologically linked [2]. There are two types of feature trees available: If almost every functional group of a molecule is assigned to a node, the feature tree depicted the molecular properties as “high resolution feature tree”. If the physicochemical and sterical properties of the whole molecule are stored in one single node, the representation is called “low resolution feature tree”.

The comparison of two feature trees is based on the matching of Ftrees onto each other [2]. Two similarity values are computed:

- For the calculation of the s*imilarity value at level 0*, each molecule is represented by a feature tree with only one node (Low resolution feature tree). These single nodes are mapped onto one another [2].

- For the *similarity value at level x*, each acyclic atom is considered as a single node and rings are merged into single nodes (High resolution feature tree). The trees are broken down recursively into smaller parts which are topologically matched onto one another [3].

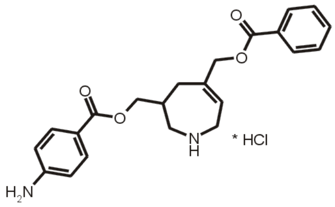
7.4 Results and discussion

7.4.1 Ftree similarity value at level X

The goal was to find compounds with a comparable biological activity to the query molecule, but displaying a low degree of chemical topology in structural terms.

First a high resolution Ftree search was performed. The azepine derivative **1**, bearing a *p*-aminobenzoyloxymethyl- and a benzoyloxymethyl-group, was used as reference compound (Table 1). It shows a K_i of 5 µM for Plm II and 7 µM for Plm IV and displays convincing selectivity with respect to the human cathepsin D [1].

Table 1. Experimental Inhibitory Activities of the Compounds Studied

Compound	Structure	Enzyme K_i (µM)		
		Plm II	Plm IV	Cat D
1		5	7	260

The generated Ftree of compound **1** is shown in figure 1. The nodes are represented by a color-code representing their property profile.

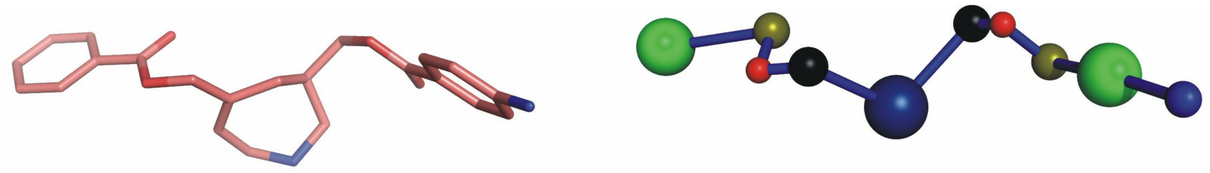


Figure 1: Compound **1** and its corresponding high resolution feature tree. The nodes are labeled with features representing steric and physicochemical properties. Features are visualized by colors and sizes of the feature tree nodes. The two aromatic ring systems are displayed as green nodes, the two hydrophobic methylene linkers by black spheres. The amino functions, representing hydrogen bond donor groups are indicated as blue spheres. Hydrogen bond accepting groups are coloured as red nodes. Nodes having a mixed profile are indicated by a corresponding mixed color (i.e. the hydrophobic seven membered ring with its basic secondary amino function is displayed by a dark blue node and the hydrogen bond accepting carbonyl group with the ester carbon as a gold node).

For the Ftree analysis about two million candidate molecules were extracted from the ZINC database [4]. Ftrees were generated and subsequently compared with the query compound. The highest ranking compound exhibits a similarity value of 0.92 (Figure 2c) while a value of 1 means very similar and 0 means dissimilar. A closer inspection on the mapping of the generated feature tree with that of the query compound illustrates the similarities (Figure 2 b, d). Each node of the query molecule is similarly present in the target compound, thereby maintaining the topology. The only exception is found for the azepine moiety, representing the core structural element of **1**. A dark blue node is assigned to the seven membered aza-cycle. During the comparison this node is matched with a hydrogen bond donor (blue) and two aliphatic (black) nodes, representing the dimethyl-amino group of the top ranked compound. Based on the asymmetric azepine reference a symmetric hit with a dimethyl-amino core structure was suggested. An advantage from a synthetical point of view in comparison to the reference compound is the absence of stereo centers.

Anyhow, the observed structural transition does not represent the expected scaffold hopping: Both molecules exhibit a central amino functionality that might interact with the catalytic aspartates. Furthermore, both, **1** and **2** are decorated with benzoic ester substituents that might interact in a similar fashion with the residues of the enzyme's active site.

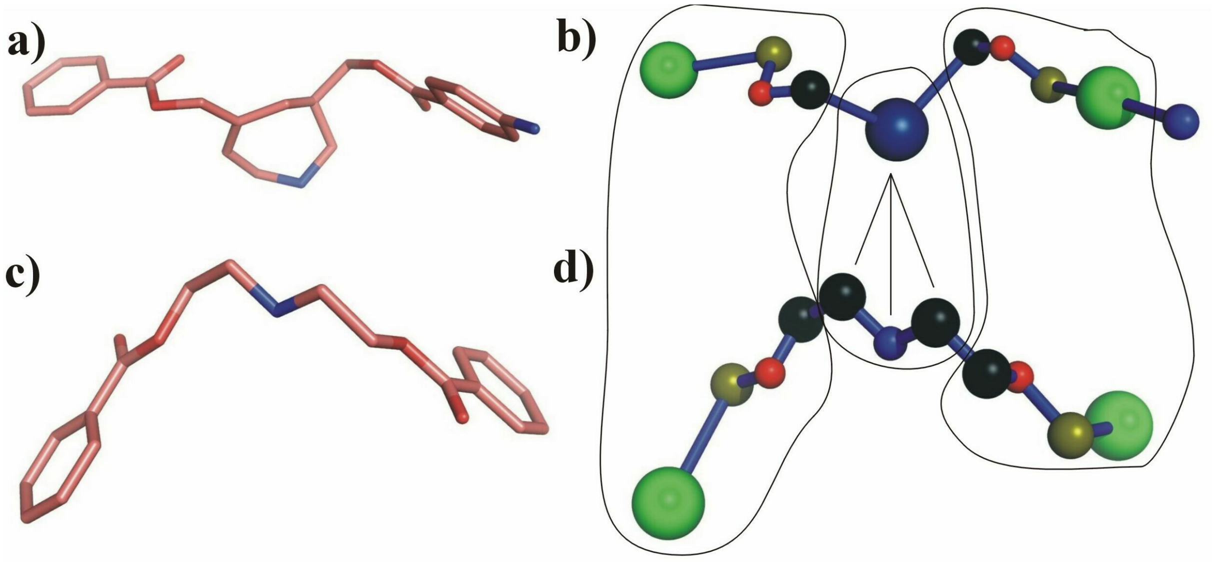


Figure 2: Compound **a) 1** and **c) 2** and its corresponding feature trees **b)**, **d)**.

Since we were interested in finding novel inhibitor scaffolds with a non-trivial degree of structural similarity, we modified our screening conditions.

7.4.2 Ftree value at level 0

As an alternative, a low resolution Ftree search has been performed. Analysis of the results in terms of the achieved *similarity value at level 0* is applied as prefiltering step, particularly as the method is very time efficient. However, considering the crudeness of the similarity measure, a subsequent second filter has been accomplished.

Again, the ZINC database was screened, now for molecules with the highest *similarity values at level 0*, with respect to the query compound **1**. Since structural information is available for Plm II, we decided to apply docking as the second filtering step. Thus, the top 0.5 % ranking molecules out of the ftree search (10,000 compounds) were selected for docking into the Plm II crystal structure with pdb-code 1LF2. For each compound, the best ranked solutions were written to disk. Subsequently, the top-scored 100 solutions were visually examined.

Compound **3** (Tab. 3) deserves particular interest: A *similarity value at level 0* of 0.94 was detected. Additionally, **3** was highly scored by the FlexX scoring function upon docking with FlexX [5]. Finally, visual inspection of the generated docking pose suggested a reasonable binding mode (Figure 3). The secondary amine is thought to interact with the two aspartates of the catalytic dyad, while the phenyl moiety could address the residues of the S1 pocket (Figure 3 right). The rather rigid fluorophenyl-thiophen group would occupy the S1' pocket. The ethylester group is assumed to point towards the S2' pocket, thereby possibly forming a hydrogen bond to Y192 (Figure 3 left).

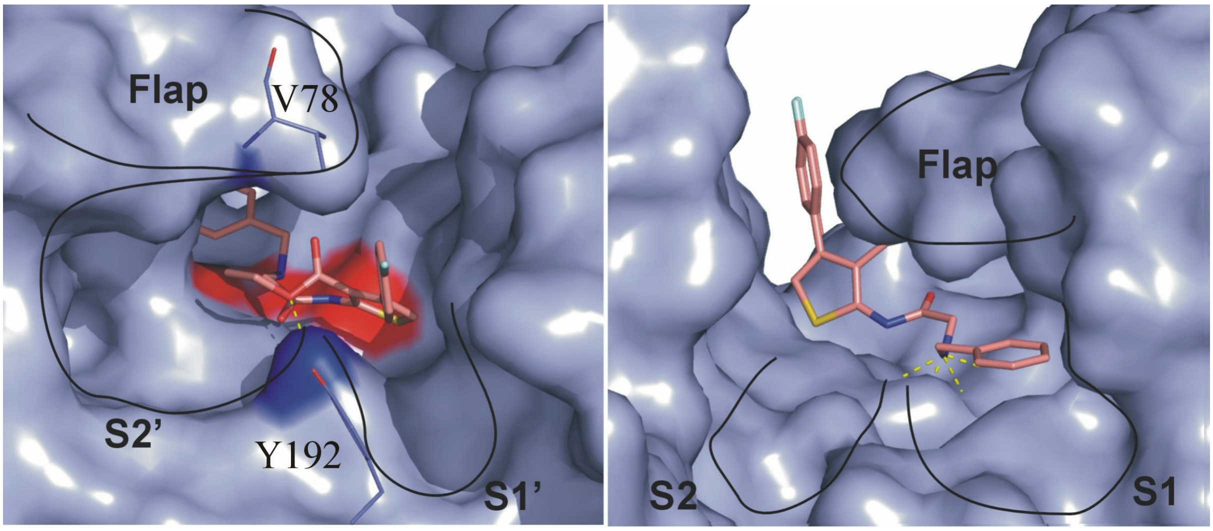


Figure 3. Compound **3** docked with FlexX into the Plm II structure with pdb-code 1LF2. Two different views on the binding pocket are illustrated. Interactions of the ligand to the flap, S1' and S2' pocket are shown on the left. Interactions to the catalytic dyad (yellow dots) and to the S1 pocket are represented on the right.

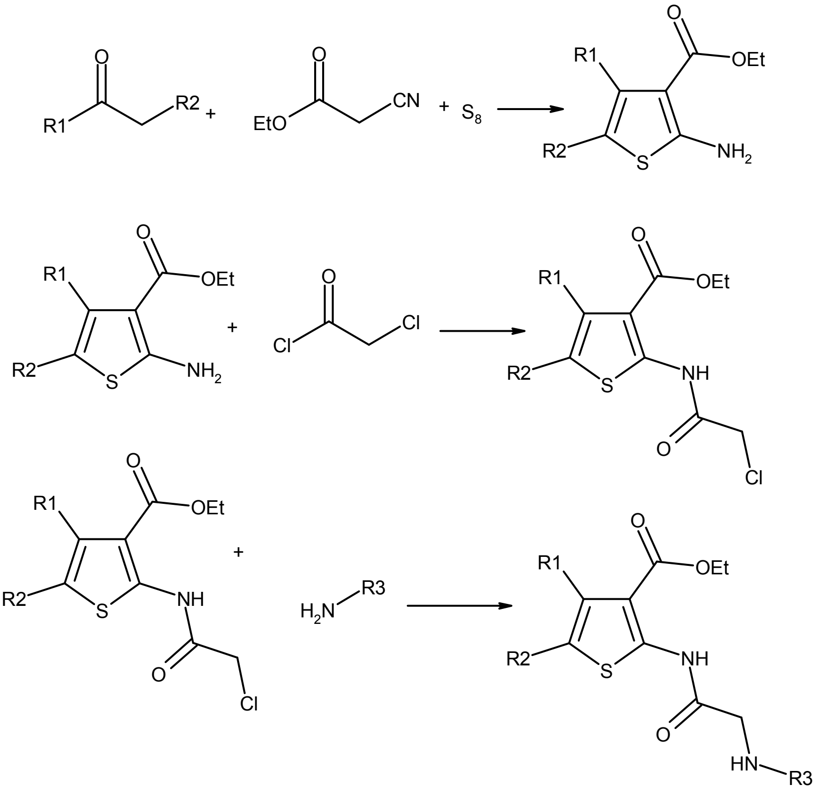
Remarkably, molecules **1** and **3** are rather dissimilar from a scaffold point of view (Fig. 4 a, b). While **1** is bearing an azepine as core structure, **3** is equipped with a substituted thiophen as central scaffold.

Compound **3** was purchased and tested for activity against Plm II and IV. K_i values in the double digit micromolar range could be obtained for both proteins (Table 3). These affinities are moderate but considering the docking solution there is apparently much more space available for optimization. Especially a substitution of the ethylester by sterically more demanding groups could lead to an increase in potency since the S2' pocket is very spacious.

Furthermore, introduction of substituents at the P1-phenyl ring or substitution of the ring system by diverse groups might improve affinity, as observed in recent studies for the azepine derivatives [1].

7.4.3 Synthesis

A triple substituted thiophen moiety was detected as the central scaffold by feature trees. This core structural element is accessible in three steps via the Gewald reaction (Scheme 1). Modifications of the P1, P1' and P2' substituents can be easily achieved by this synthesis route.



Scheme 1. Synthesis of the core scaffold in three steps.

7.4.4 Biological Evaluation

In a first attempt **3** was re-synthesized, and compared with the purchased compound. Within the estimated error range of the assay (± 40 %), the affinity data could be reproduced.

The docking solution suggests that the *p*-fluoro substituent is oriented towards the solvent (Figure 3 left) without making direct interactions with Plm II. The adjacent phenyl moiety is assumed to perform hydrophobic interactions with amino acids of the S1' pocket. Compounds **4** - **6** were synthesized to validate this hypothesis. For Plm II, replacement of the *p*-fluoro by a *p*-chloro substituent does not alter binding affinity. Also the unsubstituted derivative falls into the same range. The same holds for Plm IV binding. These findings support our hypothesis that the para position at this phenyl ring is exposed towards the solvent. However, in agreement with our docking solution, total removal of the phenyl substituent (**6**) deteriorates binding to Plm II. Interestingly, for Plm IV, **6** actually does not show a decrease in affinity data.

For optimization of the P1 substituent, compounds were synthesized varying the steric bulk and the exposed physico-chemical properties (**7**-**12**). In a previous design of azepine derivatives it could be shown that inhibitors featuring a *p*-aminophenyl or an indolyl moiety experienced improved binding compared to a phenyl substituent [1]. Both groups were introduced into the current scaffold. Indeed, compared to the initial hit **3**, both, the *p*-aminophenyl substituted **7**, and the indolyl derivative **8**, revealed stronger binding towards Plm II and Plm IV.

Instead, **9** exhibiting a pyridine moiety at P1 shows nearly unchanged affinity compared to unsubstituted phenyl derivative **5**.

Elaborate MD simulation have indicated that the S1 pocket can change its conformation and give space to accommodate significantly larger substituents. To test this hypothesis with the presently studied compound series, **10** was synthesized. It shows unchanged binding affinity which neither confirms nor speaks against occupancy of this moiety in the S1 pocket.

Interestingly, **12**, substituted with a *m*-amino group, reveals improved binding either towards Plm II and Plm IV. The synthetic nitro precursor **11** of **12**, was also tested for Plm inhibition. The detected binding demonstrates that also a H-bond acceptor is tolerated but not favored at this position.

