

Purification, co-Crystallization and Structure Determination of *Plasmodium falciparum* plasmepsins.

Master's Thesis

Matthias Barone^① and Dr. Lars Prade^②

^①*M.Sc Curriculum Nano-Sciences,*

^②*Actelion Pharmaceuticals Ltd, Gewerbestrasse 16, 4123 Allschwil, Switzerland*

October 27, 2009

Supervised by Prof. Tilman Schirmer, Biozentrum, University of Basel, Switzerland

Abstract

In the erythrocytic stage of the life cycle, the malaria parasite *Plasmodium falciparum* degrades hemoglobin inside an acidic food vacuole. Degradation is initiated by a number of aspartic proteases, called plasmepsins. Three of these members were expressed in inclusion bodies of *E. coli* purified by a protocol, which had been optimized for plasmepsin II during the master's thesis.

Plasmepsin II and IV were co-crystallized with non-peptidometric inhibitors. Protocol optimization led to crystals of three complexes which were solved in orthorhombic and triclinic crystal packings (space groups $P2_12_12_1$ and $P1$) at a resolution limits of less than 2.7 and 1.95 Å.

The complexes display a binding mode involving both hydrophobic and polar interactions. Thereby the inhibitors interact with the flap pocket and H-bonding to the catalytic dyad without need of a catalytic H_2O .

Contents

1	Introduction	2	3.1.2	PMIV	12
			3.1.3	PMI	12
			3.2	Crystallization	14
			3.2.1	Control of Protein Purity	14
			3.2.2	Crystal Growth	14
2	Materials and Methods	2	3.3	Structure Determination	18
2.1	Chemicals	2	3.3.1	Compounds	18
2.2	Main Protocol	3	3.4	Structure Description	21
2.2.1	Transformation	4	3.4.1	Interfaces	21
2.2.2	Test Expression	4	3.5	Structure Comparison	22
2.2.3	Expression	4	3.5.1	Diff. among Chains	22
2.2.4	Purification and Activation	5	3.5.2	Diff. ST47 and ST43	23
2.2.5	Crystallization	6	3.5.3	Active Site & Flap	24
2.3	Data Acquisition and Processing	6			
3	Results and Discussion	6	4	Conclusion and Look-out	27
3.1	Purification	7	4.1	Purification	27
3.1.1	PMII	7	4.2	Inhibitors	28

1 Introduction

Tropical malaria constitutes one of the most impoverishing infectious diseases, annually affecting 200–500 million people worldwide and causing the death of 2.5 million people - alone one million children in Africa^[1, 2].

Malaria Pathogens and Their Life Cycle

The tropical disease malaria is caused by protozoan parasites. There are four pathogen species which each cause a specific clinical picture: *Plasmodium falciparum* (Malaria tropica, the most serious form), *P. vivax* (Malaria tertiana, the widest spread form), *P. ovale* (Malaria tertiana) and *P. malariae* (Malaria quartana). The genus *Plasmodium* belongs to the phylum *apicomplexa* (*old: sporozoa*) and posses therefore a rather complex life cycle:

The most important mammalian reservoir is provided by humans where the parasite spends its asexual life cycle. The sexual life cycle is finished in female mosquitoes of the genus *Anopheles*. Mosquitos are infected by *gametocytes*. After maturation and fusion, the newly formed zygote resides within an oocyst on the basal side of the female *Anopheline* mosquito mid gut wall^[3]. The zygote produces in a later step infecting cells called *sporozoites*, which migrate into the anticoagulant-producing salivary gland^[4].

During piercing a person's skin the anticoagulant containing the infective form of *Plasmodium* is injected into the hepatic system of a new host. Once in the blood stream, the sporozoites are removed by the liver and infect hepatocytes. These cells serve as hosts in which the sporozoites multiply asexually and differentiate to *merozoites*. When the host cell is ruptured the mature merozoites escape from the liver into the blood stream undetectable, thus beginning the erythrocytic stage of the life cycle^[4].

The subsequent invasion of erythrocytes produces all the characteristic symptoms known from a malaria infection: fever up to 40°C, followed by periodic ague, may accompanied by vomitus and headache. The continuous lysis of red blood cells leads to anemia in a later phase.

During the erythrocytic stage, the protozoan parasite satisfies its requirements by degrading hemoglobin within an acidic food vacuole. Therein plasmepsins I to IV (PMI to PMIV) initiate degradation of hemoglobin. It is assumed that all PMs degrade host cell hemoglobin. There are three isoforms of aspartic proteases (PMI, II and IV) which all show in vitro degradation of hemoglobin. PMIII is known as histo-aspartic protease (HAP).

The predicted coding sequences of PMI, II, IV, and HAP are 60–73% identical on amino acid level. High levels of sequence identity in both pro- and mature sequence suggest that PMI, II, IV, and HAP have arisen through relatively recent gene duplication events^[5].

Compared to PMII and I, HAP has several substitutions including replacement of a catalytic aspartate with a histidine and changes in the conserved flap motif^[6].

The aim of the work was to determine structures of PM complexes with non-peptidometic inhibitors synthesized by groups of Prof. Dr. Gerhard Klebe (University of Marburg) and Prof. Dr. François Diederich (ETH Zürich).

2 Materials and Methods

2.1 Chemicals

Chemicals purchased from Fluka: polyethylene glycol 4000 (PEG 4K, ultra, 81268); glycerol 87% (92.10 g/mol+aq, 86 – 89%, 49782); tris(hydroxymethyl)aminomethane (Tris, 121.14 g/mol, ≥ 99%, 93352); sodium hydroxide (NaOH, 40.00 g/mol, 71690); glycine

(75.07 g/mol, $\geq 99\%$, 50049); ethylenediamine-tetraacetic acid disodium salt dihydrate (EDTA, 372.24 g/mol, $\geq 99\%$, 03679); hydrochloric acid 32% (HCl, $\geq 32\%$, 84421); sodium acetate anhydrous (NaAc, 82.04 g/mol, $\geq 99\%$, 71180); acetic acid (60.05 g/mol, $\geq 99.8\%$, 45731), sodium phosphate monobasic dihydrate (156.01 g/mol, $\geq 99\%$, 71500); tert.-butanol (74.12 g/mol, $\geq 99.7\%$, 19460), dimethylsulfoxide (DMSO, 78.13 g/mol, $\geq 99.9\%$, 41640), imidazole (68.08 g/mol, $\geq 99.5\%$, 56750).

Chemicals delivered by Sigma: isopropyl- β -D-thiogalactoside (IPTG, 238.3 g/mol, 367-93-1), ampicillin sodium salt (371.39 g/mol, A9518-25G); (-)-glutathione, oxidized (GSSG, 612.6 g/mol, 150568-5G); L-glutathione reduced (GSH, 307.32 g/mol, $\geq 99\%$, G4251-25G); 2-mercaptoethanol (β -MeEtOH, 78.13 g/mol, 1.115 g/l $\geq 98\%$ M3148-100ML); 1-octanol 99% (111-87-5), LB Broth (L7275-500TAB); Terrific Broth, Modified EZMix Powder (T9179-1KG); urea (U4128-5KG); ethanolamine (E-6133); potassium phosphate dibasic (174.2 g/mol, P8281-500G)

Obtained from Invitrogen were: One Shot BL21(DE3) competent cells; S.O.C. medium (15544-034); MagicMedia *E. coli* expression medium (K6802); NuPAGE 4-12% Bis-Tris Gel (10 or 12 well, NP0322BOX); NuPAGE MOPS SDS running buffer (20 \times , NP0001); NuPAGE LDS sample buffer (4 \times , NP0007); SeeBlue Plus2 pre-stained standard (1 \times protein ladder, LC5925); pH 3-10 isoelectric focusing (IEF) gel (10 well, EC6655BOX); Novex IEF pH 3-10 cathode buffer (10 \times , LC5310); Novex IEF pH 3-10 anode buffer (50 \times , LC5300).

Isopropyl- β -D-thiogalactoside (IPTG, 238.3 g/mol, 10724815 001) was provided by Roche.

NaCl (58.44 g/mol; $\geq 99.5\%$; A1149,1000) was purchased from AppliChem.

Except during expression SDS-PAGE samples

were stained as follows: 20 μ l sample, 5 μ l NuPAGE LDS sample buffer (4 \times), 2 μ l of 98% β -MeEtOH. Sample containing entire cells were stained with 80 μ l NuPAGE LDS sample buffer per OD600=1 and to every SDS sample, by default, 5 μ l of 98% (β -MeEtOH). All SDS samples were heated for 5 min at 95 °C in heat block (Techne, DRI block DB-2D) and then centrifuged at 13000 rpm for 4 min (Heraeus, Biofuge fresco). From protein ladder (MW marker) as well as samples 10 μ l were loaded onto gel. 12 well gels were run for 53 min at 200 V, 10 well gels for 44 min in XCell surelock (Invitrogen Novex mini cell).

All substances were sufficiently pure and were used without further purification. Ethanolamine, EDTA, 50% (w/v) PEG 4K, glycine, Tris, NaAc, MgCl₂, and NaCl were dissolved in deionized water (milli-Q) and afterward filtered through 0.45 μ m Durapore membrane filters (Millipore). Stock solutions of Na/K phosphate (Na/K P_i, used for SEC running buffer) and sodium citrate (used for crystallization screens) were filtered with 0.22 μ m Steriflip vacuum filtration system (Millipore).

Crystals were grown at 293 K using the hanging-drop technique and vapor-diffusion method with Cryschem plates HR3-275 (Hampton research). The crystallization solutions were first prepared manually and 200 μ l were used as reservoir. The droplet was formed by a ratio of 2:1 or 3:1 (v/v) protein to reservoir solution. For the whole pipetting procedure, a robot (Tecan, freedom evo) was used.

2.2 Main Protocol

The following protocol defines the final version achieved by continuous optimization of given protocols (written by Andrew Jones for PMII purification) and literature^[7]. Main differences between the final and the starting protocol are discussed below. For the expression media following

abbreviations were used: Terrific Broth (Tm), LB Broth (LB), and MagicMedia (Mm).

2.2.1 Transformation

Aliquots (50 μ l) of competent cells (Invitrogen, *Escherichia coli* One Shot BL21(DE3)) were thawed on ice. All tubes were held on ice until to each sample 1 μ l of plasmid (100 ng/ μ l) was added. After incubating for 30 min on ice, aliquots were heated for exactly 30 s at 42°C in a heating block (Eppendorf, Thermo-mixer Comfort) without shaking. After heat shock, tubes were placed for 2 min back on ice and afterward 250 μ l of pre-warmed (37°C) S.O.C. medium was added (tubes were kept on ice until all had received S.O.C.). After incubating at 37°C while shaking at 300 rpm for 60 min in a heating block, aliquots were spread onto 2 Petri dishes containing 5 and 20 μ l of cell medium. To one aliquot of competent cells, 1 μ l of H₂O instead of plasmid was added as control. From this aliquot the whole 301 μ l were spread onto one Petri dish. After drying for 5-10 min, the plates were turned upside-down and incubated at 37°C over night (o/n).

2.2.2 Test Expression and Glycerol Stock

For test expression 5 clones were picked and incubated o/n at 37°C in 3 ml LB (Falcon 14 ml polystyrene round-bottom tubes) containing 50 μ g/ml ampicillin (Amp). From the over-night cultures 0.5 ml were used to inoculate 50 ml (1:100 (v/v)) of prewarmed LB (50 μ g/ml Amp) and shaken (230 rpm) at 37°C until cultures reached an optical density at 600 nm (OD600) of 0.6. From every cell culture, a 3 ml sample was placed on ice.

The first 2 ml were used to form 10 \times 200 μ l glycerol stock aliquots. Cell culture was mixed in cryo tubes (Nunc, CryoTube vials 1.8 ml) with 1:1 (v/v) of 87% glycerol by finger flicking and

flash frozen in liquid nitrogen. Glycerol stock of all 5 clones were stored at -80°C. The remaining 1 ml was used for SDS-PAGE (NuPAGE 4-12% Bis-Tris Gel).

The remaining 47 ml of cell culture were induced by addition of isopropyl- β -D-thiogalactoside (IPTG) to a final concentration of 1 mM and incubated for 3 h. Another SDS sample (stained with LDS buffer according to the new OD600) was then taken and an SDS-PAGE gel run to compare the 5 clones.

2.2.3 Expression

According to SDS-PAGE one clone was picked and an over-night culture with 50 ml LB (50 μ g/ml Amp) was inoculated with 100 μ l of its glycerol stock (using sterile tips) and stirred at 160 rpm and 37°C. Together with the over-night culture the expression medium may prewarmed (adding Amp before prewarming!).

According to preferred expression media, different procedures were needed for proper inoculation and induction.

LB Broth had to be induced manually by adding 1 mM IPTG when the cell culture reached OD600 of 0.6. Therefore, it was important to inoculate the medium as early as possible in the morning. LB was inoculated with 1:100 (v/v) of o/n culture (10 ml per liter LB).

Terrific Broth contained additionally to the powder 0.8% (v/v) glycerol (9.2 ml 87% glycerol per liter). Glycerol had been added before autoclaving the medium. Tm was inoculated with 1:100 (v/v) of o/n culture, expression had to be induced manually by adding 1 mM IPTG at an optical density of 0.6.

MagicMedia was designed as a two-component medium. Component A (SoluPouch) was dissolved in 950 ml H₂O and autoclaved.

Before inoculation, the medium was completed by adding 50 ml of component B and 50 $\mu\text{g/ml}$ Amp. *E. coli* transformed PMII cultures were grown o/n at 37°C, 220 rpm and harvested in stationary phase where they reached an OD600 of about 5.

For a proper blanking of the OD meter in later steps a 80 μl sample of every uninoculated medium was stored at 4°C.

Before inducing LB and Tm, a 1 ml sample had been taken and prepared for SDS. If required one drop of 1-octanol acting as anti-foaming was added for proper aeration. After induction by IPTG, media were incubated for 4 h before harvested. Another 1 ml sample was taken for induction/expression control by SDS-PAGE .

Cells were harvested in 500 ml centrifugation tubes (same number of tubes as liter of medium used) at 7000 rpm, 10°C for 7 min (Sorvall, RC 5C plus, SLA 3000-rotor, rotor code 30). Supernatant was discarded back into empty culture flasks directly beside the ultracentrifuge. Pellets were transferred by a spatula into 50 ml-falcon tubes, briefly centrifuged (Heraeus, Multifuge 1 L-R) and frozen at -20°C.

2.2.4 Purification and Activation

Pellets stored at -20°C were thawed on ice and resuspended with approximately 40 ml of buffer A (50 mM Tris pH 8.0; 150 mM NaCl) and stirred for about 30 min at 4°C. Before further processing, 20 μl were transferred into an Eppendorf tube and put on ice for SDS-PAGE.

Cells were disrupted by two rounds of french press (1.33 kBar, 10°C). French press was rinsed and washed to a maximal final volume of 6 \times 35 ml (the maximal centrifuge tube volume), 210 ml of buffer A.

Cell lysate was transferred into 4 to 6 35 ml-centrifuge tubes and centrifuged at 14000 rpm for

15 min at 15°C (Sorvall, RC 5C plus, SS34-rotor, rotor code 5). While supernatant was kept on ice, tubes were refilled with buffer B¹ (100 mM glycine pH 11.0; 50 mM β -MeEtOH). Homogeneously colored pellets were not resuspended. Tubes were vortexed for a few seconds and afterward centrifuged at 14000 rpm for 15 min at 15°C.

After the supernatant of buffer B was also stored on ice, pellets were diluted with 500 ml of pre-cooled buffer C¹ (100 mM Tris pH 8.0; 6 M urea; 1 mM glycine; 1 mM EDTA; 50 mM β -MeEtOH). After 2 h of dissolving at 4°C, an SDS-PAGE had been run containing the stored pellet, french press samples, supernatants A and B, plus the diluted protein in urea-buffer C.

Complete dilution in buffer C took place o/n at 4°C. Aliquots (50 ml) were stored at -20°C as an oxidation during the storage led to significant lower yield. Samples of 50 ml were diluted with P960 (ÄKTApurifier 10) with 0.1 ml/min into 1 l of buffer D (10 mM Tris pH 8.5, sterile filtered and pre-cooled to 4°C). The most rapid dilution was achieved if the tubing was placed in the middle of a vigorously stirring renaturation buffer. The final dilution of 1:20 (v/v) (final 300 mM urea) was enough for rapid refolding. The yield of properly folded material decreased if renatured protein in buffer D was stored longer than 24 h.

The first liter of diluted protein solution was therefore directly processed further by loading onto a HiTrap Q HP ion exchange column (GE Healthcare, 1 ml or 5 ml bed volume) with maximum flow (depending on back pressure of the system). After washing the column with 20 – 30 \times C.V. running buffer (0.1 M Tris/HCl pH 8.5) with 1 – 2 ml/min (starting conductivity around 3 mS/cm) protein was eluted by a gradient over

¹As buffer B and C contained temperature-sensitive material as β -MeEtOH or EDTA they were mixed directly before application.

$35 \times C.V.$, 1.5 ml/min to 100% elution buffer (0.1 M Tris/HCl pH 8.5; 1 M NaCl, conductivity of around 100 mS/cm).

Depending on the elution profile, two samples were pooled: 4 ml containing the peak-top and a second sample with the adjacent fractions including the populations behind the main peak. Pools were stored at 4°C . Although 1 month old adjacent peaks were able to activate themselves, the samples contained significant populations of misfolded and aggregated protein.

For activation, the sample was brought to room temperature (RT), mixed with 1:4 (v/v) of activation buffer (1 M sodium acetate (NaAc) pH 4.5) and gently shaken for $1.5 - 2\text{ h}$. Activation was detectable after 15 min on SDS gels. The activated main-peak sample (4.8 ml) was directly loaded onto a HiLoad 16/60 Superdex 200 pg size exclusion column (SEC) at 4°C .

The HiTrap pool containing the adjacent peaks usually was bigger than 5 ml (recommended sample volume for a 16/60 SEC). Because the crystallization conditions lay around pH 3-5 and therefore the protein seemed to have a significantly lower solubility in acidic environment, the activated sample (pH 4.5) was brought to pH 8.37 before centrifugation by addition of 10% (v/v) 1 M NaOH of the HiTrap pool volume (or 50% (v/v) of activation buffer used). Samples for SEC load were concentrated at 4°C (Amicon ultra centrifugal filter, 4.5 ml max. volume, 10 kD cut-off, 2850 rpm , cycles of maximally 4 min).

SEC had been run at 4°C , 1 ml/min with following running buffer: 5 mM Na/K P_i pH 6.5; 100 mM NaCl. Correctly folded PMII passed the 16/60 SEC after 86 ml . Pooled protein was put on ice and concentrated with 4.5 ml -Amicon filter up to 10 mg/ml at 4°C , 2850 rpm , $4 - 2\text{ min}$ per run (merging the whole solution after each run).

2.2.5 Crystallization

Inhibitors were dissolved to a concentration of 30 mM in DMSO or tert.-butanol and stored at 20°C . Purified protein was acidified in SEC running buffer with 1:400 (v/v) of 1 M HCl to a final pH 4.2 and afterward mixed by finger-flicking with an equimolar, 1:20, or 1:10 (v/v) ratio with inhibitors and incubated o/n at 20°C . Before pipetting the sample was centrifuged at RT at 13000 rpm for 4 min (Heraeus, Biofuge fresco).

2.3 Data Acquisition and Processing

Crystals were measured using synchrotron radiation light source at the PXIII (X06DA) beamline of the Swiss Light Source (Paul Scherrer Institute, Villigen AG, Switzerland). Datasets were indexed using go.com^[8], imosflm^[9, 10], or XDS^[11].

Molecular replacement was done with MOLREP^[12] using a pruned PMII as a search model. This model contained only the conserved core of PMII without three flexible loops. The structures were refined using REFMAC5^[13] and rebuilt by Coot^[14] or Moloc^[15]. Assessment of the models against their native structure factors were done using Sfcheck^[12].

Protein models (C_α -trace or cartoon) were rendered using PyMOL^[16] and finished in Photoshop (highlighting, labeling, erase irrelevant chains or adding artificial objects).

3 Results and Discussion

As an overview of the crucial steps, the following itemization lists the modified and improved steps and conditions. Each point is discussed separately:

- The main differences to the starting protocol (written by Andrew Jones) lie in crucial steps before and after rapid dilution. The

old protocol used, after activation of proP-MII a concentration step, a H4848 affinity column followed by another concentration by centrifugation for SEC load. After the SEC, a third concentration step was necessary for crystallization.

Thereby the H4848 showed no significant gain in purification. After checking the literature, the first two lossy concentrations and the H4848 column were replaced by a much more efficient loading of proPMII (no activation) onto a HiTrap column (analogue original literature). Using HiTrap columns we reached more than 250 \times -concentration (1 – 4 l-load pooled as 4 ml fractions).