Table 3. Experimental Inhibitory Activities of the Compounds Studied

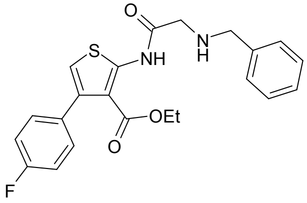
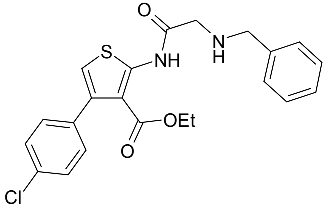
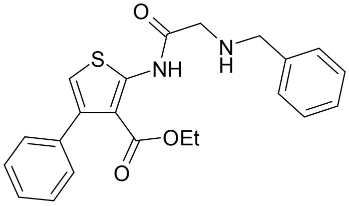
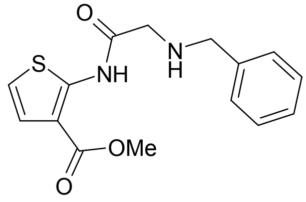
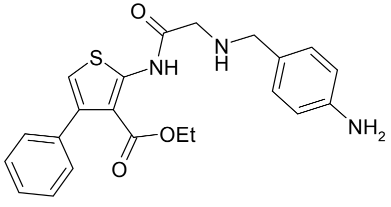
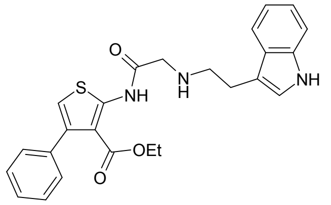
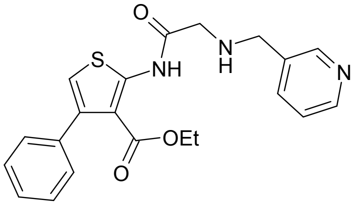
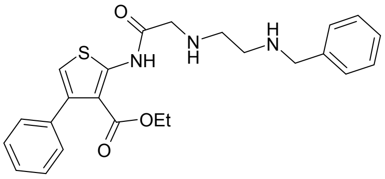
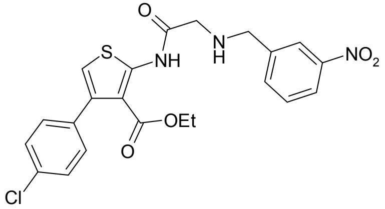
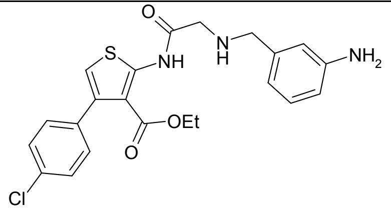
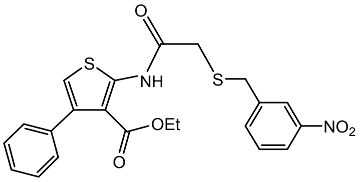
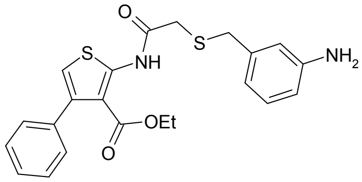
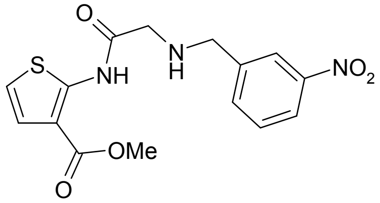
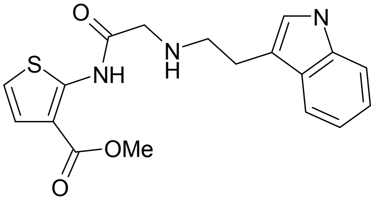
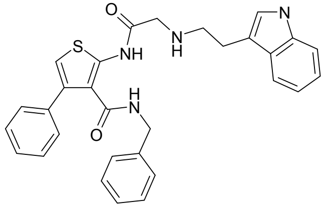
Compound	Structure	Enzyme K_i (μM)	
		Plm II	PlmIV
3		52 69	50 66
4		60	68
5		59	36
6		183	37
7		13	17
8		20	5
9		35	40

Table 3. Experimental Inhibitory Activities of the Compounds Studied

Compound	Structure	Enzyme K_i (μM)	
		Plm II	PlmIV
10		36	19
11		39	23
12		8	5
13		n.i.	n.i.
14		22	13
15		214	14
16		53	10
17		2,8	0,6

In our suggested binding model, we assume that the secondary amine plays a crucial role as it is pretended to form two hydrogen bonds to the catalytic aspartates. To test this hypothesis, sulfur analogs to **11** and **12** have been synthesized. **13** is unable to inhibit both Plms, thus confirming our model. However, the K_i values for **14** are surprisingly low even though it looses by a factor of 2-3 in affinity. This observation is quite surprising. Without further binding data and particularly lacking a corresponding crystal structure the pronounced binding of **14** is difficult to explain. Assuming unchanged overall binding mode, one might speculate that the sulfur derivatives do not directly interact with the catalytic dyad but e.g. via an interstitial water molecule. Such behavior has been detected crystallographically for a Plm II inhibitor complex (PDB-code 2BJU). The nitro derivative is also the weaker binder in case of the amino derivative **11**. Possibly the corresponding sulfur analog looses too much in binding that our assay is unable to detect any residual weak binding.

The indicated trend that an unsubstituted thiophen ring in position 3 improves binding to Plm II, but has almost no effect on Plm IV, is further confirmed by **15** and **16**. In comparison to **11** and **8**, these inhibitors show reduced inhibition to Plm II but almost unchanged affinity data to Plm IV.

To explore the importance of the S2' pocket substituent, compound **8** was selected as parent structure. Replacement of the ethylester functionality by a benzylamide (**17**) enhances, as expected, binding affinity. Seven fold improvement in potency compared to **8** is achieved for either Plm II and Plm IV, leading to the first submicromolar inhibitor of this series.

7.5 Summary and conclusion

In summary, we have discovered by a Ftree similarity search combined with docking a novel scaffold suited as lead for Plm inhibition. For this core structure, consisting of a central thiophen moiety, a series of compounds was synthesized. An appropriate synthesis scheme could be established based on the Gewald reaction, to address the S1, S1' and S2' subpockets. First inhibitors with affinities in the two-digit micromolar range could be improved by one order of magnitude decorating the molecules with an indolyl or *m*-aminophenyl group as P1 substituent. Overall, 18 inhibitors were synthesized. The experimentally determined K_i values are to a large degree in accordance with the putative binding modes. However, affinity data measured for the thioether derivative **14** show, that the presently assumed binding mode can not explain all structure-activity relationship data.

The most potent inhibitor within this series that was suggested for synthesis based on the docking studies, exhibits a K_i value in micromolar range for Plm II and submicromolar range for Plm IV. This molecule showed a 110-fold gain in affinity compared to the initial hit **3**.

7.6 Materials and methods

Assays and K_i determinations

The Plm II assays were performed in 96-well microtiter plates with a Tecan Spectra Fluor spectrometer (excitation = 360 nm, emission = 465 nm). The fluorogenic substrate Arg-Gln-Phe*Nphe-Ile-Thr was purchased from Bachem (the asterisk denotes the cleavage site and Nphe is p-nitrophenylalanine). A 0,1 M sodium acetate buffer (pH 4,5) with 2 % DMSO and 20 µM substrate was used. 180 µl of this buffer were pipetted in each well and mixed with 20 µl of a Plm II/inhibitor solution (the inhibitor was first dissolved in DMSO and then 2 µl were mixed with 18 µl of a 10 nM Plm II solution).

The change of fluorescence intensity was recorded as a function of time. IC_{50} values were taken from plots of V_i/V_0 as a function of inhibitor concentration, for which V_i and V_0 are the initial rates of reaction in presence and absence of the inhibitor, respectively. The following equation was applied to convert IC_{50} values to K_i values:

$K_i = [IC_{50} - (E_t/2)][1 + (S/K_m)]^{-1}$

The variables E_t and S are the concentrations of enzyme and substrate. The K_m for the substrate was determined to be 63 µM.

Plm IV activity assays were essentially performed as described for Plm II. The enzyme concentration was 10 nM, K_m 28 µM and S 18 µM.

Ftree search

The molecular similarity searches were performed with the program Ftrees (version 1.5.4) [2]. Compound **1** was used as query molecule. The ligand was preprocessed with Sybyl [6]: atom types as well as formal charges were assigned, hydrogens were added, 3D structures were generated and energy minimized. The ZINC database [4] (2 million molecules) was queried for similar compounds to **1**. All molecules were given in mol2-format. At first, for all molecules feature trees were generated. *Similarity values at level 0* and *similarity values at level 1* were calculated using the *match search* algorithm. Default values were assigned to the flags.

Computational Docking

Docking studies were performed with the docking program FlexX (version 2.0.3) [5]. The crystal structures of Plm II with PDB code 1LF2 was used as starting geometry. The ligand and all crystallographic water molecules were removed. For the active site determination a radius of 13 Å around the reference ligand was selected, always using complete amino acids. Ligands were extracted from the ZINC database in a ready-to-dock, mol2-format [4]. For subsequent detailed studies, **3** was preprocessed with Sybyl [6] as follows: atom and bond types as well as formal charges were assigned, hydrogens were added and the 3D structure was generated. Energy minimization was performed using conjugate gradient to a gradient of 0.05 kcal/mol. All compounds were docked by using standard parameters (*complex all* option). During placement, the FlexX scoring function was applied to obtain the final selection of the best docking solutions.

7.7 Literature

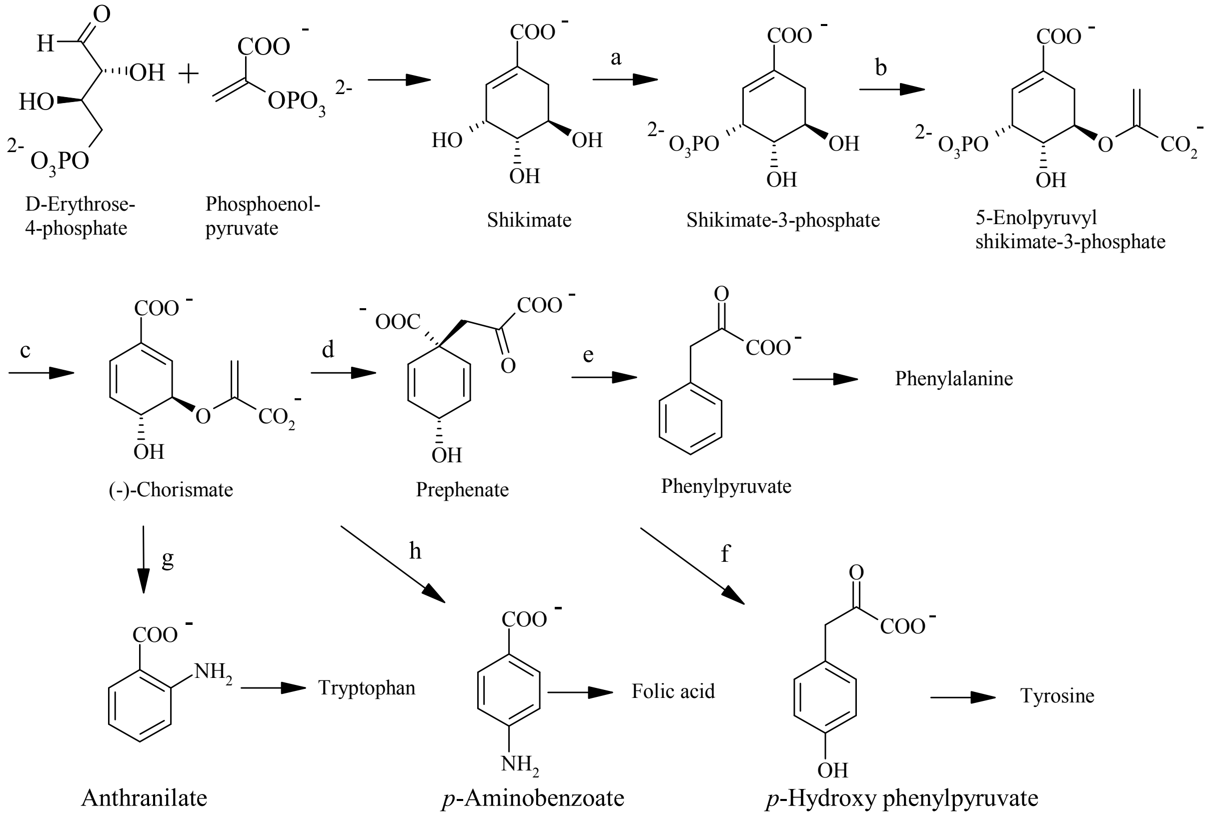
1. Luksch T, Chan N-S, Brass S, et al. Design and Synthesis of nonpeptidic antimalarial Plasmepsin II and IV inhibitors of Plasmodium falciparum. ChemMedChem submitted 2007.
2. Rarey M, Dixon JS. Feature trees: a new molecular similarity measure based on tree matching. J Comput Aided Mol Des 1998;12(5):471-90.
3. Rarey M, Zimmermann M, Hindle SA, et al. Feature Trees Ftrees v.1.5.4 User Guide.
4. Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. J Chem Inf Model 2005;45(1):177-82.
5. Rarey M, Kramer B, Lengauer T, et al. A fast flexible docking method using an incremental construction algorithm. Journal of Molecular Biology 1996;261(3):470-489.
6. Tripos A. St. Louis, MO, U.S.A, 2004.

8 Strukturlösung von Prephenat Dehydratase aus *Methanocaldococcus jannaschii*

8.1 Einleitung

8.1.1 Shikimat Biosyntheseweg

Die Biosynthese der aromatischen Aminosäuren Phenylalanin, Tyrosin und Tryptophan erfolgt über den Shikimat-Weg. Dieser Biosyntheseweg existiert nur in Pflanzen und Mikroorganismen. Tiere dagegen müssen die aromatischen Aminosäuren über Nährstoffe aufnehmen. Moleküle, die diesen Weg hemmen können, sind deshalb von grossem Interesse als mögliche Herbizide und Antibiotika. Die Prephenat Dehydratase gehört zum Shikimat-Biosyntheseweg (Schema 1), in dem D-Erythrose-4-phosphat und Phosphoenolpyruvat über Shikimat in verschiedene aromatische Aminosäuren umgewandelt werden.


D-Erythrose-4-phosphate
Phosphoenol-pyruvate
Shikimate
Shikimate-3-phosphate
5-Enolpyruvyl shikimate-3-phosphate
(-)-Chorismate
Prephenate
Phenylpyruvate
Phenylalanine
Anthranilate
Tryptophan
p-Aminobenzoate
Folic acid
p-Hydroxy phenylpyruvate
Tyrosine
a
b
c
d
e
f
g
h

Schema 1. a = Shikimat Kinase; b = 5-Enolpyruvyl Shikimat-3-phosphat Synthase; c = Chorismat Synthase; d = Chorismat Mutase; e = Prephenat Dehydratase; f = Prephenat Dehydrogenase; g = Anthranilat Synthase; h = p-Aminobenzoat Synthase [1]

8.1.2 Prephenat Dehydratase

Die meisten aromatischen Verbindungen in Bakterien, Pilzen und höheren Pflanzen entstehen im Shikimat Biosyntheseweg. Dabei ist die Prephenat Dehydratase (PDT) ein Schlüsselenzym bei der L-Phenylalanin-Biosynthese [2]. PDTs sind essenziell für das Überleben der oben genannten Organismen, wie unter anderem für PDT aus *Myobacterium tuberculosis* nachgewiesen wurde [3-5]. Das Fehlen eines analogen Enzyms im menschlichen Organismus, macht die PDT zu einem vielversprechenden Target gegen viele Krankheitserreger [2].

Prephenat Dehydratasen katalysieren die Umwandlung von Prephenat zu Phenylpyruvat, unter Freisetzung von Wasser und Kohlenstoffdioxid (Abbildung 1) [6].

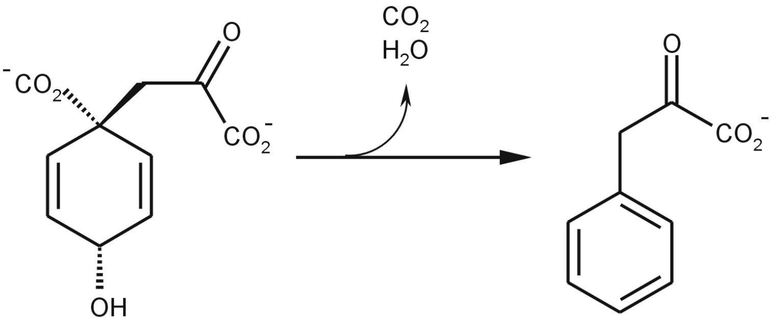


Abbildung 1. Umwandlung von Prephenat zu Phenyl-Pyruvat.