- During renaturation proPMII is susceptible to unfolding if processed or stored at RT. Refolding, purification as well as concentration yielded significantly higher amounts of correctly folded populations whenever run at 4°C.
- The activation of proPMII in acidic environment was no longer run for 2 h at 37°C. An SDS-PAGE-tracing during the activation at RT showed complete activation after 2 h. Instead of 10% we used 20% (v/v) of 1 M NaAc to reach the final pH of 4.5.
- If the activated sample was too voluminous, the concentration step for SEC load was run at high pH (8.5) to prevent precipitation. For the same reason, ultracentrifugation was run at 4°C and 2850 rpm.
- Crystal growth took place in a solution containing two different buffers. PMII was incubated in SEC running buffer (5 mM Na/K P_i pH 6.5, acidified by HCl to pH 4.2), whereas 0.1 M sodium citrate was used to obtain the pH of the crystallization condition (buffered pH 2.5 to 7.0). There is evidence that during pipetting and crystal growth the pH of

crystallization droplets did not shift to the desired value - especially at pH 2.5 to 4.0. For this purpose the initial concentration of 20 mM Na/K P_i SEC running buffer was reduced to 5 mM Na/K P_i keeping the pH at 6.5.

- We observed that after refolding by rapid dilution, proper oxidation of PMII's cystein residues does not require addition of oxidizing and reducing agents like GSSG and GSH. The crucial factor is the flow by which the denaturated protein is injected into the renaturation buffer. To control it we optimized the renaturation using the P960 membrane pump at 4°C with a flow of 0.1 ml/min. Neither the time-consuming drop-by-drop injection by hand (as in the old protocol) nor using Bernoulli's principle could reach the sample quality as a slow, flow-controlled injection by P960 did.

3.1 Purification

3.1.1 PMII

Test expression and glycerol stock forming were run according to the main protocol.

Pellet Purification

proPMII was expressed in *E. coli* in inclusion bodies, which facilitated the pellet purification by centrifugation after cell disruption. Fig. 2 shows that dissolved impurities were removed mainly by centrifugation of buffer A. The bright white color of the pellets allowed us to conclude that inclusion bodies do not get disrupted by two rounds of french press. According to an SDS the subsequent centrifugation in 100 mM glycine pH 11.0 of buffer B did not improve the purity. As this additional centrifugation step lasted only 12 min, it was kept in the protocol.

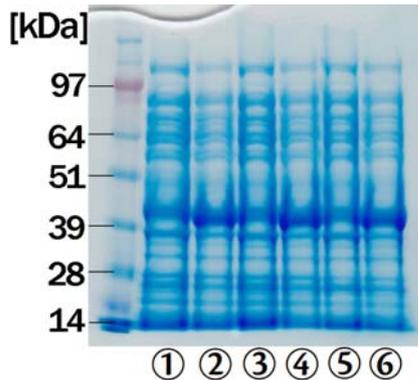


Figure 1: SDS-PAGE controlling the test expression of proPMII in three clones. Bands: MW marker, ① clone A before, ② after induction, ③ clone B before, ④ after induction, ⑤ clone C before, ⑥ after induction. Clone C was used for expression as this clone grew faster than the other (data not shown).

All subsequent steps after pellet purification were performed at 4°C or tubes were kept on ice until further processed (e.g. during concentration). Pellets were dissolved in urea with a pre-cooled buffer C while tubes were vigorously stirred over night at 4°C.

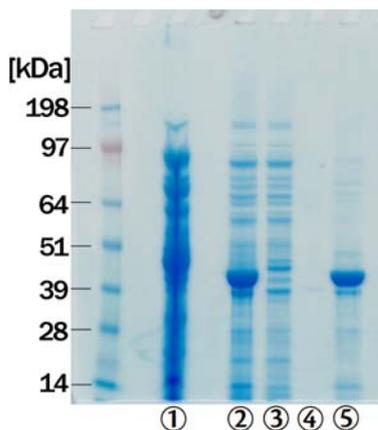


Figure 2: SDS-PAGE to check the pellet purification of proPMII. Bands from left to right: MW marker, ① pellet stored at -20°C, ② after two rounds of french press, ③ supernatant buffer A after centrifugation, ④ supernatant buffer B after centrifugation, ⑤ proPMII diluted in buffer C.

Controlling the Refolding by HiTrap Elution Profiles

The implementation of a HiTrap Q ion exchanger had two main advantages: a) proPMII diluted in rather large volumes could be bound to a HiTrap and concentrated by a NaCl-gradient elution, and b) correctly folded and misfolded proPMII got separated by the gradient as shown by isoelectric focusing (IEF, Fig. 3).

Using HiTrap columns with 1 ml bed volume and a long NaCl-gradient led to the convenient side effect of a main-peak pool with a volume around 4 ml. This was the maximum volume which could be loaded after activation directly on a SEC without concentrating the sample. The rest of the elution profile was also pooled. It was concentrated while the main-peak sample was passed through SEC. However, since all subsequent SEC flow-through profiles showed, this off-peak material was of less quality (SEC peak was thereby often flanked by shoulders).

By comparing the elution profiles of a HiTrap Q we were able to optimize the refolding step which was not possible with the old protocol. Thereby, we found several parameters which all influence the efficiency of the refolding. A sample of good quality should have a low ratio between misfolded and correctly folded protein. This ratio can be easily estimated by the peak areas of a HiTrap elution profile.

There are mainly four criteria which influence the efficiency of the refolding by rapid dilution:

There is evidence to suggest that the expression medium influences the quality of the protein in inclusion bodies. Compared to LB and especially Tm, the elution profiles of proPMII expressed from *E. coli* in Mm showed throughout better quality.

The time in which proPMII stays in renaturation buffer is crucial and should be kept as short as possible because the subsequent step - the load-

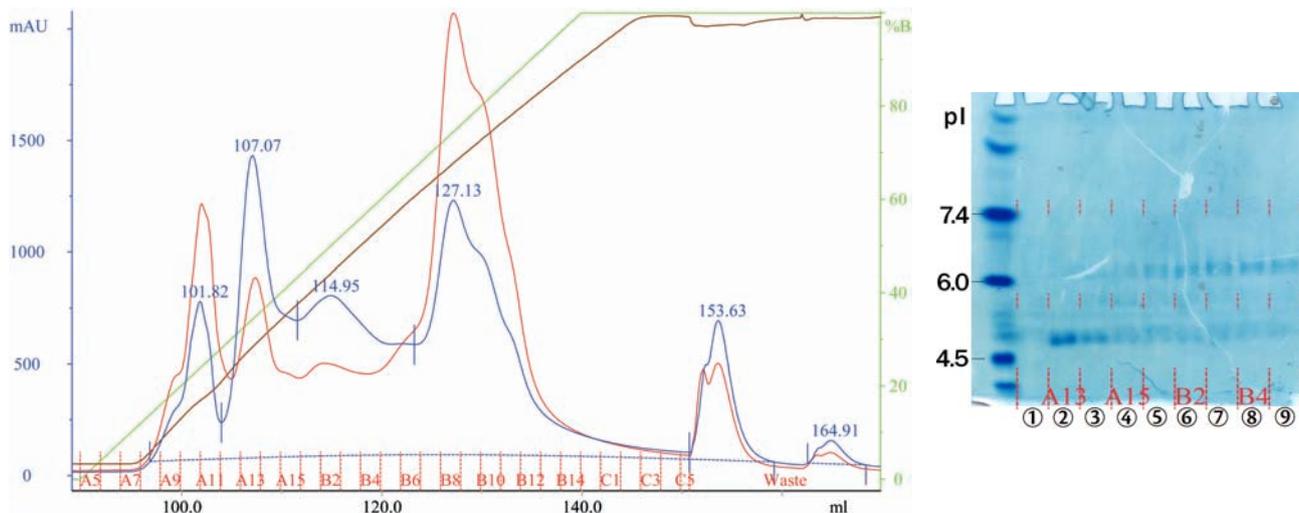


Figure 3: Elution profile of the first batch proPMII expressed in LB and processed according to the old protocol. The IEF bands ① to ⑨ correspond to fractions A12 to B5 of the elution profile. Behind the main peak (A13 and A14) a second population (which is thought to be misfolded proPMI) with a different isoelectric point (pI) is detectable. According to the IEF-marker the population in A13 has a pI around pH 4.8, whereas the population behind the main peak reveals a pI at pH 6.3. IEF

ing on a HiTrap with $1\text{ ml}/\text{min}$ still lasts 17.5 h per 1 l . For this purpose only small aliquots (50 ml) were diluted in 1 l of renaturation buffer. Such small samples can be diluted with a small flow rate which increases the quality significantly without losing much time in which already renaturation protein stays in solution.

It seems that the ratio between folded and misfolded proPMII cannot be influenced by oxidizing/reducing agents. Even if used at non-equimolar amounts, the ratio could not be shifted toward correctly folded protein. To highlight this conclusion, Fig. 4 shows an overlay of elution profiles run at 4°C . One sample (dotted line) was renaturation in buffer D according to the main protocol. For comparison, the elution profile of a sample renaturation in buffer D containing 1 mM GSSG and 0.5 mM GSH is printed in straight lines.

The most relevant factor is the temperature. Samples which were renaturation, loaded, and eluted at 20°C showed significantly higher amounts of misfolded populations compared to samples at 4°C . This fact holds for both normal buffer D (Fig. 5) and buffer D containing

GSSG/GSH (Fig. 6). In both Figures the sample renaturation and eluted at 20°C are plotted with dotted lines. These curves show beside the main peak a dominant second population which is thought to be misfolded proPMII, as an IEF showed (Fig. 3).

Figures 5 and 6 show that proPMII detached from HiTrap Q resin at around 175 (buffer D plus GSSG/GSH) to 195 mM (normal buffer D) NaCl. The second population got eluted at around 350 (buffer D plus GSSG/GSH) to 380 mM NaCl (normal buffer D). The longer the gradient was chosen the better the two populations got separated.

Activation and Concentration

proPMII activates itself at low pH (4.5) by cleavage of the N-terminal pro-sequence. Following to the old protocol, the incubation time for activation was 2 h at 37°C . After activation, PMII degrades in a later step the hydrolysed pro-sequence and other impurities, mainly misfolded protein populations. This is impressively visible on the right SDS gel in Fig. 7. Within 2 h , nearly all

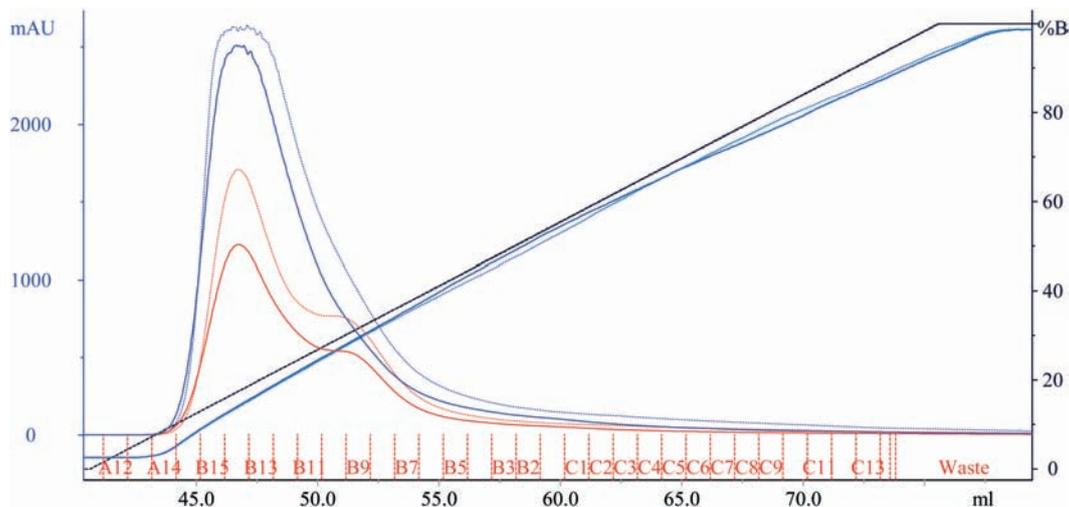


Figure 4: Comparison of HiTrap elution profiles run at 4°C. Straight lines: renaturation buffer containing 1 mM GSSG and 0.5 mM GSH, dotted line: normal renaturation buffer. Absorbance at 280 nm blue, absorbance at 254 nm red. The black broken line represents the actual percentage of elution buffer (called %B in right scale). The difference of peak-top absorption had arisen of loading different amounts of material (normal renaturation buffer contained 35% less material). There is no qualitative difference between the two renaturation buffers.

impurities as well as the pro-sequence fade away. After 15 min (band ② in the right SDS gel), the cut pro-sequence is easily visible.

Complete activation and degradation of impurities were finished after about 2 h at RT. Longer activation - up to one month at 4°C - showed no loss of material. Thus, the activation of PMII per se can be used as purification of the sample when run long enough. Note that the band ② of the left SDS gel in Fig. 7 did not attenuate during 4.5 h. Compared to this main-peak sample, bands ③ and ④ in Fig. 9 show an activation of an adjacent peak pool. Therein only a small amount of proPMIV could be activated while the rest got degraded during activation.

After activation, the small degraded fragments have to be separated from the active protein. This task might be fulfilled by either running an affinity column (H4848) or a SEC (Superdex 16/60). The H4848 would have the advantage of eluting the protein in 20 mM Tris pH 8. This pH first improves the solubility of PMII during a later concentration step and inactivates the protease, thereby possibly slowing auto-degradation.

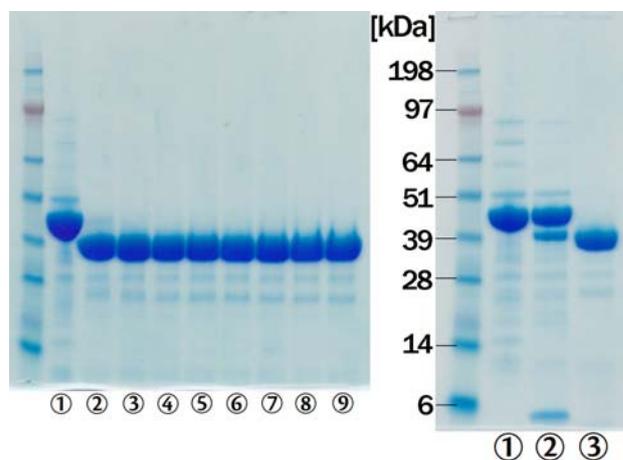


Figure 7: SDS-PAGE monitoring the activation of a HiTrap main-peak sample expressed in Mm(2/4), renatrated in buffer D containing 1 mM GSSG and 0.5 mM GSH. Bands from left to right: MW marker, ① before activation, ② after 0.5 h, ③ 1 h, ④ 1.5 h, ⑤ 2 h, ⑥ 2.5 h, ⑦ 3 h, ⑧ 3.5 h, and ⑨ 5 h. In the right SDS the two populations of proPMII, active form and the cut pro-sequence are visible after 15 min of activation at RT (②). After 2 h (③) the activation appears to be complete.

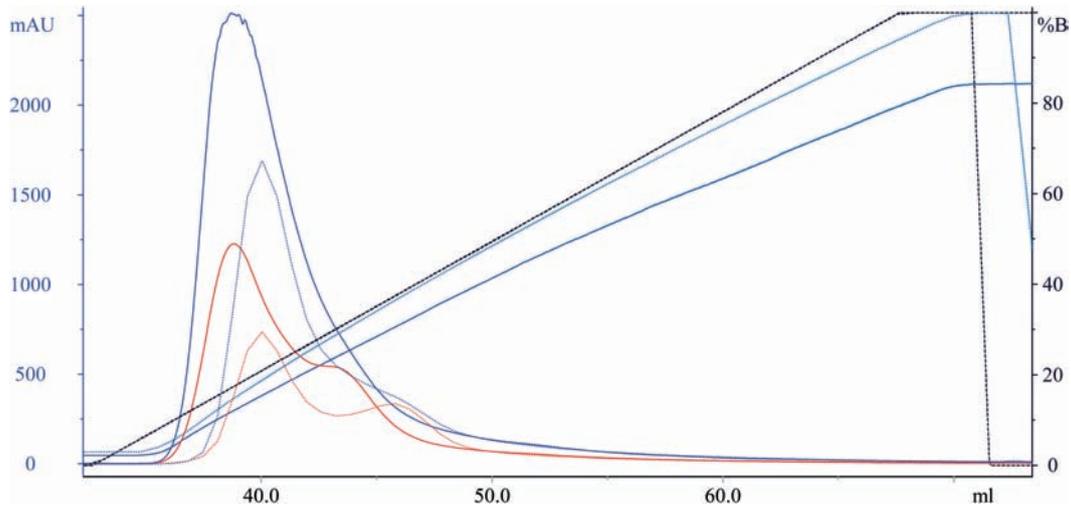


Figure 5: Comparison of HiTrap elution profiles run at 4 and 20°C loaded from normal renaturation buffer. Straight lines: 4°C, dotted line: 20°C. Absorbance at 280 nm blue, absorbance at 254 nm red. The black broken line represents the actual percentage of elution buffer (called %B in right scale) The difference of peak-top absorbtion had arisen of loading different amounts of material (the sample at 20°C contained 38% less material). At 20°C the elution profile shows clearly a second population eluted later then the main peak. Especially at 254 nm the second (misfolded) proPMII population is visible.

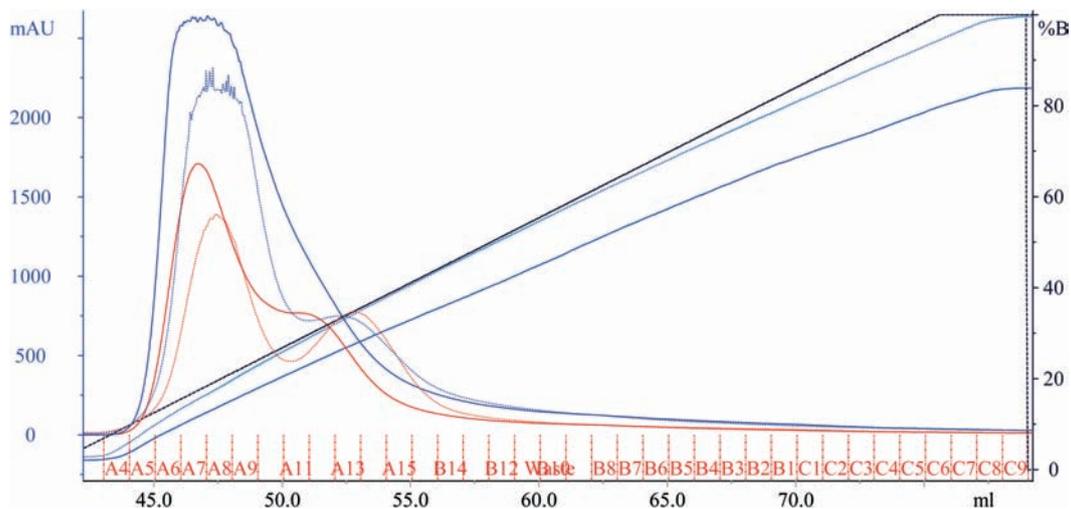


Figure 6: Comparison of HiTrap elution profiles run at 4 and 20°C eluted proPMII in renaturation buffer containing 1 mM GSSG and 0.5 mM GSH. Straight lines: 4°C, dotted line: 20°C. Absorbance at 280 nm blue, absorbance at 254 nm red. The black broken line represents the actual percentage of elution buffer (called %B in right scale) The difference of peak-top absorbtion had arisen of loading different amounts of material (the sample at 20°C contained 20% less material). Renaturation buffer at 20°C shows at 280 nm a dominant peak shoulder.

We decided to skip the H4848 because of several practical reasons. First of all H4848 showed no better purification of the sample compared to SEC. Additionally, the H-4848 inhibitor is expensive and the column has to be prepared and packed manually. In contrast to the H4848, size exclusion columns have two advantages: In the elution profile aggregates as well as peptide fragments are easily detectable and SECs have a more convenient handling: the sample is directly loaded via super-loop and flowed through within 2 h without time consuming elution and regeneration steps.

3.1.2 PMIV

Test expression and glycerol stock forming were run according to the main protocol.

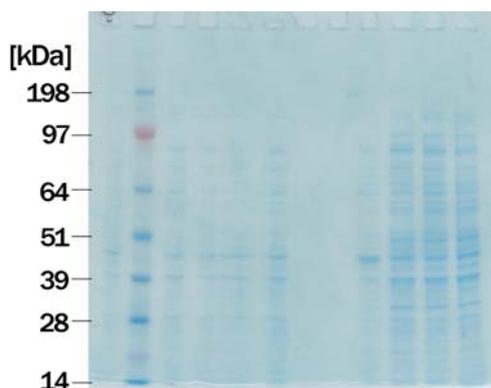


Figure 8: SDS-PAGE controlling the test expression of proPMIV in four clones. Bands: before induction ① clone A, MW marker, ② clone B, ③ clone C (two bands), after induction ④ clone A, ⑤ clone B, ⑥ clone C, ⑦ clone D. Clone A was used for expression as this clone used most of the energy to express proPMIV.