Bei der Abwesenheit dieses Enzyms findet die Umwandlung Säure-katalysiert in zwei Schritten über eine Carbokation-Zwischenstufe statt. Kinetische Messungen haben gezeigt, dass PDTs diese Reaktion um einen Faktor von $> 10^6$ beschleunigen [7]. Obwohl einige PDTs biochemisch charakterisiert wurden, gab es bis zu Beginn dieser Arbeit keine PDT-Kristallstrukturen. Zusätzlich ist der Reaktionsmechanismus weitgehend ungeklärt [8-10]. PDTs liegen in den verschiedenen Organismen als monofunktionale oder multifunktionale Enzyme vor. Das *E.coli* P-Protein ist zum Beispiel ein bifunktionales Enzym, das aus einer PDT Domäne und einer AroQ Chorismatmutase Domäne am N-Terminus zusammengesetzt ist [7]. Meistens besitzen PDTs eine C-terminale regulatorische ACT Domäne, die L-Phenylalanin oder andere Aminosäuren bindet [11]. Dabei kommt es beim Binden von L-Phenylalanin in der Regel zur Hemmung des Proteins, während Tyrosin häufig als Aktivator wirkt. Wechselwirkungen zu anderen Aminosäuren wie Tryptophan, Methionin, Leucin und Isoleucin und zu verschiedenen Aminosäurederivaten wurde ebenfalls beobachtet. Um wichtige strukturelle und funktionelle Eigenschaften zu erörtern, wurden multiple Sequenz-Alignments in der Arbeitsgruppe von Prof. Dr. D. Hilvert von allen bekannten PDT Sequenzen aus der Pfam Datenbank [12] erstellt, die aus dem Reich der Archaeen, Bakterien und Eukaryonten kommen [13]. Die Abfolge der konservierten Reste lässt ein für die ganze Familie gemeinsames Faltungsmotiv vermuten. Aus dem vorhandenen Satz an Sequenzen wurde PDT aus *Methanocaldococcus jannaschii* (MjPDT) als repräsentativer Vertreter der monofunkionalen Enzyme aus Archaeen für detailliertere Studien ausgewählt [13]. MjPDT ist ein thermostabiles und hoch aktives Enzym, mit einem Molekulargewicht von 31 kDa. Der berechnete pI-Wert liegt bei 8.8. Eine biophysikalische Charakterisierung ergab, dass PDT als Dimer aktiv ist (62 kDa) und sowohl aus α-Helices, als auch aus β-Faltblättern als Sekundärstruktur-Elementen besteht. MjPDT entfaltet sich bei 94°C und die freie Entfaltungsenergie [$\Delta G_U(H_2O)$] beträgt 32.0 kcal/mol. Die Umwandlung von Prephenat zu Phenylpyruvat folgt der Michaelis-Menten Kinetik ($k_{cat} = 12.3\ s^{-1}$ und $K_M = 22\ \mu M$ bei 30°C), dabei bleibt die Enzymaktivität zwischen pH 5 und pH 10 unverändert. L-Phenylalanin wirkt als Inhibitor, mit einem K_i-Wert von 0.5 µM, L-Tyrosin und L-Tryptophan zeigen dagegen keine Hemmung [13].

8.2 Aufgabenstellung

Die Prephenat Dehydratase stellt ein vielversprechendes Target für die Entwicklung von Antibiotika und Pflanzenschutzmittel dar. Inhibitoren der 5-Enolpyruvyl Shikimat-3-phosphat Synthase, ebenfalls ein Enzym aus dem Shikimat Biosyntheseweg, haben bereits eine Anwendung als Breitbandspektum-Herbizid gefunden: Der Bedarf an *N*-Phosphonomethylglycin, auch bekannt als Roundup®, ist sogar so hoch, dass jährlich Erträge bis zu 700 Millionen Euro erzielt werden.

Durch Anwendung der Methoden des strukturbasierten Designs sollten Moleküle mit hoher Affinität zu PDTs gefunden werden. Für das rationale Design von Hemmstoffen ist jedoch das Vorhandensein einer Röntgenstruktur von essenzieller Bedeutung, da Verbindungen gesucht werden, die sowohl geometrisch als auch chemisch komplementär zur Proteinbindetasche sind.

Da es bis zu Beginn dieser Arbeit keine Kristallstruktur einer Prephenat Dehydratase gab, war die primäre Aufgabe eine Kristallstruktur einer PDT in unkomplexierter und/oder komplexierter Form zu erhalten. Um dies zu erreichen, sollten zunächst geeignete Kristallisationsbedingungen gefunden werden.

8.3 Kristallisation und Datensammlung von MjPDT

*Mj*PDT wurde von A. Kleeb aus der Arbeitsgruppe von Prof. Dr. D. Hilvert (ETH Zürich) hergestellt [13]. Die Überproduktion fand in *E. coli* statt. Die finalen Schritte, der folgenden Protein-Aufreinigung, bestehend aus Kationenaustausch- und Größenausschluss-chromatographie, lieferten Protein in hoher Reinheit, das zur Kristallisation verwendet wurde (Abb. 2).

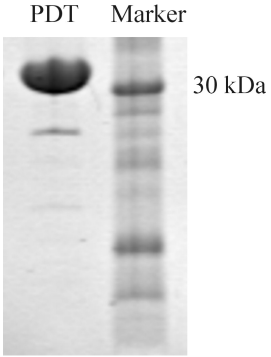


Abb. 2. SDS-Gel mit Banden von aufgereinigter PDT und Marker.

8.3.1 Kristallisationsoptimierung

Erzeugung erster Kristalle

Alle Experimente wurden bei 18°C unter Anwendung der „Sitting Drop“-Methode durchgeführt. Dabei wurden 1 µl Protein-Lösung mit gleichem Volumen entsprechender Reservoir-Lösung zusammenpipettiert. Die Protein-Lösung besaß eine Konzentration von 36 mg/ml in 20 mM Tris/HCl (pH 8.0). Als Reservoir-Lösungen wurden sämtliche Bedingungen des Wizard I und Wizard II Screen von Emerald Biostructures und der Crystal Screen von Hampton Research getestet. Bereits nach einem Tag bildeten sich erste Kristalle mit nadelförmiger Gestalt, die sich von einem gemeinsamen Zentrum aus sternförmig anordneten (Abb. 3 links). Die Reservoirlösung dieses Tropfen bestand aus 20 % (w/v) PEG-1000, 0.1 M Na_2HPO_4/Zitronensäure-Puffer (pH 4.2) und 0.2 M Lithiumsulfat. Mit Kristallen dieser Art konnten bei Röntgenbeugungs-Experimenten Diffraktionsbilder bis ca. 10 Å erhalten werden. Durch Variation der Präzipitanz- und Proteinkonzentration, des pH-Wertes, der Kristallisationstemperatur und der Lithiumsulfat-Konzentration gelang es, die Kristallqualität so zu verbessern (Abb. 3 rechts), dass das Streuvermögen bis auf 6 Å anstieg. Bei den optimierten Bedingungen wurde die Proteinkonzentration auf 4 mg/ml und die PEG-1000 Konzentration auf 16 % gesenkt.

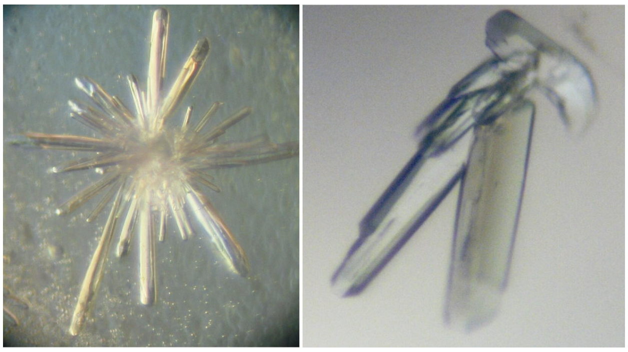


Abbildung 3. Erste Kristalle der PDT (links). Kristalle, die nach erster Optimierung entstanden (rechts).

Entstehung einer zweiten Kristallform

Es wurde weiter versucht, die Kristallqualität zu verbessern, um Datensätze höherer Auflösung zu erhalten, mit denen die *Mj*PDT Struktur gelöst werden kann. Eine neue Proteincharge wurde geliefert, die veränderte Kristallisationseigenschaften aufwies. Bei identischen Bedingungen (4 mg/ml Proteinkonzentration, 16 % (w/v) PEG-1000, 0.1 M Na_2HPO_4/Zitronensäure-Puffer (pH 4.2) und 0.2 M Lithiumsulfat) entstanden bereits innerhalb einer Minute die ersten Kristalle. Interessanterweise wuchsen über Nacht zusätzlich Kristalle mit einer veränderten Kristallform. Während die ursprüngliche Form nadelartig war, besaß die neue Variante eine zweidimensionale, rautenartige Gestalt (Abb. 4). Diffraktionsbilder mit Auflösungen zwischen 3.0 Å und 2.4 Å wurden erhalten.

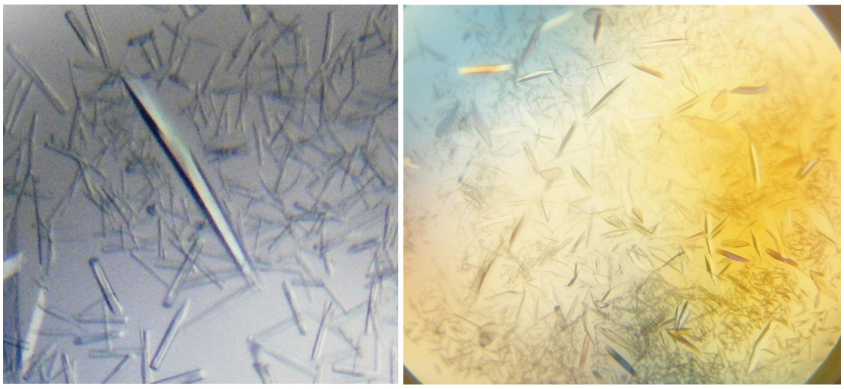


Abbildung 4. 2 unterschiedliche Kristallformen sind in einem Tropfen zu beobachten.

Micro-seeding und Macro-seeding

Bei weiteren Proteinchargen konnte die spontane Enstehung der rautenartigen Kristallform nicht mehr beobachtet werden. Durch „Micro-Seeding“ gelang es aber in allen Ansätzen, die Kristalle mit zweidimensionaler Gestalt zu reproduzieren (Abb. 5 links).

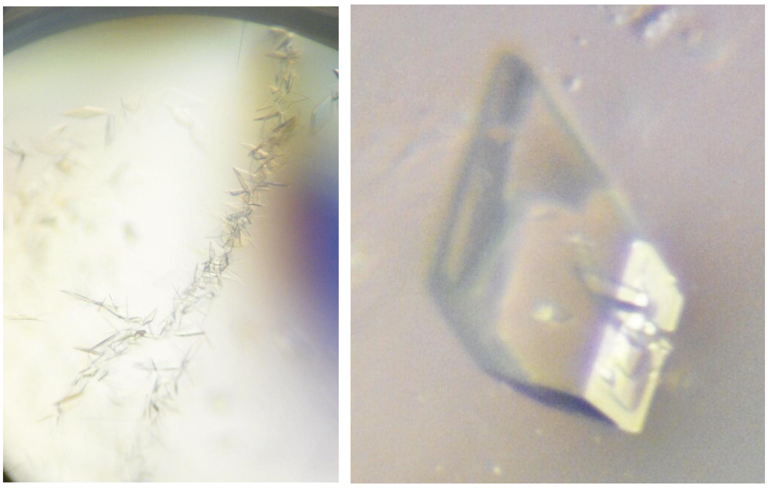


Abb. 5. Seeding Linie von PDT Kristallen mit rautenförmiger Gestalt (links). Kristall erhalten durch Macro-Seeding (rechts).

Makro-Seeding gelang dagegen nur in Einzelfällen. Meistens bildeten sich trotz mehrfachen Waschens der Seeding-Kristalle zusätzlich neue, kleine Keim-Kristalle, was zu Verwachsungen führte (Abb. 5 rechts).

Kokristallisation

Prephenat Dehydratasen katalysieren die Umwandlung von Prephenat zu Phenylpyruvat. Eine Hemmung des Proteins wird durch Phenylalanin erreicht. Um Komplexstrukturen zu erhalten, wurden Kokristallisationsexperimente durchgeführt. Dazu wurde eine PDT-Lösung (5mg/ml) mit 10 % einer wässrigen Phenylpyruvat- bzw. Phenylalanin-Lösung versetzt, die jeweils eine Konzentration von 20 mg/ml besaßen. Kokristallisationsansätze wurden analog zu den oben beschriebenen Bedingungen durchgeführt. Erste Kristalle entstanden nach wenigen Tagen. Nach der Testung von ca. 40 Kristallen, wurden Kristalle gefunden, die bis ca. 2.5 Å streuten.

Optimierung der Cryobedingungen

Beim Versuch geeignete Cryobedingungen zu finden, wurde zunächst beobachtet, dass eine 25 % Glycerin-Lösung zum klaren Gefrieren der Kristalle führt, ohne dass diese optisch beschädigt werden. Ein erster nativer Datensatz konnte bis 2.4 Å gesammelt werden (siehe unten). Um den Einfluß des Cryomittels auf das Streuvermögen der Kristalle zu untersuchen, wurde eine Reihe verschiedener Verbindungen in unterschiedlichen Konzentrationen getestet (z. B. Glycerin, PEG-400, PEG-500, MPD). Die besten Ergebnisse wurden beim schrittweisen Transfer des Kristalls in zunächst 5 % MPD, dann 10 % MPD und schließlich 15 % MPD erzielt. Die Umsetzung der Kristalle in die jeweils höhere MPD-Konzentration erfolgte jeweils nach 30 minütiger Inkubation. Der Wechsel von Glycerin zu MPD machte sich besonders in den Mosaizitäten bemerkbar, die beim Verwenden von MPD deutlich sanken. Zudem streuten die Kristalle bei ähnlichem äußerem Erscheinungsbild um durchschnittlich 0.2 Å bis 0.3 Å besser.

8.3.2 High Throughput Kristallisation

Kristalle konnten bereits gezüchtet werden, von denen Diffraktionsbilder mit einem Beugungsvermögen bis knapp über 2 Å erhalten wurden. Um zu analysieren, ob es andere Kristallisationsbedingungen gibt, die zu noch besser streuenden Kristallen führen, wurde ein Kristallisationsscreen von ca. 2400 Bedingungen mit Hilfe eines Kristallisationsroboters angesetzt. Bei 38 Bedingungen wurde Kristallwachstum mit teilweise sehr unterschiedlichen pH-Werten (zwischen pH 4.0 und pH 10), Präzipitanzien und Additiven beobachtet (Tabelle 1). Obwohl einige Kristalle optisch sehr vielversprechend für Röntgenbeugungs-Experimente aussahen (Abb. 6), zeigten die Streubilder Auflösungen bis nur maximal 8 Å.

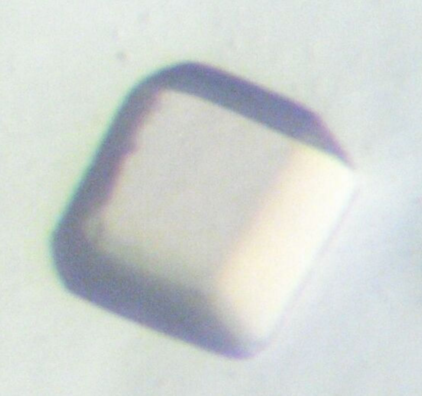


Abb. 6. Kristall von PDT, entstanden beim Kristallisationsscreen von über 2400 Bedingungen.

Präzipitanz	**Puffer**	**c**	**pH**	**Additiv**	**c**	**Additiv**
PEG3350 20%			5,1	$CaCl_2$	0,2 M	
PEG3350 20%			6,9	NaSCN	0,2 M	
PEG3350 20%			7,1	NH_4OAc	0,2 M	
PEG3350 20%			6,4	Li_2SO_4	0,2 M	
PEG3350 20%			5,9	$MgSO_4$	0,2 M	
PEG3350 20%			6,6	Na_2SO_4	0,2 M	
PEG3350 20%			4,5	NaH_2PO_4	0,2 M	
PEG3350 20%			5	$(NH_4)_2$HCitrate	0,2 M	
PEG1500 30%						
PEG3000 20%	HEPES	0,1M	7,5	NaCl	0,2 M	
PEG8000 20%	Na_3PO_4/Citronensäure	0,1M	4,2	NaCl	0,2 M	
PEG1000 20%	Na_3PO_4/Citronensäure	0,1M	4,2	Li_2SO_4	0,2 M	
2-Propanol 10%	MES	0,1M	6	$Ca(OAc)_2$	0,2 M	
PEG3000 20%	NaOAc/HOAc	0,1M	4,5			
PEG3000 30%	Tris/HCl	0,1M	7	NaCl	0,2 M	
2-Propanol 10%	Cacodylat	0,1M	6,5	$Zn(OAc)_2$	0,2 M	
PEG8000 20%	MES	0,1M	6	$Ca(OAc)_2$	0,2 M	
PEG3000 10%	Na_3PO_4/Citronensäure	0,1M	4,2	NaCl	0,2 M	
MPD 40%	Cacodylat	0,1M	6,5	PEG8000 5%		
PEG200 50%	CHES	0,1M	9,5			
PEG200 50%	Tris/HCl	0,1M	7			
PEG200 50%	Cacodylat	0,1M	6,5	$MgCl_2$	0,2 M	
PEG200 40%	NaOAc/HOAc	0,1M	4,5	NaCl	0,2 M	
PEG200 30%	CAPS	0,1M	10,5	$(NH_4)_2SO_4$	0,2 M	
PEG200 50%	Na_3PO_4/Citronensäure	0,1M	4,2	NaCl	0,2 M	
PEG2000 15%	NaOAc/HOAc	0,1M	4,6	$(NH_4)_2SO_4$	0,2 M	
PEG2000 15%	NaOAc/HOAc	0,05 M	4,6	$(NH_4)_2SO_4$	0,1 M	
PEG4000 25%	NaOAc/HOAc	0,1M	4,6	$(NH_4)_2SO_4$	0,2 M	
PEG8000 20%				K_3PO_4	0,05 M	
PEG4000 30%	NaOAc/HOAc	0,1M	4,6	NH_4OAc	0,2 M	
PEG4000 15%	NaOAc/HOAc	0,05M	4,6	NH_4OAc	0,1 M	
PEG4000 15%	Tris/HCl	0,05M	8,5	NaOAc	0,1 M	
K-Na-Tartrat 0,5M	HEPES	0,05M	7,5	K-Na-Tartrat	0,4M	
PEG4000 25%	NaOAc/HOAc	0,085M	4,6	NH_4OAc	0,17M	Glycerol 15%
PEG4000 12%			7,5	$(NH_4)_2SO_4$	0,085M	Glycerol 7,5%
PEG8000 16%				KH_2PO_4	0,04 M	Glycerol 20%
PEG4000 10%	NaOAc/HOAc	0,04M	4,6	$(NH_4)_2SO_4$	0,08 M	Glycerol 10%

Tabelle 1. Kristallisationsbedingungen für PDT, gefunden beim Screen von 2400 unterschiedlichen Bedingungen.

8.3.3 Gesammelte Datensätze

Trotz optimierter Kristallisationsbedingungen und optisch sehr ähnlich aussehender Kristalle streute durchschnittlich nur jeder zwanzigste Kristall besser als 3 Å. Erschwerend kam bei der Datensammlung hinzu, dass das Streuvermögen der Kristalle sich als anisotrop erwies.

Nach zahlreichen Streutests wurden insgesamt ca. 20 Datensätze von unmodifizierter PDT, basierend auf den optimierten Kristallisationsbedingungen gesammelt. Eine Auswahl der am höchsten aufgelösten Datensätze befindet sich in Tabelle 2. Das Enzym kristallisierte in allen Fällen in der Raumgruppe $C222_1$ und die Einheitszelle ist sowohl für PDT in unkomplexierter, als auch in komplexierter Form nahezu isomorph (Tabelle 2). Alle Daten wurden mit HKL2000 prozessiert [14].

	PDT_nativ_1	**PDT_nativ_2**
Raumgruppe	$C222_1$	$C222_1$
Einheitszelle [Å] **[°]**	55.8 115.2 188.6 90 90 90	55.3 114.5 188.5 90 90 90
Gesamtzahl der Reflexe	187233	104887
Anzahl einzigartiger Reflexe	24295	29211
Vollständigkeit [%]	99.8	94.5
Redundanz	7.7	3.6
Auflösung	2.40	2.20
R_{sym}	0.13	0.069
Mosaizität	0.63	0.58

	PDT/Phenylalanin	**PDT/Phenylpyruvat**
Raumgruppe	$C222_1$	$C222_1$
Einheitszelle [Å] **[°]**	55.3 114.4 187.8 90 90 90	55.9 115.0 188.2 90 90 90
Gesamtzahl der Reflexe	85794	73587
Anzahl einzigartiger Reflexe	22661	22203
Vollständigkeit [%]	94.2	97.4
Redundanz	3.8	3.3
Auflösung	2.38	2.45
R_{sym}	0.08	0.06
Mosaizität	0.443	0.643

Tabelle 2. Auswahl der am höchsten aufgelösten Datensätze von PDT in nativer, als auch komplexierter Form.

8.4 Strukturlösungsversuche durch Molekularen Ersatz

8.4.1 Sequenzalignment

Zu Beginn des Jahres 2007 wurden die Koordinaten einer potenziellen Prephenat Dehydratase aus dem Organismus *Staphylcoccus aureus subsp. aureus Mu50* in der PDB freigegeben. Bis zu diesem Zeitpunkt waren keine Proteinstrukturen bekannt, die dazu geeignet gewesen wären, ein Homologie-Modell zu erstellen, mit dem die PDT Struktur durch Molekularen Ersatz hätte gelöst werden können. Ein Sequenzalignment zeigt (Abb. 7), dass die hinterlegte Struktur mit dem PDB-Code 2IQ8 eine Sequenzidentität von 32 % zur PDT aus *Methanocaldococcus jannaschii* besitzt. Darüber hinaus werden insgesamt 55 % der Aminosäuren im Sequenzvergleich als ähnlich eingestuft. Diese Daten liegen in einem Bereich, bei dem die Strukturlösung von *Mj*PDT durch Molekularen Ersatz möglich sein könnte.

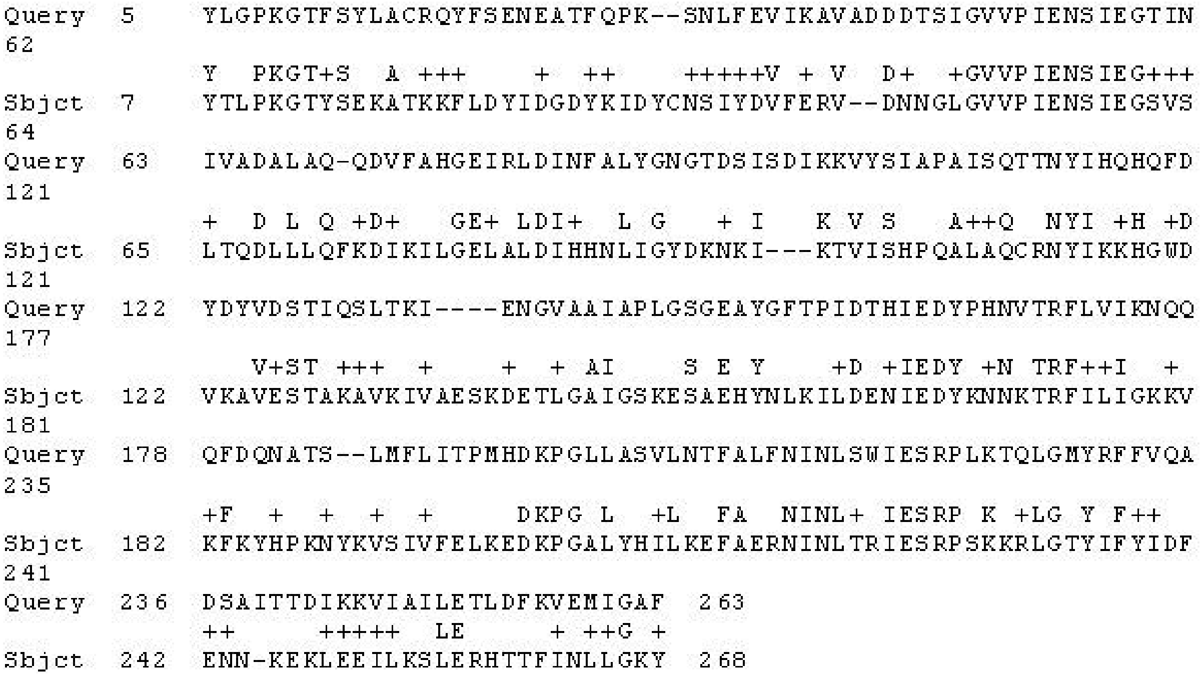
Score = 130 bits (326)

Expect = 1e-28

Identities = 87/268 (32 %)

Positives = 149/268 (55 %)

Gaps = 15/268 (5 %)



```
Query  5     YLGPKGTFSYLACRQYFSENEATFQPK--SNLFEVIKAVADDDTSIGVVPIENSIEGTIN
62
             Y  PKGT+S  A +++    +  ++    +++++V + V  D+  +GVVPIENSIEG+++
Sbjct  7     YTLPKGTYSEKATKKFLDYIDGDYKIDYCNSIYDVFERV--DNNGLGVVPIENSIEGSVS
64
Query  63    IVADALAQ-QDVFAHGEIRLDINFALYGNGTDSISDIKKVYSIAPAISQTTNYIHQHQFD
121
             +  D L Q +D+   GE+ LDI+  L G   + I   K V S   A++Q  NYI +H +D
Sbjct  65    LTQDLLLQFKDIKILGELALDIHHNLIGYDKNKI---KTVISHPQALAQCRNYIKKHGWD
121
Query  122   YDYVDSTIQSLTKI----ENGVAAIAPLGSGEAYGFTPIDTHIEDYPHNVTRFLVIKNQQ
177
                V+ST +++  +    +  + AI    S E Y    +D +IEDY +N TRF++I  +
Sbjct  122   VKAVESTAKAVKIVAESKDETLGAIGSKESAEHYNLKILDENIEDYKNNKTRFILIGKKV
181
Query  178   QFDQNATS--LMFLITPMHDKPGLLASVLNTFALFNINLSWIESRPLKTQLGMYRFFVQA
235
             +F  +  +  +  +     DKPG L  +L  FA  NINL+ IESRP K +LG Y F++
Sbjct  182   KFKYHPKNYKVSIVFELKEDKPGALYHILKEFAERNINLTRIESRPSKKRLGTYIFYIDF
241
Query  236   DSAITTDIKKVIAILETLDFKVEMIGAF  263
             ++     +++++  LE     + ++G +
Sbjct  242   ENN-KEKLEEILKSLERHTTFINLLGKY  268
```

Abbildung 7. Sequenzalignment durchgeführt mit dem Programm BLAST. *Mj*PDT (Query) und PDT aus *Staphylcoccus aureus subsp. aureus Mu50* (Sbjct).

8.4.2 Modelle für den Molekularen Ersatz

Die Struktur (2IQ8) besteht wie *Mj*PDT aus einer PDT- und einer ACT-Domäne (Abb.8). ACT-Domänen haben eine regulatorische Funktion und kommen bei vielen am Metabolismus beteiligten Enzymen vor. Sie können in der Regel durch Konzentrationsunterschiede verschiedener Aminosäuren reguliert werden, bei PDT durch L-Phenylalanin.

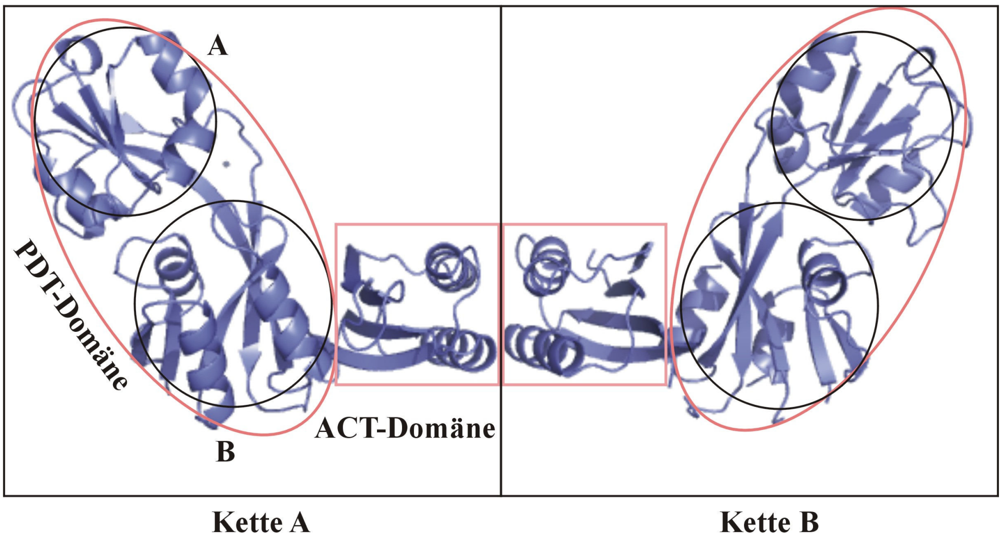


Abbildung 8. PDT mit dem PDB-Code 2IQ8 dargestellt als Cartoon. Die zwei Ketten bestehen jeweils aus einer PDT- und einer ACT-Domäne.

Um Ausgangsphasen für *Mj*PDT zu erhalten, wurden verschiedene Modelle ausgehend von der Struktur 2IQ8 erzeugt. Bei den Modellen handelte es sich einerseits um Polyalanin-Modelle aber auch um Homologiemodelle erzeugt durch MOE [15], MODELLER [16, 17] und SWISS MODEL [18, 19]. Eine genaue Übersicht über die generierten Modelle zur Erzeugung der Ausgangsphasen befindet sich in Tabelle 3. Auch wenn sich die RMSD-Werte, der mit unterschiedlichen Programmen erzeugten Modelle, im Bereich der α-Helices und β-Faltblätter nur wenig unterschieden, wurden alle Modelle getestet, da sich größere Abweichungen in den Loop-Regionen zeigten.

	Polyalanin-Modell	**Homologiemodelle**		
		MOE	**MODELLER**	**SWISS MODEL**
Kette A und Kette B	X			
Kette A	X	X	X	X
PDT-Domäne	X	X	X	X
ACT-Domäne	X	X	X	X
Teil A der PDT-Domäne	X	X	X	X
Teil B der PDT-Domäne	X	X	X	X

Tabelle 3. Generierte Modelle, die zur Erzeugung von Ausgangsphasen für *Mj*PDT dienen sollten.

8.4.3 Molekularer Ersatz

Für die Phasenbestimmung der Prephenat Dehydratase mit Hilfe des Molekularen Ersatzes wurden die oben beschriebenen Startmodelle (Tab. 3) und die oben angegebenen Datensätze (Tab. 2) verwendet. Zur Anwendung kamen dabei die Programme Phaser [20], MOLREP [21] und CNS [22]. Rotations- und Translationssuchen mit Phaser, bei der die ACT-Domäne als Startstruktur verwendet wurde, zeigten zunächst zuversichtlich stimmende Ergebnisse. Lösungen mit einem LLG von 70 und einem Z-Score von 7 setzten sich von den darauffolgenden Lösungen ab und erste Dichten zeigten bei einem Sigma-Level von 2 zusammenhängende Dichten von bis zu 50 Aminosäuren (Abb. 9). Durch weitere Refinement-Schritte mit CNS konnte diese Elektronendichte jedoch nicht verbessert werden, die R-Werte wurden schlechter (>50 %). Durch die Methode des Molekularen Ersatzes konnte folglich nicht die richtige Lösung gefunden werden.