SDS-PAGE in Figure 8 showed significantly slower bacterial growth compared to PMII. Interestingly, clone A grew much faster if expressed afterward in Tm or Mm.

Purification

Pellet purification of proPMIV expressed in Tm was run according to an ‘intermediate’ protocol.

This means that the protocol was already optimized in certain steps whereas some values were not yet modified. The differences to the final protocol in section 2.2 were added in parentheses.

From PMIV only one batch (a pellet from 1 l of Tm) was purified. This procedure is briefly summarized in the following section:

Around 50 ml of buffer A was used to resuspend the pellet at RT under stirring. French press was run two times and rinsed with buffer A to a final volume of 110 ml. After centrifugation of buffer A in 3 tubes, the pellet showed less impurities and supernatant was a little muddier compared to PMII Tm samples. SDS-PAGE of pellet purification and activation is shown in Fig. 9.

After dilution of proPMIV in 250 ml buffer C (containing 8 M urea), the sample was diluted with P960 at 4°C, 0.2 ml/min into 10 l buffer D and loaded afterwards on a 5 ml-HiTrap with 2 ml/min. This lasted 3.5 days.

The elution profile with a gradient over $30 \times C.V.$ (150 ml, 2 ml/min) is shown in Fig. 10. Main peak sample was activated and directly loaded on SEC, where PMIV passed the SEC after the same volume as PMII (86 ml).

3.1.3 PMI

Compared to the other two PMs, PMI was cloned with a (His)₁₀-tag. From PMI only one batch (a pellet from 1 l of Tm) was purified. This procedure is briefly summarized in the following section:

Around 30 ml of buffer A was used to dilute the pellet thawed on ice. As the pellet did not resuspend properly and stayed slimy, it was vortexed and afterward stirred at 4°C for 10 min. French press was run two times and rinsed with buffer A to final volume of around 130 ml.

After centrifugation of buffer A, subsequent refilling and centrifugation with buffer B, pellets were dissolved in 500 ml of buffer C.

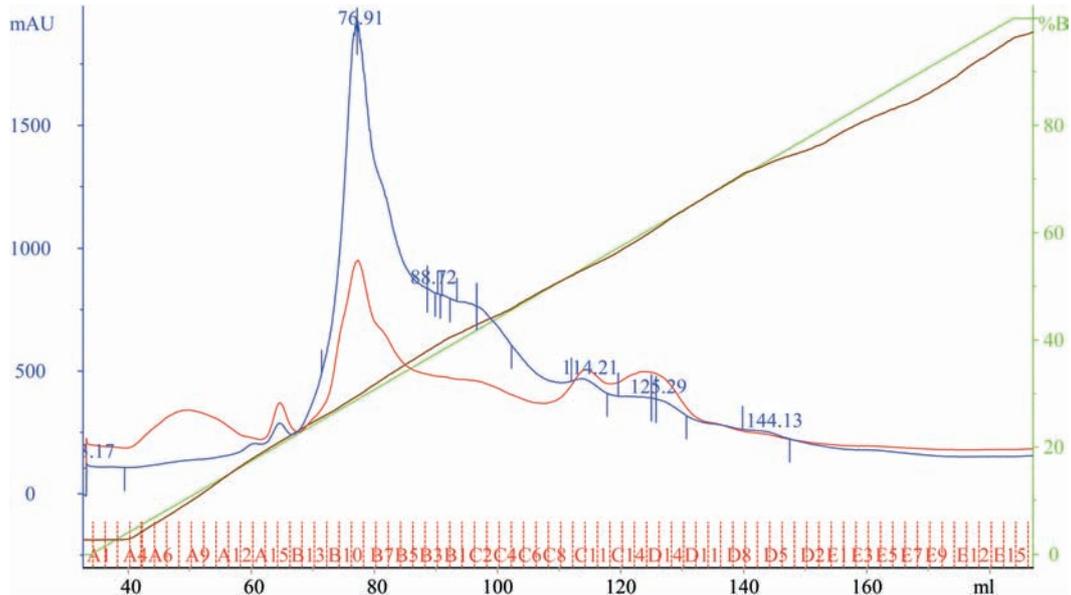


Figure 10: HiTrap elution profile of proPMIV expressed in Tm. Absorbance at 280 nm blue, absorbance at 254 nm red. The green line represents the actual percentage of elution buffer (called %B in right scale), the conductivity is drawn in brown.

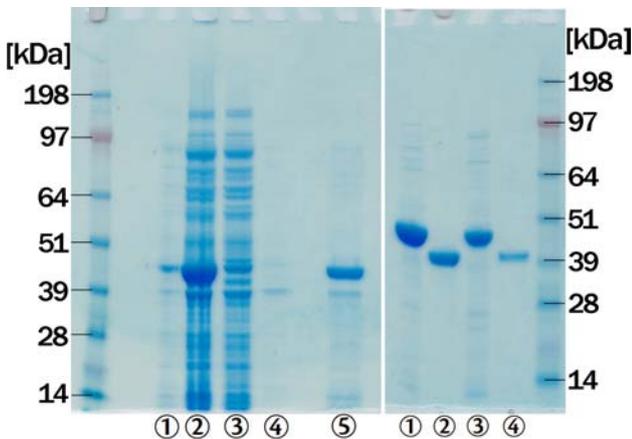


Figure 9: Left: Pellet purification of proPMIV expressed in Tm. Bands from left to right: MW marker, ① -20°C stored pellet, ② after two rounds of french press, ③ supernatant buffer A after centrifugation, ④ supernatant buffer B after centrifugation, ⑤ diluted in buffer C. Right: Activation of HiTrap main peak (before ① and after 90 min ②) and adjacent peak (before ③ and after activation ④). During activation of adjacent peak pool, the misfolded protein running at the same height is degraded by active PMIV.

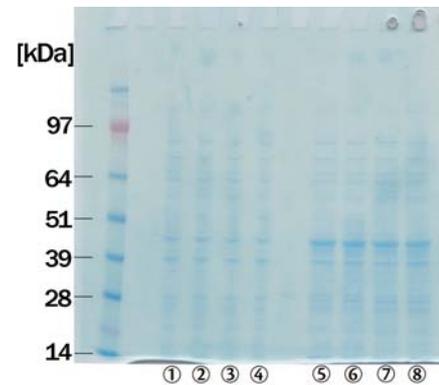


Figure 11: SDS-PAGE controlling the test expression of PMI in four clones. Bands: MW marker, before induction ① clone A, ② clone B, ③ clone C, ④ clone D, after induction ⑤ clone A, ⑥ clone B, ⑦ clone C, ⑧ clone D. Clone C was used for expression as this clone grew faster than the other (data not shown).

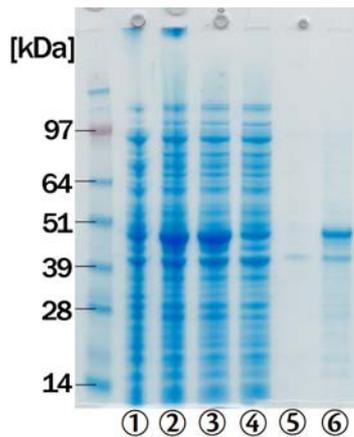


Figure 12: SDS-PAGE controlling the pellet purification of proPMI. Bands: MW marker, ① - 20°C stored pellet, ② first and ③ second run french press, ④ supernatant buffer A after centrifugation, ⑤ supernatant buffer B after centrifugation, ⑥ proPMI dissolved in buffer C

SDS-PAGE showed material running at the same height as proPMI in supernatant A. 130 ml supernatant of buffer A were directly loaded on a HisTrap HP (GE Healthcare, 1 ml bed volume) with 1 ml/min. HisTrap binding buffer contained 20 mM Tris/HCl pH 8.0; 500 mM NaCl; 20 mM imidazole (buffer was adjusted to pH 8.07 with HCl, final conductivity around 45 mS). Supernatant buffer A after centrifugation was loaded on HisTrap o/n.

HisTrap was washed with 20×C.V. binding buffer and proPMI eluted with a gradient over 25 × C.V., 1 ml/min to 100% elution buffer (20 mM Tris/HCl pH 8.0; 500 mM NaCl; 500 mM imidazole (adjusted with HCl to pH 8.04) conductivity of 48 mS). The elution profile shown in Fig. 13 displayed no PMI in supernatant A. Therefore it was assumed that proPMI was also expressed in inclusion bodies.

A small sample (20 ml) of proPMI dissolved in buffer C was renatured with 0.1 ml/min into 500 ml buffer D. Afterward, the sample was loaded on a 1 ml-HisTrap (as the backpressure rose constantly, HisTrap was eluted after loading 200 ml buffer D).

Elution HiTrap elution profile visible in Fig. 14 showed a massive amount of misfolded protein. There was no time to optimize the refolding of PMI.

3.2 Crystallization

3.2.1 Control of Protein Purity

The quality of purified PMII according to the optimized protocol was tested by co-crystallization with a known in-house inhibitor (ACT-061133, deposited PDB code: 2BJU). PMII (8.96 mg/ml; 5 mM Na/K P_i pH 6.5; 100 mM NaCl) was mixed in a 1:20 ratio with ACT-061133 (30 mM in tert.-butanol, stored at -20°C for 5 years) and pipetted with a drop ratio of 3:1 (v/v) protein to reservoir solution. Crystals grew in pH 6.5 and 18% PEG after 3-4 weeks and showed the known P₂₁₂₁₂ crystal form with an according resolution around 1.7 Å. This corresponds to previous quality.

3.2.2 pH Shift During Crystal Growth

Native crystals of the first batches in which the SEC running buffer contained 20 mM Na/K P_i grew within 3 to 4 weeks. All of them showed strong diffraction up to 2.1 Å. Because of huge mosaicity and twinned diffraction pattern crystals were not suitable for processing. Fig. 15 shows the diffraction image of a crystal from the first batch.

An experiment showed that the pH - using a drop ratio of 2:1 (v/v) protein (20 mM Na/K P_i pH 6.5) to reservoir solution (0.1 M sodium citrate) - lay around 0.37 to 0.41 above the desired pH of the reservoir solution (measured pH values 3.0; 3.25; 3.75). This problem was detected as crystals showed less mosaicity if soaked with cryo buffer in which the pH was shifted +0.25 (relative to the pH of the reservoir solution).