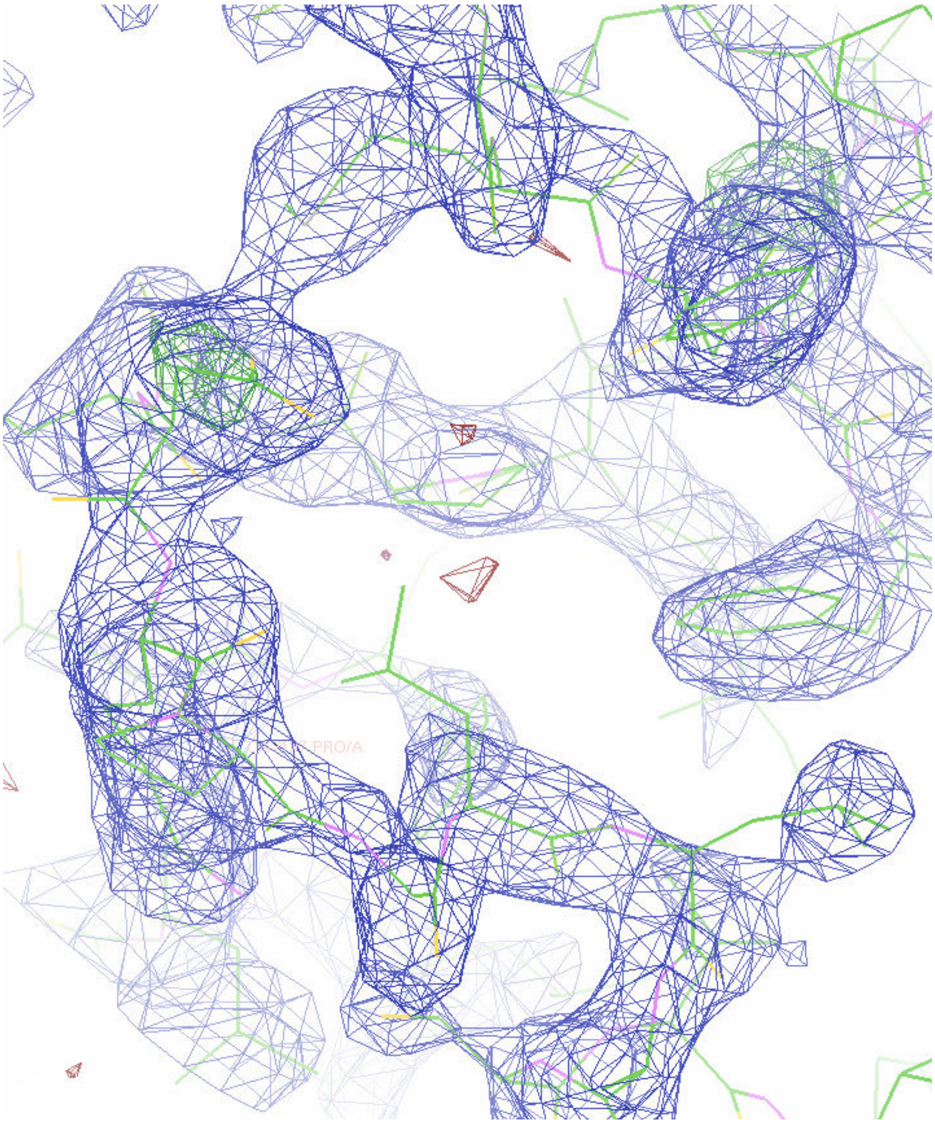


Abbildung 9. Dichte erzeugt mit Phaser. Ein Homologiemodell der ACT-Domäne diente hierbei als Startmodell.

8.5 Experimentelle Phasen via SeMet

8.5.1 Einfügen von Methioninen durch Mutation

Da die für die Kristallisation eingesetzte Prephenat Dehydratase keine internen Methionine besitzt, wurden von den Kooperationspartnern durch gezielte Mutagenese drei Methionine eingefügt. Um später die Positionen der Selenomethionine lokalisieren zu können, sollten die Mutationen in rigiden Bereichen des Proteins erfolgen. Es wurde zunächst eine Sekundärstruktur-Vorhersage durchgeführt und nur zu mutierende Aminosäuren aus nicht flexiblen Bereichen ausgewählt. Die Dreifach-Mutante V52M/I91M/H118M zeigte sich im Assay unverändert aktiv und behielt ihre Thermostabilität. Die mutierte PDT wurde als Selenomethionin-Derivat überproduziert und dann zur Kristallisation eingesetzt.

Für den Fall, dass die Strukturlösung mit der Dreifachmutante nicht funktioniert, wurden nach dem gleichen Prinzip zwei unterschiedliche Fünffachmutanten geplant:
V52M/L69M/I91M/H118M/A136M
V52M/I91M/H118M/V195M/I209M

8.5.2 Kristallisationsoptimierung

Die für unmodifizierte PDT optimierten Kristallisationsbedingungen dienten als Ausgangspunkt für einen Screen, um Kristallisationsbedingungen für die 3-Se-Met-PDT-Variante zu finden (4 mg/ml Proteinkonzentration, 16 % (w/v) PEG-1000, 0.1 M Na_2HPO_4/Zitronensäure-Puffer (pH 4.2) und 0.2 M Lithiumsulfat). Über Nacht wuchsen erste Kristalle mit nadelförmiger Gestalt. Im Gegensatz zu den nativen Ansätzen, ließ sich die rautenförmige Kristallform nicht durch Micro-Seeding erzeugen. Röntgenbeugungsexperimente mit den PDT-Nadeln bestätigten das schwache Streuvermögen dieser Kristallform (~ 10 Å). Alle Additive des Hampton Additive Screen wurden ausgetestet, um Kristalle mit größerem Volumen zu erhalten, was jedoch nicht gelang.

Analoge Kristallisations-Experimente wurden mit einer Protein-Charge höherer Reinheit durchgeführt. Aber interessanterweise führte erst das Mischen der ersten mit der zweiten Proteincharge im Verhältnis 2 zu 3 unter den oben beschriebenen Kristallisationsbedingungen (inklusive Micro-Seeding) in Einzelfällen zu Kristallen mit rautenförmiger Gestalt. Diese Kristalle zeigten wiederum in Einzelfällen ein Streuvermögen von ~ 2-3 Å.

Zu einem späteren Zeitpunkt wurde versucht die beiden 5-Se-Met-PDT-Varianten zu kristallisieren. Neben den nicht zur Strukturlösung geeigneten Nadeln wuchs ein einziger Kristall in rautenförmiger Gestalt, der geeignet zur Datensammlung schien.

8.5.3 Gesammelte Datensätze

Es gelang einen Peak-, Inflection-Point und Remote-Datensatz der V52M/I91M/H118M-PDT-Mutante zu sammeln (Tabelle 4, 3-Se-Met-PDT_1). Da der Kristall bei der Bestrahlung beschädigt wurde, konnten keine vollständig kompletten Datensätze aufgenommen werden. Da zusätzlich die R_{sym}-Werte recht hoch waren, wurden erneut drei Datensätze eines Kristalls, bestehend aus der 3-Se-Met-PDT-Variante gesammelt (Tabelle 4, 3-Se-Met-PDT_2). Interessanterweise wechselte die Raumgruppe beim zweiten Kristall von $P2_12_12$ zu $P2_1$.

Nur ein einziger, gut streuender Kristall wurde von der Fünffach-Mutante V52M/L69M/I91M/H118M/A136M erhalten, von dem ebenfalls drei Datensätze gesammelt wurden (Tabelle 4, 5-Se-Met-PDT).

3-Se-Met-PDT_1	**Peak**	**Inflection**	**Remote**
Wellenlänge [Å]	0.9774	0.9789	0.9338
Raumgruppe	$P2_12_12$		
Einheitszelle [Å] **[°]**	55.4 113.3 190.6 90 90 90		
Gesamtzahl der Reflexe	66942	69038	68038
Anzahl einzigartiger Reflexe	29579	20087	20087
Vollständigkeit [%]	77.7	79.0	79.0
Redundanz	2.3	3.4	3.4
Auflösung	2.6	2.6	3.0
R_{sym}	0.075	0.124	0.124
Mosaizität	0.992	0.903	0.903

3-Se-Met-PDT_2	**Peak**	**Inflection**	**Remote**
Wellenlänge [Å]	0.9800	0.9800	0.9401
Raumgruppe	$P2_1$		
Einheitszelle [Å] **[°]**	56.4 197.5 61.6 90.0 113.3 90.0		
Gesamtzahl der Reflexe	132166	107893	103919
Anzahl einzigartiger Reflexe	29763	27731	29387
Vollständigkeit [%]	97.1	91.7	98.1
Redundanz	4.4	3.9	3.5
Auflösung	2.8	2.8	2.8
R_{sym}	0.096	0.072	0.073
Mosaizität	0.282	0.314	0.364

5-Se-Met-PDT	**Peak**	**Inflection**	**Remote**
Wellenlänge [Å]	0.9797	0.9798	0.9343
Raumgruppe	$P2_1$		
Einheitszelle [Å] **[°]**	63.0 51.5 93.7 90.0 102.5 90.0		
Gesamtzahl der Reflexe	43809	43342	34284
Anzahl einzigartiger Reflexe	12582	12696	10434
Vollständigkeit [%]	99.2	99.7	99.5
Redundanz	3.5	3.4	3.3
Auflösung	2.95	2.95	3.15
R_{sym}	0.124	0.112	0.142
Mosaizität	2.000	1.796	1.752

Tabelle 4. Peak-, Inflection- und Remote-Datensätze von modifiziertem PDT: Drei bzw. fünf Methionine wurden durch Mutagenese eingefügt. Die Überproduktion fand als Selenomethionin-Derivat statt.

8.5.4 Strukturlösungsversuche mit HKL2map

Für alle gesammelten Datensätze konnte die Inkorporation von Selenomethionin durch Absorptionsscans nachgewiesen werden. Im Folgenden wurde zur Lösung der Struktur durch die Methode der Anomalen Streuung das Programm HKL2MAP [23] eingesetzt.

Die Qualität des Anomalen Signals wurde mit SHELXC [24] abgeschätzt. Für alle drei Sets gesammelter Datensätze wurde bis ca. 5 Å ein anomaler Korrelationskoeffizient größer als 30 % gefunden, was zur Strukturlösung ausreichen könnte, aber als grenzwertig zu bewerten ist.

Mit SHELXD [25] wurden anschließend die Selenpositionen gesucht. Mit Hilfe des Matthews-Koeffizienten wurde die Anzahl der zu erwartenden Selene in der asymmetrischen Einheit berechnet.

Bei SHELXD gibt es die Möglichkeit, folgende Parameter zu variieren: Zum einen wie viele Suchdurchläufe durchgeführt werden sollen, um die Schweratompositionen zu finden und zum anderen bis zu welcher Auflösung die Daten bei der Suche verwendet werden sollen. Die Anzahl der Suchdurchläufe wurde von 100 bis 20000 variiert, die Auflösung von 5.1 Å bis 3.0 Å in 0.1 Å Schritten. Für 3-Se-Met-PDT_1 und 5-Se-Met-PDT konnten jedoch die Positionen nicht gefunden werden.

Vielversprechend sah ein Ergebnis für die 3-Se-Met-PDT_2-Datensätze aus. 4 Protein-Moleküle wurden in der asymmetrischen Einheit durch den Matthews-Koeffizienten vorausgesagt. Vier Selenomethionine liegen in der PDT vor (das N-terminale Selenomethionin eingeschlossen), so dass maximal 16 Selene in der asymmetrischen Einheit gefunden werden könnten. Für 12 dieser Selene wurden mit einer signifikanten Wahrscheinlichkeit Positionen mit SHELXD [25] gefunden (Abb. 10), bei denen sich herausstellte, dass die Positionen untereinander symmetrieverwandt sind.

Phasierung mit SHELXE [25] führte jedoch weder zu hohen Korrelationskoeffizienten noch zu hohem Dichteunterschied der beiden Enantiomorphe, was darauf hinwies, dass die Struktur noch nicht gelöst wurde. Bestätigt wurde dieses Ergebnis durch eine visuelle Betrachtung der für die potenziellen Lösungen erzeugten Elektronendichten.

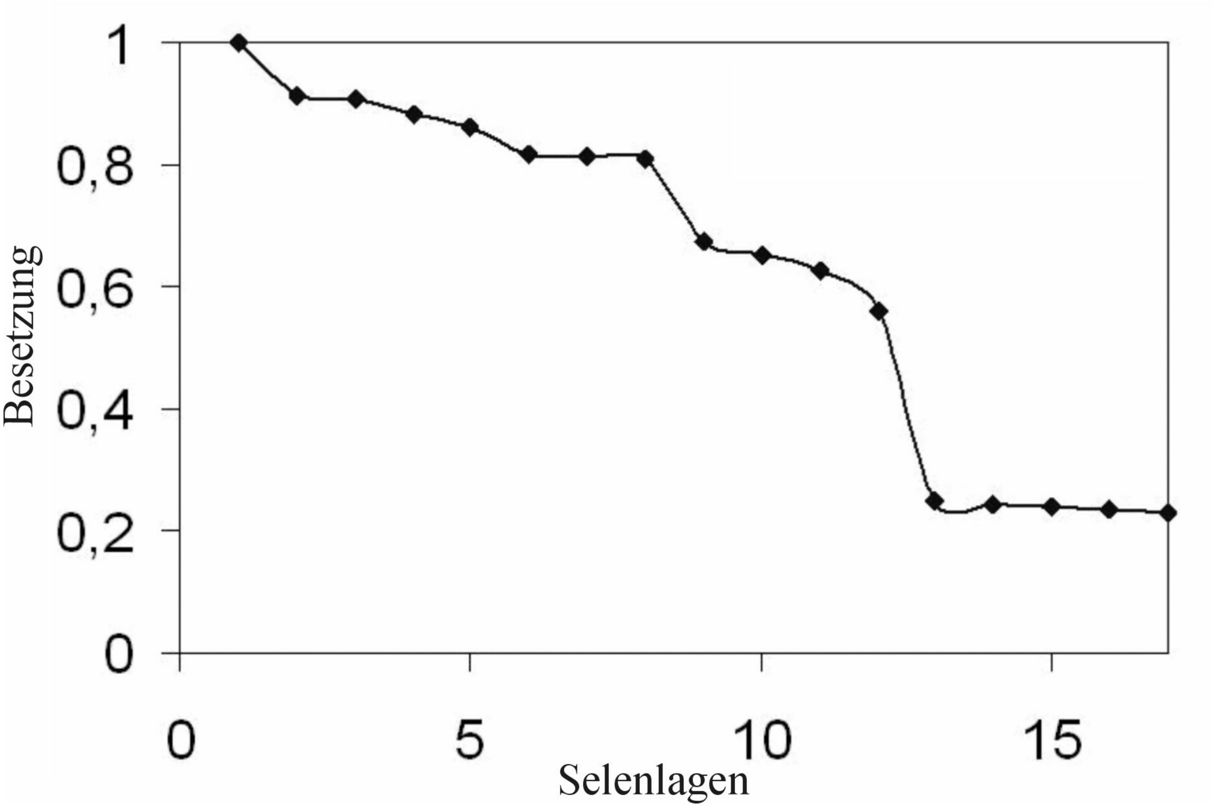


Abbildung 10. Darstellung der Selenlagen mit zugehöriger Besetzungszahl. 12 Selenlagen weisen eine hohe Bestetzungszahl vor.

8.6 Experimentelle Phasen via Schweratomderivate

8.6.1 Geeignete Schweratomderivate

Um geeignete Schweratomderivate zu finden, wurde der Heavy Atom Kit von Hampton Research eingesetzt. Dabei wurden insbesonders Schweratomverbindungen berücksichtigt, die nicht mit Bestandteilen der optimierten Kristallisationbedingung reagieren können. Da in der Reservoirlösung Sulfat präsent ist, waren die meisten Quecksilber- und Blei-Salze für die Anwendung uninteressant, da sie schwerlösliche Verbindungen mit Sulfat eingehen würden. Da beim Einstellen des pH-Wertes Salzsäure verwendet wurde, befinden sich auch Chlorid-Ionen im Puffer, die mit Silberionen zu Silberchlorid reagieren und ausfallen würden. Viele weitere Substanzen sind nicht kompatibel mit den sauren Bedingungen der Kristallisationslösung wie zum Beispiel einige Platin- und Cadmiumverbindungen.

Aus den oben genannten Gründen wurden vorwiegend Schweratom-Soaking-Experimente mit Seltenerd-Metall-Ionen, Goldverbindungen und einigen Platinkomplexen durchgeführt.

8.6.2 Soakingbedingung-Optimierung

Für alle Soaking-Experimente wurden immer frische Schweratom-Stammlösungen mit einer Konzentration von 100 mM in einem 0.01 M Na_2HPO_4/Zitronensäure-Puffer hergestellt. Jeweils 1 µl der Stammlösung wurde anschließend mit 9 µl Reservoirlösung vermischt. In die so entstandenen Soaking-Lösungen mit finalen Schweratomkonzentrationen von 10 mM, wurden Kristalle für 10 min. überführt (Quick-Soaking). Um zu erreichen, dass bei späteren Absorptionsscans nur spezifisch gebundene Schweratome detektiert werden, wurden im Anschluss die Kristalle in eine Reservoirlösung überführt, in der sich keine Schweratome mehr befanden (Back-Soaking).

Das Überführen der Kristalle in die Schweratomlösungen hatte zur Folge, dass sehr viele Kristalle sich auflösten, Risse entstanden oder ohne sichtbaren optischen Schaden an Streuvermögen verloren. Aus diesem Grund mussten zahlreiche Kristalle getestet werden, um daraus wenige zu erhalten, die zumindest bis 5.0 Å streuten.

Alle durchgeführten Soaking-Experimente führten jedoch nicht zu geeigneten Derivaten, die für die Strukturlösung durch IR, SIRAS, SAD oder MAD geeignet gewesen wären.

8.6.3 Druckbegasung mit Xenon

Bei einem weiteren Versuch, ein geeignetes Schweratomderivat zu finden, wurden PDT-Kristalle zehn Minuten mit Xenon unter hohem Druck begast. Ein Datensatz mit einer Auflösung von 2.65 Å wurde gesammelt (Tabelle 5). Auch mit diesen Daten konnte die Struktur von *Mj*PDT nicht über Isomorphen Ersatz gelöst werden.

	PDT_Xe_10-min
Raumgruppe	$C222_1$
Einheitszelle [Å] [°]	55.7 114.9 189.0 90 90 90
Gesamtzahl der Reflexe	56410
Anzahl einzigartiger Reflexe	17918
Vollständigkeit [%]	97.8
Redundanz	3.1
Auflösung	2.65
R_{sym}	0.066
Mosaizität	0.28

Tabelle 5. Datensatz von einem mit Xenon begasten PDT-Kristall.

8.7 Zusammenfassung und Ausblick

Es ist gelungen Kristallisationsbedingungen für *Mj*PDT zu finden und zu optimieren, so dass ein Datensatz bis zu einer Auflösung von 2.2 Å gesammelt werden konnte. Bis Ende 2006 gab es in der PDB keine Kristallstrukturen, die als Modell für die Strukturlösung der PDT durch Molekularen Ersatz geeignet gewesen wären. Anfang 2007 wurden die Koordinaten eines Proteins mit 32 %iger Sequenzidentität zu *Mj*PDT in der PDB hinterlegt. Diese Struktur diente als Vorlage zur Generierung verschiedener *Mj*PDT-Homologiemodelle, mit denen jedoch die Strukturlösung durch Molekularen Ersatz ebenfalls nicht gelang.

Bei der Strukturlösung durch Isomorphen Ersatz oder durch Anomale Streuung ist es notwendig, Schweratome in das Protein einzufügen. Eine Standardmethode ist der Einbau von Selenomethionin anstelle von Methionin. Da *Mj*PDT nur über ein N-terminales Methionin verfügt, wurden durch gerichtete Mutation drei bzw. fünf Aminosäuren durch Methionin ersetzt. Absorptionsscans bestätigten, dass die Überproduktion der mutierten PDT als Selenomethionin-Derivat erfolgreich war. Bei der Auswertung der gesammelten Datensätze wurde ein schwaches anomales Signal festgestellt, welches nicht ausreichte, das Phasenproblem zu lösen. Zahlreiche Versuche weitere geeignete Schweratomderivate zu finden, waren nicht erfolgreich.

Um die Struktur dennoch lösen zu können, müsste weiter nach geeigneten Schweratomderivaten gesucht werden. Ein kürzlich durchgeführter Absorptionsscan an einem Kristall, der mit $(NH_4)_2WS_4$ gesoakt wurde, verlief positiv. Da die Streubilder keine ausreichende Qualität zur Datensammlung besassen, konnte nicht verifiziert werden, ob die Strukturlösung möglich gewesen wäre.

Weitere Hoffnungen liegen in der Bestimmung der Phasen durch Molekularen Ersatz, was gelingen könnte, sobald zur Modellbildung zusätzliche Strukturen in der PDB hinterlegt würden.

8.8 Referenzen

1. Kienhöfer A. Mechanistic studies of a Chorismate Mutase from Bacillus subtilis. Chemistry. Zürich: ETH Zürich, 2005:193.
2. Prakash P, Pathak N, Hasnain SE. pheA (Rv3838c) of Mycobacterium tuberculosis encodes an allosterically regulated monofunctional prephenate dehydratase that requires both catalytic and regulatory domains for optimum activity. J Biol Chem 2005;280(21):20666-71.
3. Parish T, Stoker NG. The common aromatic amino acid biosynthesis pathway is essential in Mycobacterium tuberculosis. Microbiology 2002;148(Pt 10):3069-77.
4. Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol 2003;48(1):77-84.
5. Gordhan BG, Smith DA, Alderton H, et al. Construction and phenotypic characterization of an auxotrophic mutant of Mycobacterium tuberculosis defective in L-arginine biosynthesis. Infect Immun 2002;70(6):3080-4.
6. Knaggs AR. The biosynthesis of shikimate metabolites. Nat Prod Rep 2001;18(3):334-55.
7. Zhang S, Pohnert G, Kongsaeree P, et al. Chorismate mutase-prephenate dehydratase from Escherichia coli. Study of catalytic and regulatory domains using genetically engineered proteins. J Biol Chem 1998;273(11):6248-53.
8. Zhang S, Wilson DB, Ganem B. Probing the catalytic mechanism of prephenate dehydratase by site-directed mutagenesis of the Escherichia coli P-protein dehydratase domain. Biochemistry 2000;39(16):4722-8.
9. Fischer R, Jensen R. Prephenate dehydratase (monofunctional). Methods Enzymol 1987;142:507-12.
10. Davidson BE. Chorismate mutase-prephenate dehydratase from Escherichia coli. Methods Enzymol 1987;142:432-9.
11. Aravind L, Koonin EV. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. J Mol Biol 1999;287(5):1023-40.
12. Bateman A, Coin L, Durbin R, et al. The Pfam protein families database. Nucleic Acids Res 2004;32(Database issue):D138-41.
13. Kleeb AC, Kast P, Hilvert D. A monofunctional and thermostable prephenate dehydratase from the archaeon Methanocaldococcus jannaschii. Biochemistry 2006;45(47):14101-10.
14. Otwinowski Z, Minor W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology 1997;276:307-326.
15. MOE. Montreal: Chemical Computing Group.
16. Sali A, Blundell TL. Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. J Mol Biol 1990;212(2):403-28.
17. Sali A, Overington JP, Johnson MS, et al. From comparisons of protein sequences and structures to protein modelling and design. Trends Biochem Sci 1990;15(6):235-40.
18. Schwede T, Kopp J, Guex N, et al. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res 2003;31(13):3381-5.
19. Arnold K, Bordoli L, Kopp J, et al. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006;22(2):195-201.
20. McCoy AJ, Grosse-Kunstleve RW, Adams PD, et al. Phaser crystallographic software. J. Appl. Cryst. 2007;40:658-674.

21. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. J. Appl. Cryst. 1997;30:1022-1025.
22. Brünger AT, Adams PD, Clore GM, et al. Crystallography and nmr system (cns): A new software for macromolecular structure determination. Acta Cryst. 1998;D54:905-921.
23. Pape T, Schneider TR. HKL2MAP: a graphical user interface for phasing with SHELX programs. J. Appl. Cryst. 2004;37:843-844.
24. Sheldrick GM. Göttingen University, 2003.
25. Sheldrick GM, Schneider TR. Methods Enzymol 1997;277:319-343.

9 Zusammenfassung

Malaria stellt eine der weltweit am weitesten verbreiteten Krankheiten dar und ist nach Tuberkulose und AIDS die bedrohlichste aller Infektionskrankheiten. Heute gibt es eine Reihe von Medikamenten, die auf unterschiedliche Weise gegen Malaria wirken. Durch aufkommende Resistenzentwicklungen in vielen Regionen gegenüber zahlreichen Wirkstoffen, ist es jedoch notwendig geworden, neue Arzneimittel zu entwickeln. Ein alternativer Ansatz zur Entwicklung chemisch verwandter Verbindungen zu den existierenden Wirkstoffen, besteht aus der Identifizierung neuer Drug Targets und dem anschließenden Design potenzieller Inhibitoren. Durch die Sequenzierung des *Plasmodium falciparum* Genoms wurden einige potentielle Zielstrukturen entdeckt, unter denen sich die Plasmepsine befinden, die Hauptgegenstand dieser Arbeit waren.

Während der erythrozytären Phase baut *P. falciparum* einen Großteil des in der Wirtszelle vorhandenen Hämoglobins ab, um den eigenen Stoffwechselbedarf an Aminosäuren zu decken. Der Abbau erfolgt entlang einer Enzymkaskade, bei der die Plasmepsine, die zu den Aspartylproteasen gehören, beteiligt sind. Studien mit Proteaseinhibitoren konnten belegen, dass der Hämoglobinabbau für das Überleben des Parasiten essentiell ist.

Zu Beginn der Arbeit waren bereits eine Reihe potenter, peptidomimetischer Inhibitoren und eine Kokristallstruktur von Plasmepsin II in Komplex mit Pepstatin A in der Literatur beschrieben. Um einen Beitrag im Bereich des strukturbasierten Designs von Plasmepsin Inhibitoren leisten zu können, wurden zunächst die erforderlichen Rahmenbedingungen geschaffen. Dazu gehörte die Etablierung eines Expressionssystems für Plasmepsin II und IV, die im Rahmen einer vorausgegangenen Diplomarbeit erreicht wurde. Während der Doktorarbeit ist es gelungen einen Assay zu entwickeln, der die Gesetze der Michaelis-Menten Kinetik befolgt. Weiterhin wurde versucht, ein funktionsfähiges Kristallisationssystem für Plasmespin II zu entwickeln, das die Aufklärung von Protein-Ligand-Bindungsmoden mittels Röntgenbeugungsexperimenten ermöglichen sollte. Obwohl die Reproduktion der in der Literatur bekannten Kristallstruktur mit Pepstatin A gelang, erwies es sich im Verlauf der Arbeit als schwierig, Kristalle des Zielproteins in Komplex mit neuartigen Inhibitoren zu erhalten. Die Qualität dieser Kristalle stellte sich bezüglich ihres Streuvermögens zudem als sehr niedrig heraus, so dass eine Strukturlösung mit weiteren Inhibitoren nicht gelang.

Nach Etablierung der für das strukturbasierte Liganden-Design erforderlichen Methoden galt es in Zusammenarbeit mit synthetisch arbeitenden Gruppen Inhibitoren mit neuartigen, nichtpeptidomimetischen Eigenschaften zu entwerfen und zu optimieren. Erste Modelling-Studien ergaben, dass 2,3,4,7-Tetrahydro-1*H*-Azepin ein geeignetes Grundgerüst für die

Entwicklung von Plasmepsin II und IV Inhibitoren darstellen könnte (Abbildung 1). Dabei wird angenommen, dass über das protonierte basische Stickstoffatom des siebengliedrigen Rings Wasserstoffbrückenbindungen zu den katalytischen Aspartaten ausgebildet werden.

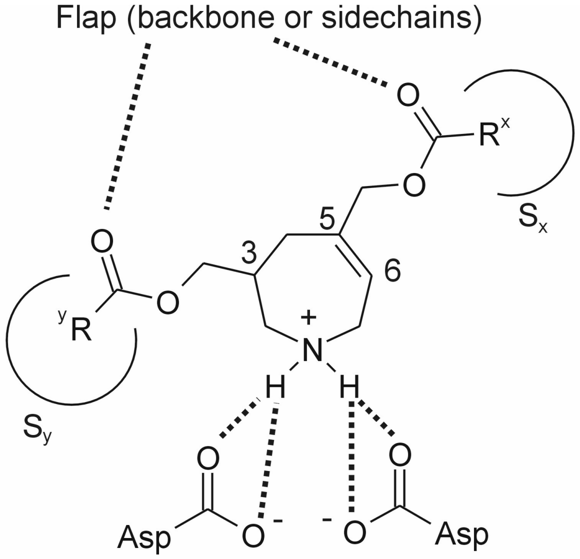


Abbildung 1. 3,5-Bis-carbonyloxymethyl-substituiertes Azepin-Grundgerüst mit seinen vorhergesagten Wasserstoffbrückenbindungen zu Plasmepsin II

Die Ausstattung des Liganden mit geeigneten Linkergruppen und die Einführung von maßgeschneiderten Substituenten in Position 3 und 5 sollte es ermöglichen, die entsprechenden Substrat-Erkennungstaschen der beiden Enzyme Plasmepsin II und IV zu adressieren. Die von den Kooperationspartnern verwendete Synthesestrategie führt zu einem razemischen Gemisch. Um eine genauere Vorstellung über einen mutmaßlichen Bindungsmodus der beiden Enantiomere zu erhalten, wurde Docking mit dem Programm FlexX durchgeführt. Hierbei stellte sich heraus, dass das *R*-Enantiomer von der FlexX-Scoring-Funktion besser bewertet wird als das *S*-Enantiomer. Ein Vergleich der Docking Pose des Azepin-Derivats mit einer Kristallstruktur von Plasmepsin II in Komplex mit einem Hydroxypropylamin-Derivat (PDB-Code: 1LF2) zeigte, dass bei beiden Liganden ein ähnliches Wasserstoffbrückenbindungsmuster erhalten wird. Dies wurde als weiterer Hinweis, dass substituierte Azepin-Derivate potente Plasmepsin Inhibitoren darstellen können, interpretiert (Abbildung 2).

Da nicht nur Inhibitoren für Plasmepsin II entwickelt werden sollten, sondern Verbindungen, die gleichzeitig Plasmepsin II und IV hemmen, wurde nach einer in Hinsicht auf das strukturbasierte Design von Azepinderivaten geeigneten Plasmepsin IV Struktur gesucht. In der Literatur sind bisher keine Strukturen beschrieben, die für diesen Zweck brauchbar wären. Aufgrund der hohen Sequenzidentität von Plasmepsin IV zu Plasmepsin II konnte jedoch ein Homologiemodell für Plasmepsin IV erstellt werden. Eine anschließende Überlagerung der beiden Plasmepsine verdeutlichte, dass das Design von Inhibitoren, die an beide Enzyme binden, grundsätzlich möglich sein sollte.