Despite shifting the pH of the cryo buffers by +0.25 or +0.5 units, crystals from the first

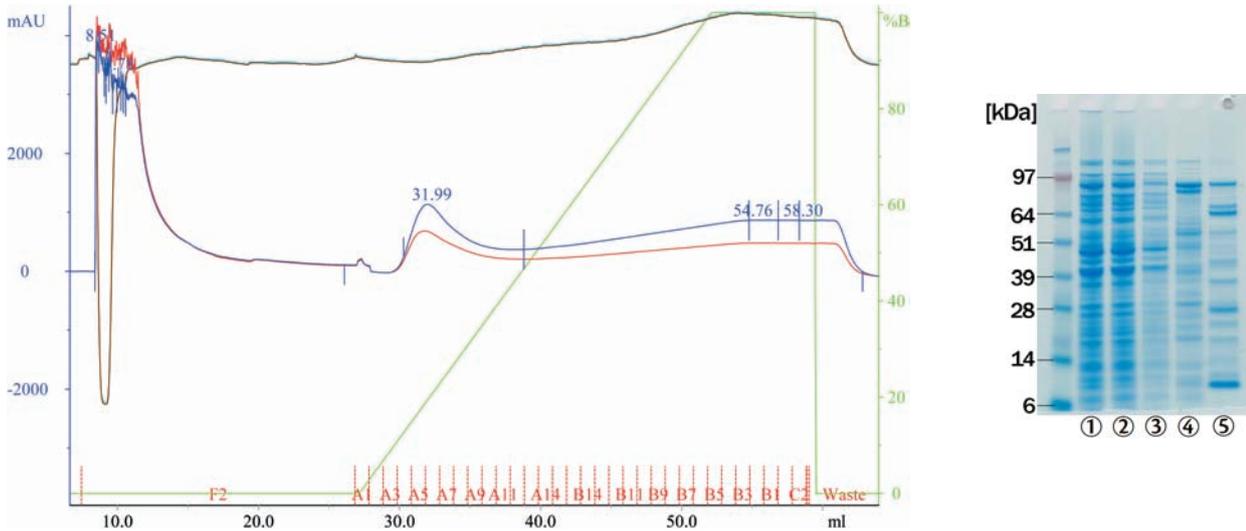


Figure 13: HisTrap elution PMI expressed in Tm. Left image: Elution profile, absorbance at 280 nm blue, absorbance at 254 nm red. Green line represents actual percentage of elution buffer (called %B in right scale), the conductivity is drawn in brown. Right image: SDS-PAGE controlling the loading and elution of HisTrap. Bands: MW marker, ① HisTrap load, ② HisTrap flowthrough during load, ③ 1. ml of wash (visible in elution profile), ④ 20 × C.V. wash, ⑤ pool of peak fractions A4-A9.

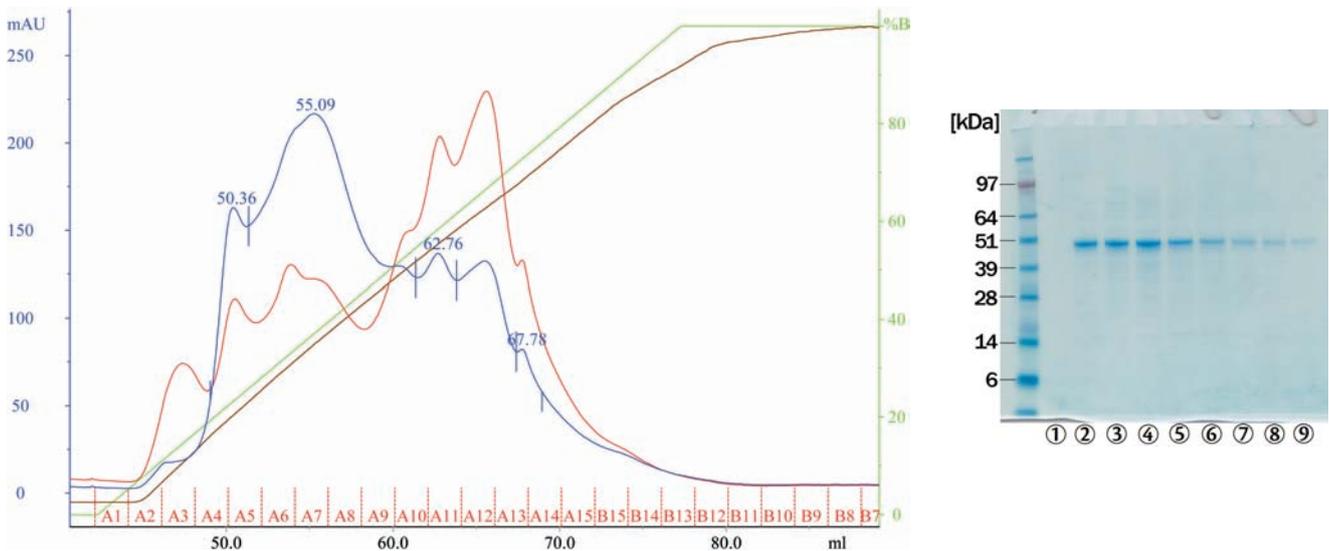


Figure 14: HiTrap elution PMI expressed in Tm. Left image: Elution profile, absorbance at 280 nm blue, absorbance at 254 nm red. Green line represents actual percentage of elution buffer (called %B in right scale), the conductivity is drawn in brown. Right image: SDS-PAGE controlling the loading and elution of HiTrap. Bands: MW marker, ① fraction A3, ② A5, ③ A6, ④ A7, ⑤ A9, ⑥ A10, ⑦ A11, ⑧ A12, ⑨ A13

batches still showed huge mosaicity problems. To ensure that the pH stayed constant during the growth of a crystal, the concentration of the SEC running buffer was reduced to 5 mM Na/K P_i pH 6.5. Crystals with co-crystallized inhibitor ST47 grew thereby within 24 h , other crystals within two to three weeks. Crystals which grew even slower showed huge mosaicity problems and could not be processed. Soaking experiments using native ST43 and ST47 crystals with more affine inhibitors were not successful.

Acidification for Incubation

Inhibitors were added in a ratio of 10:1 or 20:1 (v/v) or equimolar amount directly to active protein in SEC running buffer (5 mM Na/K P_i pH 6.5; 100 mM NaCl). To ensure a complete protonation of the inhibitors, the pH was shifted before incubation with 1:400 (v/v) of 1 M HCl to a final pH 4.2. This shift led to much smaller pellets in the subsequent centrifugation step (before pipetting the samples were centrifugated to remove unbound inhibitors and suspended solid) which may have been caused by a higher occupancy of the active site. Fig. 16 shows a diffraction image of a crystal grown with acidified solution. Compared to the first batches (Fig. 15) the problem of mosaicity had been solved successfully.

With the acidification during incubation a third inhibitor (mzd_1) crystallized. We assume that further investigations in acidification provide even more stable complexes and might lead to more crystals.

Screening

Initial screening covered nearly the whole pH range between 3.0 and 5.75 (0.25 increments) and 10-30% PEG 4K (2% increments). In the beginning each screen was expanded by adding $30\text{ }\mu\text{l}$ 4 M NaCl as additive to reservoir solution. This led to a reservoir concentration of

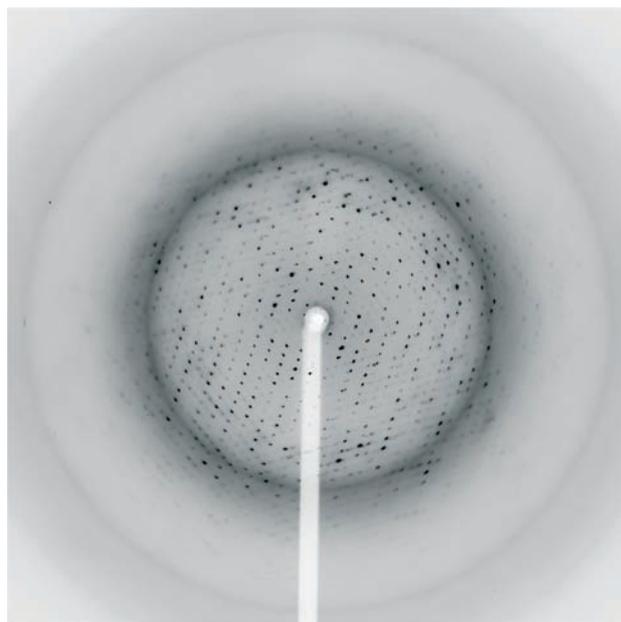


Figure 15: Diffraction image of PMII-ST47 complex. This crystal belonged to the first batch. SEC running buffer contained 20 mM Na/K P_i , crystal grew in pH 3.0 and 20% PEG 4K. Cryo buffer 0.1 M sodium citrate pH 3.0; 35% PEG 4K. Detector distance 220 mm , $\lambda = 1.0001\text{ \AA}$

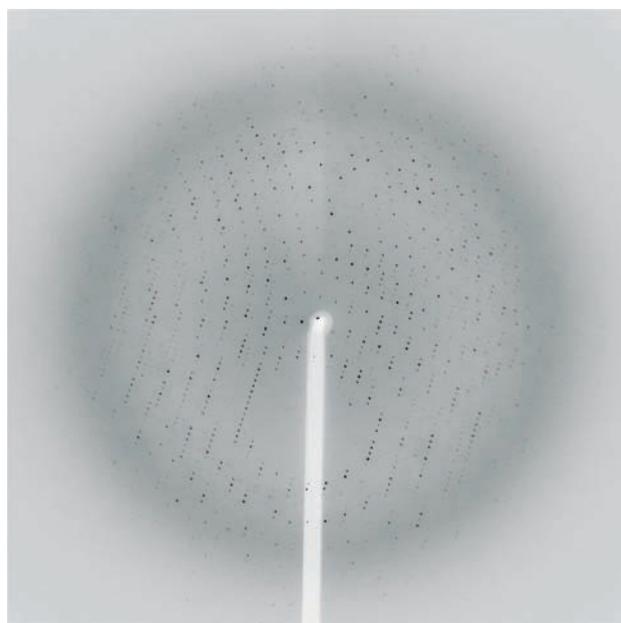


Figure 16: Diffraction image of PMII-mzd_1 complex. PMII was incubated in 5 mM Na/K P_i and acidified with HCl. Inhibitor was added to equimolar amount, crystal grew in pH 4.0 and 20% PEG 4K. Cryo buffer 0.1 M sodium citrate pH 3.75; 25% PEG 4K; 20% glycerol. Detector distance 280 mm , $\lambda = 1.0097\text{ \AA}$

522 *mM* NaCl (200 μ l reservoir plus 30 μ l 4 *M* NaCl). By a drop ratio of 2:1, the final NaCl concentration in the droplet reached theoretically 722 *mM* in equilibrium with the reservoir (protein buffer contained in all screens 100 *mM* NaCl). Droplets (2:1) without additive reached theoretically 200 *mM* NaCl as final concentration.

Each screen was pipetted in parallel with and without NaCl as additive. From the beginning we easily could observe that the final NaCl concentration was too high: Plates with additive showed precipitation in all conditions whereas without additive the same conditions stayed clear. Therefore in later batches the additive volume was reduced to 20 μ l (364 *mM* NaCl in reservoir) leading to a final theoretical droplet concentration of 564 *mM* (2:1) and 664 *mM* (3:1).

The influence of 20 μ l NaCl additive by a drop ratio of 2:1 was impressive mainly with two inhibitors (screens with 8.66 *mg/ml* LBM-09-2 and 9.85 *mg/ml* PMII-mzd_1 complex). In total 8 screens (mainly covering pH 2.75 to 3.75) the additive shifted the phase boundary to around 6-10% less PEG 4K (e.g mzd_1 from 22% PEG 4K to 14% at pH 3.25) indicating that NaCl pushes the precipitation.

Especially in screens of high pH and high PEG 4K the supplemental NaCl led to heavy precipitation in all conditions. Therefore, the addition of any NaCl was stopped to save material. In return, the ratio of 2:1 was changed to 3:1 (v/v) protein to reservoir to increase slightly the concentration of NaCl stemming from SEC buffer. This new ratio was successful with inhibitor ST47 but not with ST43 (which crystallized in 2:1 but precipitated in 3:1 screens).

None of the screens above pH 5 led to diffracting crystals. We assume that above a critical pH the inhibitors get deprotonated and therefore the complex is not stable anymore. Especially the compounds synthesized by Diederich *et*

al. (ETH) contained an electron withdrawing sulfonic acid beside the protonated secondary amide. From these compounds only one complex could be measured and processed (complex with inhibitor mzd_1 scattered to 2.7 Å).

Interestingly the most affine inhibitors did not crystallize. Two of them have values of K_i of 30 and a third inhibitor 100 *nM*. In the first batches these complexes stayed completely clear. After removing glycerol from the SEC buffer (not mentioned in the main protocol or in the discussion) and reducing the SEC running buffer to 5 *mM* Na/K P_i pH 6.5; 100 *mM* NaCl phase boundaries were noted between pH 3-3.75; 14-24% PEG 4K. An addition of 20 μ l 4 *M* NaCl to reservoir led in nearly all conditions to precipitate, except for those at pH 3 and low PEG 4K. After designing a new screen with even lower pH and less PEG 4K (pH 2.5-3.25 and 10-20% PEG 4K) phase boundary was furthermore observable but at the boundary no crystals grew.

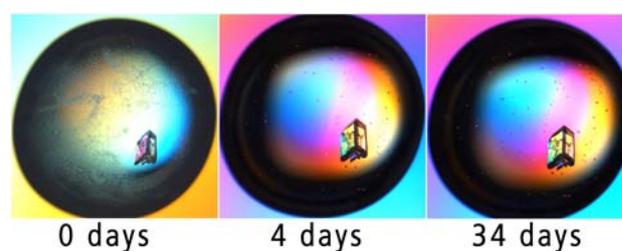


Figure 17: Observation during one month of a PMII-ST43 complex twinned crystal growing in pH 4.5; 24% PEG 4K.

Crystals of the PMII-ST43 complex grew within 2 weeks in pH 4.75; 24% PEG 4K. PMII (10.07 *mg/ml*; 5 *mM* Na/K P_i; 100 *mM* NaCl) was mixed by 20:1 (v/v) ratio with ST43 (30 *mM* in DMSO). Droplets were pipetted with a drop ratio of 2:1 (v/v) reservoir to protein solution. Crystals grew as huge polarizing plates, most of them twinned. During growth of crystals precipitate dissolved reversibly. Figure 17 shows a twinned crystal growing at pH 4.5; 24% PEG 4K. During the first days, crystals grew while the pre-

precipitate dissolved. In a later phase, continuous forming of microcrystals in the droplet led to a slower growth of the crystal.

Crystals of the PMII-ST47 complex grew within 24 h in pH 3.0; 20% PEG 4K. PMII (10.07 mg/ml; 5 mM Na/K P_i; 100 mM NaCl) was mixed with ST47 (30 mM in DMSO) at a 20:1 ratio and pipetted with a drop ratio of 2:1 (v/v) reservoir to protein. Crystals grew as huge polarizing plates shown in Fig. 18.

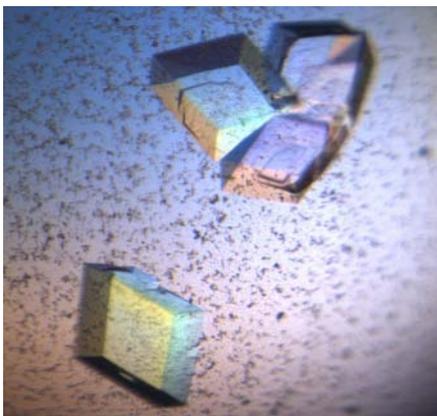


Figure 18: Image of PMII co-crystallized ST47. These crystals grew in pH 3.0; 20% PEG 4K and were used for solving the structure.

3.3 Structure Determination

Only PMII complexes could be processed. All three inhibitors are non-peptidomimetic chiral compounds synthesized by the group of Prof. Dr. Gerhard Klebe.

3.3.1 Compounds

Data collection and structure refinement statistics are listed in Table 2.

Inhibitor ST43

Assignment to space group P1 (No. 1) with the cell constants $a = 63.380$ $b = 74.220$ $c = 74.840$ Å, $\alpha = 113.98$ $\beta = 106.08$ $\gamma = 104.89$ yielded an R_{merge} of 4.1 (26.4 for the highest

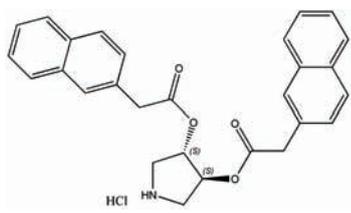
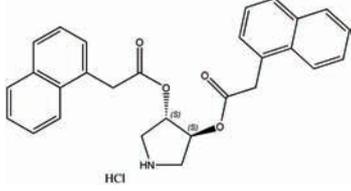
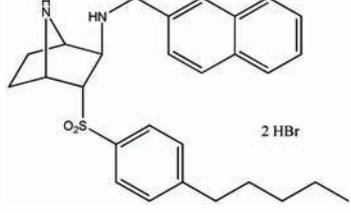
Name	Chemical Structure	K_i
ST43		1 μ M
ST47		100 nM
mzd.1		680 nM

Table 1: Listing of the chemical structures of inhibitors ST43, ST47 and mzd.1 and the corresponding affinities to PMII.

resolution shell between 2.30 and 1.90 Å). According to the Matthews coefficient calculation, 3 monomers are in the asymmetric unit ($V_M = 2.53$ Å³/Da, 36.9 kDa).

The phase problem was solved by molecular replacement (MR) using CCP4 suite and a pruned PMII (source PDB code: 2BJU) as search model. Molecular replacement ended with a R_{fact} of 29.51% and R_{free} = 30.38%. Initial NCS-based refinement was done without inhibitors. After improving the electron density by adding H₂O energy minimized inhibitors were fit in.

Inhibitor ST47

Assignment to space group P1 (No. 1) with the cell constants $a = 64.000$ $b = 74.510$ $c = 74.590$ Å, $\alpha = 113.47$ $\beta = 106.89$ $\gamma = 104.66$ yielded an R_{merge} of 4.8 (31.5 for the highest resolution shell between 2.30 and 1.90 Å). According to the Matthews coefficient calculation, 3 monomers are in the asymmetric unit ($V_M =$

2.56 Å³/Da, 36.9 kDa).

Data reduction of ST47 complex displayed the same crystal form as ST43. ST47 was solved without MR, but was directly refined with the structure of ST43 (first refinement cycles without inhibitor). After improving the electron density by adding H₂O energy minimized inhibitors were fit in.

Inhibitor mzd_1

The structure of a third complex with inhibitor mzd_1 synthesized by the group of Prof. Dr. François Diederich was solved right before finishing this report. Because of lack of time during finishing the report, this complex (its packing and crystal contacts) will not be mentioned in the discussion about structure determination and will be used only in reference to non-peptidomimetic inhibitors in section 3.5.3.

Assignment to space group P2₁2₁2₁ with the cell constants $a = 94.99$ $b = 110.63$ $c = 134.13$, $\alpha = \beta = \gamma = 90$ yielded an R_{merge} of 6.3 (58 for the highest resolution shell between 2.9 and 2.70 Å).

Data reduction of mzd_1 complex displayed 4 enzymes in the asymmetric unit. The phase problem was solved by molecular replacement (MR) using CCP4 suite and a pruned PMII (source PDB code: 2BJU) as search model.

Inhibitor Name	ST43	ST47	mzd.1
Crystal System	Triclinic	Triclinic	Orthorhombic
Space group	P1	P1	P2 ₁ 2 ₁ 2 ₁
Cell dimensions [Å, °]	$a = 63.380$ $b = 74.220$ $c = 74.840$ $\alpha = 113.98$ $\beta = 106.08$ $\gamma = 104.89$	$a = 64.000$ $b = 74.510$ $c = 74.590$ $\alpha = 113.47$ $\beta = 106.89$ $\gamma = 104.66$	$a = 94.99$ $b = 110.63$ $c = 134.13$ $\alpha = \beta = \gamma = 90$
Wavelength [Å]	1.000	1.000	1.000
Resolution [Å]	35.95 – 1.98 (2.30 – 1.90)	35.68 – 1.98 (2.30 – 1.90)	38.93 (2.9 – 2.7)
Unique Reflections	68445 (23340)	66612 (22206)	38368 (6761)
Redundancy	1.95 (1.90)	1.95 (1.89)	7.37 (7.41)
Completeness [%]	79.6 (62.3)	76.7 (58.7)	97.1 (90.1)
I/ σ (I)	14.61 (3.08)	12.76 (2.89)	26.08 (4.05)
R _{merged} [%]	4.1 (26.4)	4.8 (31.5)	6.3 (58)
Values in parentheses refer to data in the highest resolution shell.			
Resolution [Å]	35.95 – 1.98	35.68 – 1.98	38.93 – 2.71
Reflection used	65022	63280	36449
R _{fact} [%]	20.69	23.26	22.44
R _{free} [%]	26.07	28.78	32.95
Average B-factor [Å ²]	44.54	41.49	63.93
rmsd _{bond} length [Å]	0.013	0.019	0.010
rmsd _{bond} angle [°]	1.556	1.858	1.590

Table 2: Data collection and structural refinement statistics of co-crystallized PMII with inhibitor ST43, ST47 and mzd.1.