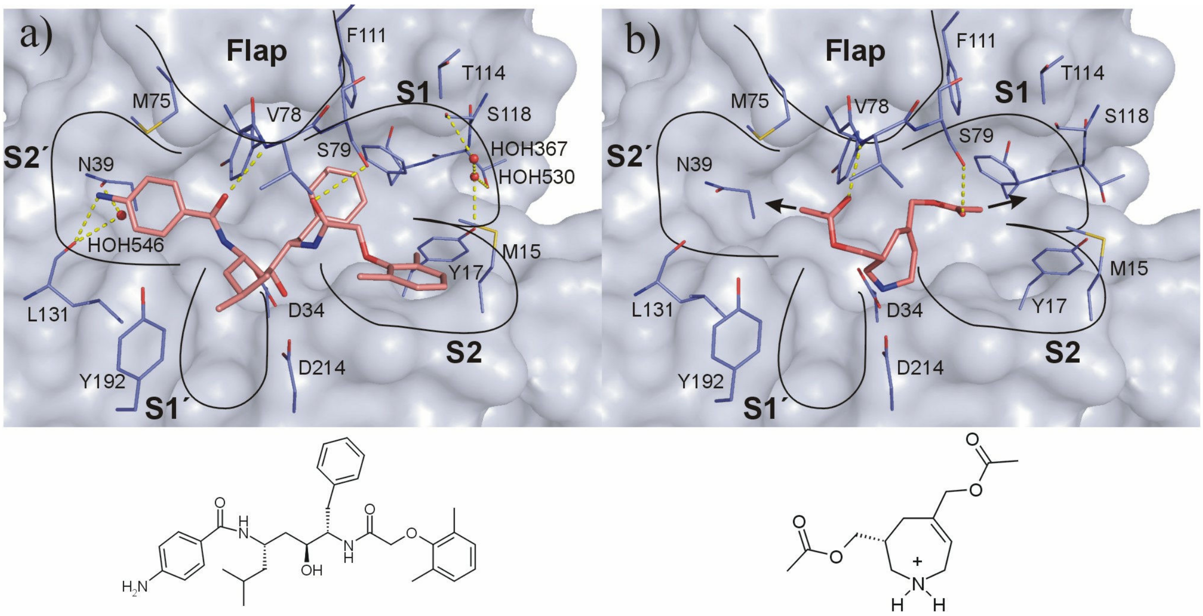
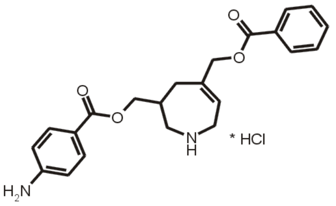
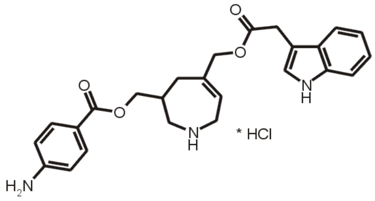
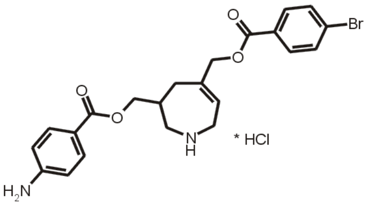


Abbildung 2. a) Kristallstruktur 1LF2 in Komplex mit RS367. **b)** Docking Lösung des Azepingrundgerüsts (*R*-Enantiomer) in 1LF2. Sowohl in der Kristallstruktur als auch in der Docking Lösung werden ähnliche Wasserstoffbrücken zum Protein gebildet.

Aus dem Docking des Azepin-Grundgerüsts in die Plasmepsin Struktur 1LF2 (Abb. 2b) geht hervor, dass Substituenten in Position 3 und 5 des Azepin Grundgerüsts vermutlich in Richtung der S1 und S2’ Taschen zeigen. Aus diesem Grund wurden die beiden Taschen sorgfältig daraufhin untersucht, welche Substituenten gute Wechselwirkungen mit ihnen ausbilden könnten. Zunächst wurde eine Datenbanksuche mit dem Programm Cavbase durchgeführt, was zum Ziel hatte, ähnliche Subtaschen zu Plasmepsin II in anderen Kristallstrukturen aufzufinden. Der Anteil der Liganden, der in ähnlichen Subtaschen gefunden wurde, wurde extrahiert und genauer analysiert in Hinsicht auf ihre Eignung als Baueinheit für die Synthese von Azepin-Inhibitoren. Für die Plasmepsin II S1-Tasche wurde eine Phenylgruppe als vielversprechende Einheit gefunden, während die Suche für die S2’-Tasche weniger schlüssig verlief. Um einen geeigneten Substituenten für die S2’-Tasche zu finden, wurde eine gründliche Suche mittels kombinatorischem Dockings mit FlexX^C durchgeführt. Die am besten bewerteten Verbindungen wurden einer visuellen Inspektion unterzogen, woraus sich die ersten Synthesevorschläge ergaben. Die Testung der ersten synthetisierten Verbindungen bestätigte die Design-Vorschläge und lieferte für Plasmepsin II und für Plasmepsin IV K_i Werte im einstellig mikromolaren Bereich (Tabelle 1). Gleichzeitig zeigte sich eine gute Selektivität gegenüber dem humanen Enzym Cathepsin D. Strukturelle Modifikationen des ersten Hits, die basierend auf den Erkenntnissen der verfeinerten, vorhergesagten Bindungsmoden synthetisiert wurden, lieferten verbesserte K_i-Werte, die zum Teil im submikromolaren Bereich liegen.

Table 1. Experimental Inhibitory Activities of the Compounds Studied

Compound	Structure	Enzyme K_i (µM)		
		Plm II	Plm IV	Cat D
1		4.7	7.2	260
2		1.0	1.5	262
3		0.4	5.7	165

Insgesamt wurden 13 Inhibitoren synthetisiert, von denen die experimentell bestimmten K_i Werte generell gut im Einklang mit der Design Hypothese stehen. Die beobachteten Struktur-Wirkungsbeziehungen stützen dabei den vorhergesagten Bindungsmodus.

Weitere interessante Grundgerüste wurden kürzlich zur Hemmung der HIV-1 Protease innerhalb der Arbeitsgruppe synthetisiert, bei denen es sich um Pyrrolidin-Derivate handelte. In der Literatur sind viele Fälle beschrieben, bei denen HIV-1 Protease Inhibitoren gleichzeitig gute Hemmstoffe für Plasmepsine darstellen, obwohl die Plasmepsine strukturell eine Pepsin-ähnliche Faltung einnehmen, während die HIV-1 Protease zu den Retroviralen Proteasen gehört. Beide Proteine besitzen aber Gemeinsamkeiten, die für ähnliche Bindungsprofile verantwortlich sein könnten. Zu diesen Gemeinsamkeiten gehören die katalytische Diade, ein Peptid-Erkennungsmotiv und der Flap, der an der Bindung des Substrats beteiligt ist. Die Flaps schließen sich bei der Substratbindung. Bei diesem Prozess wird Lösungsmittel vom Inneren der Bindetasche der Proteine ausgeschlossen. Da die Pyrrolidinderivate synthestisiert wurden, um genau diese familienweiten Merkmale zu adressieren, schien eine Hemmung der Plasmepsine mit Pyrrolidinderivaten als sehr wahrscheinlich (Abbildung 3).

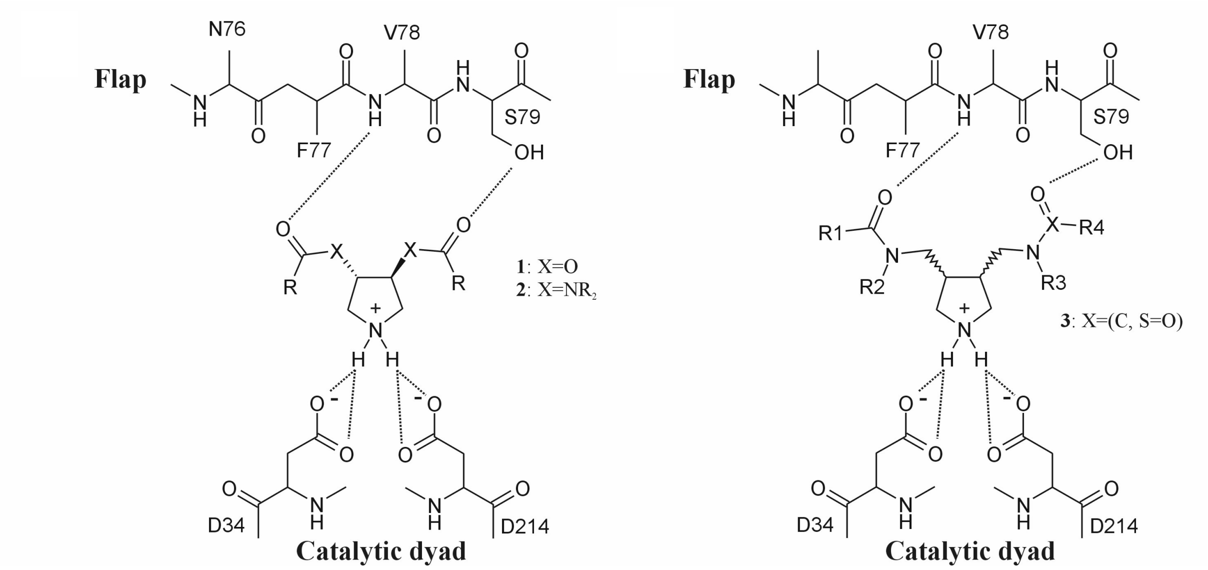
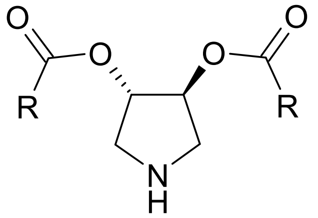
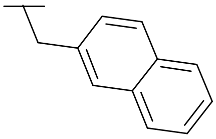
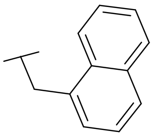


Abbildung 3. Mutmaßliche Wasserstoffbrückenbindungen der verschiedenen Pyrrolidingrundgerüste zu Plasmepsin II.

Von insgesamt zwölf Molekülen wurden K_i Werte für Plasmepsin II und IV bestimmt. Dabei zeigten die affinsten Liganden K_i Werte im nanomolaren Bereich (Tabelle 2).

Table2. Experimental Inhibitory Activities of the Compounds Studied

Compound	Structure	Enzyme K_i (µM)		
		Plm II	PlmIV	Cat D
4				
4.1		1.0	0.8	0.5
4.2		0.1	0.089	0.374

Um die entstandenen Struktur-Wirkungsbeziehungen besser verstehen zu können, wurde versucht plausible Bindungsmoden zu generieren. Dies gelang für die Pyrrolidin-diol-diester

Derivate und das Pyrrolidindiamid (Abbildung 3a). Mit den vorhandenen kristallographischen Informationen war es jedoch nicht möglich, einen sinnvollen Bindungsmodus für die Pyrrolidindimethylenediamin-Derivate (Abbildung 3 b) vorherzusagen.

Allerdings zeigten einige strukturelle Untersuchungen, dass die Plasmepsine sehr flexible Proteine sind. Aus diesem Grund wurden Moleküldynamik (MD) Simulationen durchgeführt und mobile Bereiche innerhalb der Bindetasche genauer untersucht. Unterstützt durch die MD Simulation konnte ein Bindungsmodus für den Typ der Pyrrolidindimethylenediamin-Derivate vorgeschlagen werden (Abbildung 4).

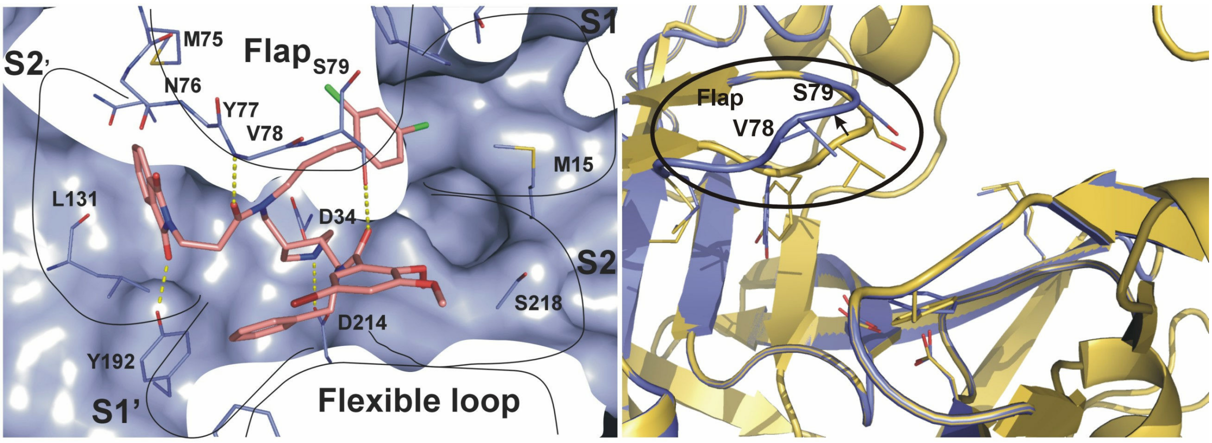


Abbildung 4. Eine konformative Öffnung des Flaps ist erforderlich, um Schlüsselinteraktionen zwischen Ligand und Protein ausbilden zu können (links). Eine solche Öffnung wird innerhalb der MD Simulation beobachtet (rechts).

Auch bei der Affinitätsdaten-Analyse einer Norstatin-Bibliothek, bei der die Verbindungen sich nur im P1 Substituenten unterschieden, konnten zunächst überraschende Ergebnisse mit Hilfe der MD Simulation erklärt werden. Eine konformative Öffnung der Plasmepsin II S1 Tasche wurde beobachtet, wodurch eine Anpassung an sperrige P1 Substituenten ermöglicht wird.

Konformative Übergänge aller Subtaschen wurden entlang der Trajektorie genau untersucht und dabei konnte I133 eine Schlüsselrolle bei der Generierung von diskreten Konformationen der S2' Tasche zugewiesen werden. Die Beweglichkeit dieses Restes scheint nicht auf die Seitenkette limitiert zu sein, sondern schliesst das Peptid-Rückgrat mit ein. Die Tasche wird weiter geöffnet und die Einpassung von Liganden mit sterisch anspruchsvollen Resten sollte möglich sein (Abbildung 5).

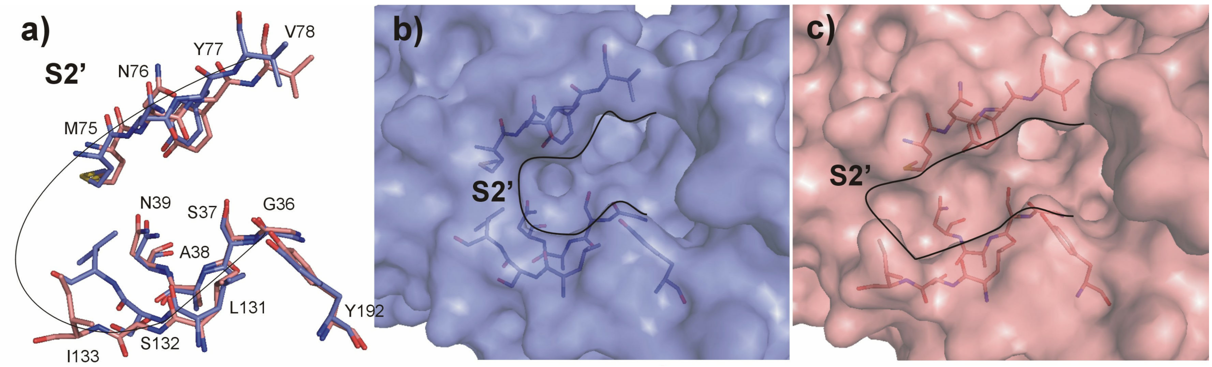


Abbildung 5. a) Überlagerung zweier Schnappschüsse aus der MD Simulation. Dargestellt sind Reste aus der S2' Tasche. Die Position von I133 hat sich stark verändert. **b, c)** Eine Oberflächenbetrachtung zeigt die Form der S2' Tasche, die sich abhängig von der I133 Position stark verändert.

Es wurde deutlich, dass die Plasmepsine durch ihre hohe strukturelle Flexibilität schwierige Targets für das strukturbasierte Wirkstoffdesign sind und dass Routine Drug Design Protokolle, basierend auf einer einzigen rigiden Proteinkonformation nicht ausreichend sind, um korrekte Bindungsmoden für alle Inhibitoren vorherzusagen. Daher kann eine einzelne ausgewählte Konformation beim Virtual Screening auch nur dazu dienen, lediglich einen Teil der möglichen Hemmstoffe aufzufinden.

Ausgehend von den Azepin Derivaten wurde in dieser Arbeit ein weiterer Ansatz verfolgt, um Inhibitoren für Plasmepsin II und IV mit neuartigen Grundgerüsten zu entwickeln, bei dem der Einfluss der Proteinkonformation bei der Suche reduziert wurde. Ftrees (Feature Trees), ein Computerprogramm, das es ermöglichen soll, neue Grundgerüste zu entdecken, die eigenschaftsverwandt mit dem Suchmolekül sind, wurde für diesen Zweck verwendet. Es stellt eine zeiteffiziente Methode dar, mit der riesige Kleinmoleküldatenbanken durchsucht werden können. Feature Trees werden aus Knoten und Kanten zusammengesetzt. Dabei repräsentieren Knoten die physikochemischen und sterischen Eigenschaften eines Moleküls, während Kanten die Information über die topologische Verknüpfung der Knoten wiedergeben. Es gibt zwei Typen von Feature Trees: Wenn nahezu jeder funktionellen Gruppe eines Moleküls ein Knoten zugeordenet wird, dann wird das Feature Tree „Hochaufgelöstes Feature Tree" genannt. Wenn die physikochemischen und sterischen Eigenschaften der gesamten Verbindung in einem einzigen Knoten gespeichert werden, dann wird die Repräsentation als „Niedrigaufgelöstes Feature Tree" bezeichnet. Ziel war es, eine Verbindung mit vergleichbarer biologischer Aktivität zu Verbindung **1** (Tabelle 1) zu finden, die aber gleichzeitig eine niedrige strukturelle Ähnlichkeit zu diesem Molekül besitzt. Von Verbindung **1** wurde das entsprechende Ftree generiert (Abbildung 6). Die Knoten sind in einem Farbcode dargestellt, der die Moleküleigenschaften repräsentiert.

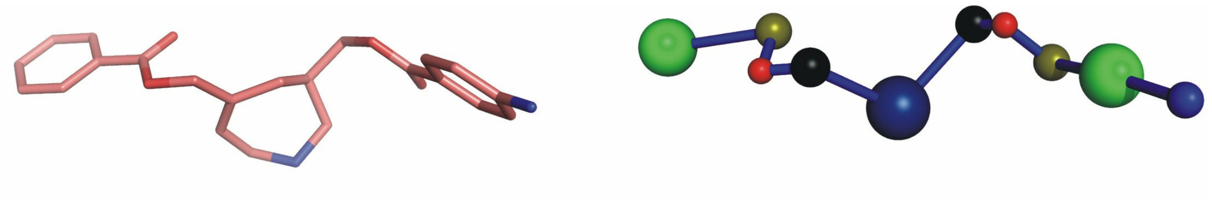
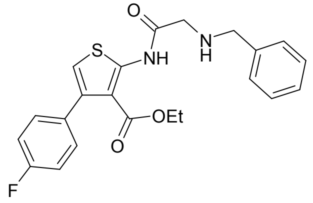
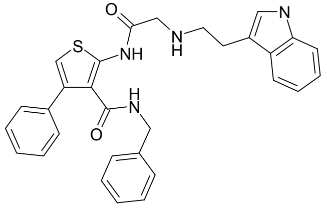


Abbildung 6. Verbindung **1** (links) und seine entsprechende Darstellung als "Hochaufgelöstes Feature Tree" (links). Die beiden aromatischen Ringsysteme werden als grüne Knoten gezeigt, die beiden hydrophoben Methylgruppen als schwarze Kugeln. Die Aminogruppen, als Vertreter von Wasserstoffbrückendonorgruppen sind als blaue Kugeln abgebildet. Wasserstoffbrückenakzeptorgruppen werden als rote Knoten angegeben. Knoten mit mehreren Eigenschaften werden mit einer Mischfarbe gezeigt. So wird dem hydrophoben siebengliedrigem Ringsystem mit seiner Aminogruppe eine dunkelblaue und dem Carbonylsauerstoff als Wasserstoffbrückendonorgruppe mit dem Ester Kohlenstoff eine goldfarbene Kugel zugewiesen.

Für die Ftree Analyse wurde die ZINC Datenbank mit ca. zwei Millionen Molekülen durchsucht und mit dem Suchmolekül verglichen. Die Auswertung der Suche ergab, dass das am besten bewertete Molekül zwar vermutlich Affinität zu den Plasmepsinen zeigen würde, jedoch konnte ein Übergang zu einem Grundgerüst mit sehr niedriger struktureller Ähnlichkeit nicht erreicht werden. Aus diesem Grund wurden die Screening-Bedingungen verändert. Als Alternative wurde eine „niedrigaufgelöste" Feature Tree Suche durchgeführt, die aufgrund ihrer Grobheit, in Hinsicht auf die Ähnlichkeitsmessung, nur als Vorfilter diente. Folglich wurde ein zweiter Filterschritt hinzugefügt, bei dem die bestbewerteten Moleküle in eine Plasmepsinstruktur gedockt wurden. Verbindung **5** erwies sich als besonders interessant, da sie eine hohe Bewertung bei der Ftree Suche und durch die FlexX Scoring Funktion erhielt und zudem strukturell sehr stark von der Suchverbindung **1** abweicht (Tabelle 3).

Table 3. Experimental Inhibitory Activities of the Compounds Studied

Compound	Structure	Enzyme K_i (µM)	
		Plm II	PlmIV
5		52 69	50 66
6		2,8	0,6

Von der gekauften Verbindung **5** wurden die K_i Werte für Plasmepsin II und IV bestimmt. Die Werte befanden sich im zweistellig mikromolaren Bereich für beide Enzyme, was nur als mittelmäßig einzustufen ist. Eine genauere Analyse der Docking Ergebnisse zeigte aber, dass **5** großes Potential für eine Affinitätssteigerung bietet (Abbildung 7).

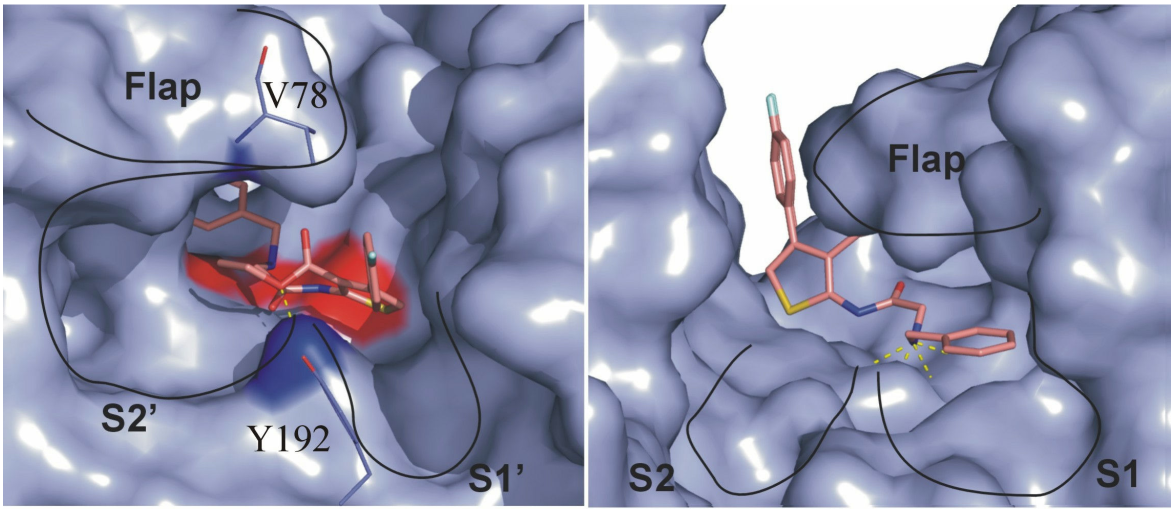


Abbildung 7. Verbindung **5** wurde mit FlexX in eine Plasmepsin II Struktur gedockt (1LF2). Beide Ansichten (links und rechts) zeigen, dass besonders Modifikationen des P2' uns P1 Substituenten zu einer Affinitätssteigerung beitragen könnten.

Besonders die leicht zugängliche Synthese des Thiophengrundgerüstes über die Gewald Reaktion ermöglichte die Erstellung einer Serie von Verbindungen. Die experimentell bestimmten K_i Werte sind dabei bei den meisten Inhibitoren zu einem hohen Grade im Einklang mit dem vorhergesagten Bindungsmodus. Der affinste Hemmstoff innerhalb dieser Serie (**6**) konnte mit K_i Werten im mikromolaren (Plasmepsin II) und submikromolaren Bereich (Plasmepsin IV) eine 110 fache Verbesserung zum ursprünglichen Hit vorweisen.

Insgesamt gelang es in dieser Arbeit durch unterschiedliche Ansätze Hemmstoffe für Plasmepsin II und IV vorherzusagen, bei denen die Struktur-Wirkungsbeziehungen mit den erzeugten Bindungsmoden konsistent zu diskutieren sind. Strukturlösungsversuche gelangen nur mit Pepstatin A, Moleküldynamiksimulationen gewährten dennoch einen Einblick in die hohe Flexibilität und stabile Konformationen der Plasmepsine. Daraus wurden Erkenntnisse für das strukturbasierte Wirkstoffdesign gewonnen, die einen Beitrag zur Entwicklung neuer Malaria-Medikamente leisten könnten.

10 Anhang

Danksagungen

Herrn Prof. Dr. Gerhard Klebe danke ich dafür, dass ich diese interessante und methodisch vielfältige Arbeit in seinem Arbeitskreis durchführen konnte und dass mir dabei sämtliche Freiheiten offenstanden. Zusätzlich danke ich Ihm für die Ermöglichung des Besuchs von internationalen Tagungen, wodurch stimulierende Ansätze für die eigene Arbeit gewonnen werden konnten.

Bei Herrn Prof. Dr. Christof Sotriffer möchte ich mich dafür bedanken, dass er sich auch bei gefülltem Terminkalender immer Zeit genommen hat, Probleme zu diskutieren und Sachverhalte zu erklären. Weiterhin möchte ich Ihm auch für die perfekte Organisation des Erstsemester-Praktikums bedanken, die allen Assistenten viel Arbeit abnahm.

Dr. Andreas Heine danke ich für die ausführliche Einführung in die Röntgenstrukturanalyse und die spannenden aber auch amüsanten Fahrten zum Synchrotron (Plumbonium).

Dr. Wibke E. Diederich, Nan-Si Chan, Ruben Bartholomäus, Dr. Andreas Blum, Dr. Steffen Weik und Dr. Sascha Brass danke ich für die Zusammenarbeit, die zur Synthese vieler Plasmepsin Inhibitoren führte.

Serghei Glinca möchte ich für seine tatkräftige Hilfe bei der Erzeugung der letzten Ergebnisse danken.

Für das Korrekturlesen von Teilen dieser Arbeit bzw. darin enthaltenen Publikationen bedanke ich mich bei Prof. Dr. Christoph Sotriffer, Dr. Klaus Reuter, Dr. Andreas Heine und Dr. Wibke E. Diederich.

Den Mitgliedern der Arbeitsgruppe Klebe danke ich für die angenehme Arbeitsatmosphäre.

Dr. Daniel Rauh danke ich für eine schöne und lehrreiche Zeit als Vertiefungsstudent, die mich neugierig auf mehr gemacht hat.

Allen Vertiefungsstudenten danke ich für Ihre Hilfe bei der Proteinproduktion.

Danken für eine schöne Zeit in Marburg möchte ich Naomi, Maja, Stephan, Kian, Markus, Paddy, Jark, Andreas, Paul, Edgar, Alex, Dani, Sarah, Daniel, Katrin, Jürgen

Meiner Freundin, meinen Eltern und meinen Freunden danke ich für die Unterstützung, ohne die diese Arbeit nicht möglich gewesen wäre.

Erklärung

Ich versichere, dass ich meine Dissertation

Structure-based Inhibitor Design for the Antimalarial Target Plasmespin

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 11.12.07

(Torsten Luksch)

Curriculum Vitae

Since 11/2003	Thesis: „Structure based inhibitor design for the antimalarial drug target plasmepsin” (Research Group Prof. Dr. G. Klebe, Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Germany), supported by a DFG Graduiertenkolleg grant „Protein function at the atomic level“.
9/2003	Diploma
3/2003 – 9/2003	Diploma thesis: „Expression, inclusion body purification, refolding, crystallisation and structure determination of plasmepsin II and IV” (Prof. Dr. T. Carell/Prof. Dr. G. Klebe, Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Germany)
8/2001 – 4/2002	Research stay (Dr. J-P. Pillot, CNRS: Laboratoire de chimie organique et organométallique, Université de Bordeaux) supported by a Sokrates grant
10/2000	Intermediate diploma
10/1998	Begin of the chemistry studies (Philipps-Universität Marburg, Germany)
9/1997 – 7/1998	Military service (Fernmeldebataillon Koblenz, Germany)
1997	High School Diploma (Mons Tabor Gymnasium Montabaur, Germany),
1989 - 1997	High School (Mons Tabor Gymnasium Montabaur, Germany)
Date of birth:	Sep. 2, 1977 in Dernbach (WW), Germany

Poster presentation:

8/2006	1st European Chemistry Congress, Budapest, Hungary: *Structure Based Drug Design for the highly flexible protein plasmepsin II* (Luksch,T., Weik, S., Rademann, J., Sotriffer, C., Heine, A., Klebe, G.)
05/2006	8th international school on the crystallography of biological macromolecules, Como, Italy: *Protein Engineering: Redesigning of Prephenate Dehydratase* (Luksch,T., Kleeb, A., Hilvert, D., Klebe, G., Heine, A.)
03/2006	GDCh Medicinal Chemistry: Frontiers in Medicinal Chemistry, Frankfurt, Germany: *Inhibitor design strategy for the antimalarial drug target plasmepsin II using protein-flexibility information generated by MD simulation* (Luksch,T., Weik, S., Rademann, J., Sotriffer, C., Heine, A., Klebe, G.)
03/2006	GDCh Medicinal Chemistry: Frontiers in Medicinal Chemistry, Frankfurt, Germany: *Facile synthesis of substituted azepines as putative non-peptidic inhibitors of Plasmepsin II* (Brass, S.,Chan, N., Luksch, T., Klebe, G., Diederich, W.E.)
08/2005	DPhG–Jahrestagung, Mainz, Germany: *Inactivation of the antimalarial drug target plasmepsin II by new classes of inhibitors* (Luksch,T., Weik, S., Rademann, J., Heine, A., Klebe, G.)
08/2005	DPhG–Jahrestagung, Mainz, Germany: *Design and Synthesis of substituted azepines as potential non-peptidic inhibitors of plasmepsin II* (Brass.S., Luksch, T., Klebe, G., Diederich, W.E.)
07/2005	Eurocombi 3: 3rd European Symposium on Combinatorial Sciences, Winchester, England: *The potential of P1-site mutations in peptide isostere inhibitors explored via the use of CC-coupling linker reagents* (Weik, S., Luksch, T. Klebe, G., Loeffler, HG., Hasilik, A., Rademann, J.) (awarded with a poster prize)
09/2004	28th European Peptide Symposium, Prague, Czech Republic: *The potential of P1-site mutations in peptide isostere inhibitors explored via the use of CC-coupling linker reagents* (Weik, S., Luksch, T. Klebe, G., Loeffler, HG., Hasilik, A., Rademann, J.)

Publication:

The Potential of P1 Site Alterations in Peptidomimetic Protease Inhibitors as Suggested by Virtual Screening and Explored by the Use of C_C-Coupling Reagents
Weik, S., Luksch, T., Evers, A., Böttcher, J., Sotriffer, C., Hasilik, A., Löffler, H-G., Klebe, G., Rademann, J. ChemMedChem, **2006**, 4, 445-457

Synthesis of 2,3,4,7-tetrahydro-1H-azepines as privileged ligand scaffolds for the design of aspartic protease inhibitors via a ring-closing metathesis approach
Brass, S., Chan, N., Gerlach, C., Luksch, T., Böttcher, J., Diederich, W.E.
Journal of Organometallic Chemistry, **2006**, *691*, 5406-5422