3.4 Structure Description

PMII shows the typical fold of eukaryotic aspartic proteases like mammalian (pepsin, renin, gastric-sin and cathepsin D) and fungal (endothiapepsin and rhizopuspepsin) enzymes^[17] (see Fig. 19 for overview).

The active enzyme is formed by a single chain consisting of 329 amino acids. It is folded into two topologically similar N and C-terminal domains^[18], each of which contributes a catalytic aspartate residue to the active site. The two domains contact each other along an intersection in which the binding cleft containing the catalytic dyad, Asp34 and Asp214, is buried. Behind the active site, a six-stranded interdomain β -sheet anchors the two domains. This structure is composed of 5 β -strands which were stabilized at one end by an N-terminal sheet.

The N-terminal domain consists mainly of 7 β -sheets which form a rib cage. At the top end of the "costal arch" a distinctive single β -hairpin, known as flap, forms directly above the binding cleft a finger-like structure and interacts with a bound substrate.

3.4.1 Interfaces of the Crystal Packing

Both complexes PMII-ST47 and -ST43 reveal three enzymes in the asymmetric unit (called chain A to C, see Fig. 20). In the asymmetric unit two different crystal contacts were found:

The first contact between chain A and B consists of several H-bonds between the two C-domains. Thereby 3 amino acids in a row form a short antiparallel β -sheet structure with the neighboring PMII: Lys163A-Thr165A and Tyr184B-Leu182B with totally 3 H-bonds. Directly beside this contact, a fourth H-bond between the amide group of the side chain Asp328A and the backbone of Glu262B completes this polar crystal contact. Its contact area between chain A and B in the unit cell is computed to 560.6 \AA^2 (ST43, 3.7% of total

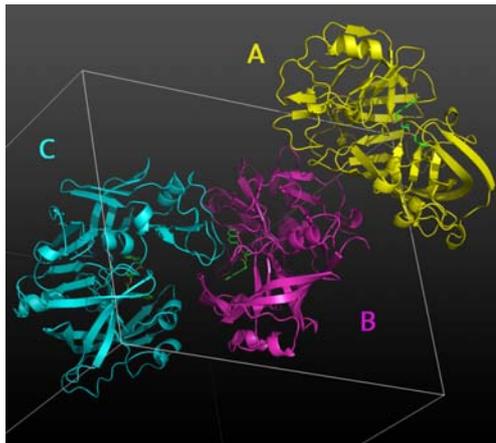


Figure 20: Crystal packing of PMII-ST43 complex. The cartoon shows the unit cell containing chains A (yellow), B (magenta), and C (cyan) with the respective inhibitors in green. Both crystal contacts are visible: hydrophobic contact between chains B and C in the proximity of the binding cleft of chain B, and the polar interaction between chains A and B.

surface) and 549.8 \AA^2 (ST47, 3.6% of total surface).

PISA^[19] found several supplemental interactions. The local density at the crucial amino acids is not accurate enough to determine the exact position: There might be weak H-bonds between Ser262A/Asn250A and Arg64B and a second weak interaction between Asn145A/Asp316A and Lys72B. All four interactions are more than 3.2 \AA long and therefore are mentioned here only for the sake of completeness.

The contact interface between chain A and B contributes only little to the entire solvation free energy (ΔG) loss upon formation of the interfaces. This is because the contact interface between chain A and B involves no hydrophobic interactions. The major driving force of crystal growth arose upon formation of the second contact interface (between chains B and C in unit cell). PISA calculated 757.1 \AA^2 (ST43, 5.0% of total surface) and 733.8 \AA^2 (ST47, 4.5% of total surface) for the hydrophobic interface between B and C in the unit cell.

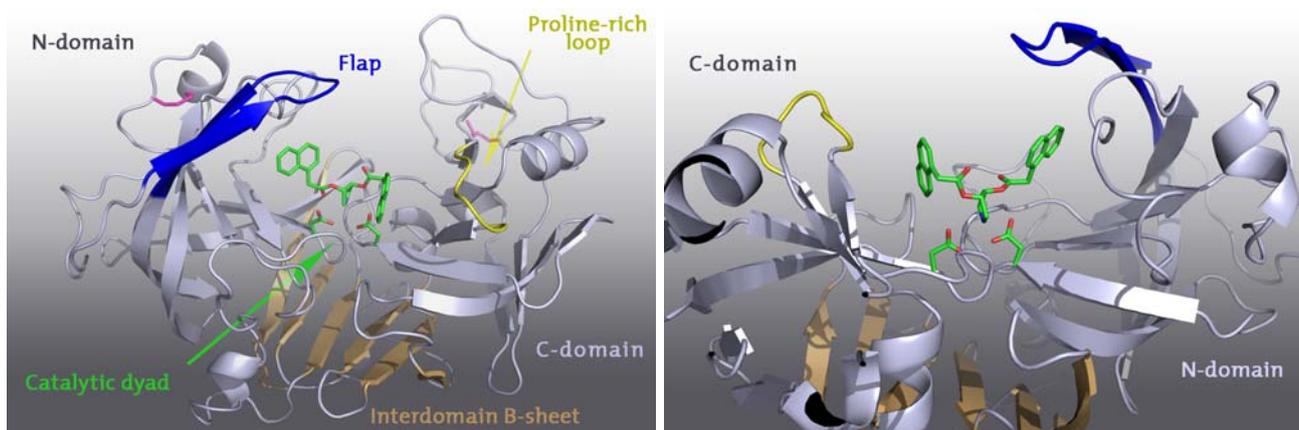


Figure 19: Ribbon diagram of the complex of PMII with ST47 showing disulfide bridges (magenta), catalytic dyad with inhibitor (green), flap (blue), interdomain β -sheet (brown), and proline-rich loop in yellow. Right image: 180° turned and slabbed ribbon diagram, same coloring.

The interaction between chain B and C was already found in other protease structures (PDB codes: 2BJU, 2IGX and 2IGY). The mainly hydrophobic character explains the high contribution to ΔG (PISA calculated more than 50% of the total loss of solvation Gibbs energy due to formation of this hydrophobic interface). It has arisen from a special conformation of hydrophobic loop interacting in the proximity of the neighboring binding cleft and the bound inhibitor:

The space at the far side of the pyrrolidine ring is filled by a loop of the coterminous chain C interacting directly with the inhibitor bound in the active site of chain B. The interacting loop is natively positioned at the surface but reveals within its 10 amino acids an extremely atypical hydrophobic pattern: V²³⁶IKVPFLPFYV²⁴⁶. Due to the hydrophobic properties, this loop is thought to interact with the heterodimeric hemoglobin by partially disrupting the intermolecular interactions, thereby making the quaternary structure susceptible for degradation by PMs.

PISA found 3 residues of this loop staying in hydrophobic contact with the inhibitor: Pro240C and Phe241C at the very end of the loop and the aliphatic chain of Lys238C.

Already two amino acids away from Lys238C

- around 16 Å from the hydrophobic interaction with the inhibitor - the same loop forms a hydrophobic pocket in which the proline-rich loop of chain B fits in. Thereby mainly Phe294B, Pro295B, and Val296B of the proline-rich loop are involved in interactions with the pocket-forming Leu227C, Leu231C, Val236C, Tyr245C, and Val246C.

Due to the hydrophobic interactions at the far side of the inhibitor, the loop of chain C gets sandwiched between the inhibitor and the flap of chain B, consequentially also interacting with the flap. As a result, the hydrophobic loop forms a 2.92 Å long H-bond from the backbone of Phe241C to the backbone of Val78B (tip of the flap). This H-bond fixes the flap in a well defined position and completes this crystal contact. It is shown in Fig. 21. Beside these interactions PISA found one additional H-bond between Gly134B and Val280C (3.1 Å).

3.5 Structure Comparison

3.5.1 Differences among Chains in Crystal

Although both datasets of ST43 and ST47 could be indexed with rhombohedral lattice (space group R3, No. 146), processing yielded a R_{merge}

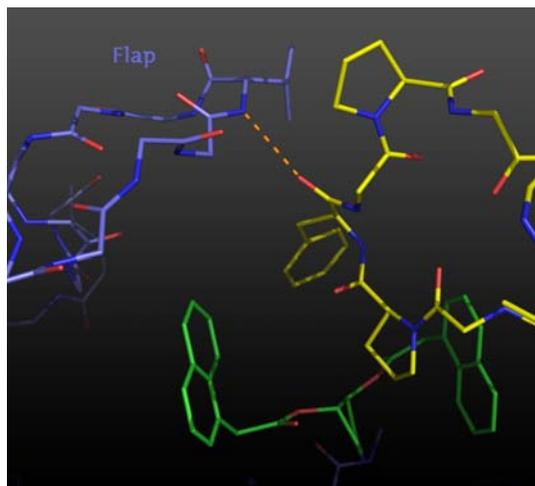


Figure 21: Stick model of the H-bond between the backbone of Phe241C to the backbone of Val78B. The drawn 2.92 Å long H-bond presumably fixed the flap in a highly conserved arrangement found throughout in all chains of ST47 and ST43.

value of 50%. Processing was only possible with triclinic crystal system (space group P1). Slightly differing arrangements at distinct positions among chain A to C may broke the symmetry down to P1.

In the ST47 structure mainly three irregularities were found (all shown in Fig. 22):

1. The N-domain-near naphthyl ring in chain C is turned by 180° still lying in the same plane as the other two naphthyl rings in front of the flap pocket.
2. The ester bond of the C-domain-near naphthyl ring bound to chain B is turned away and oriented toward a H₂O molecule which stays in H-bonding contact (2.7 Å) to Thr217B.
3. An overlay shows striking irregularities only in a flexible loop in the C-domain. This loop contains amino acids His276 to Leu284.

3.5.2 Differences between ST47 and ST43

Both structures show very similar packing and nearly the same cell constants in P1. An overlay

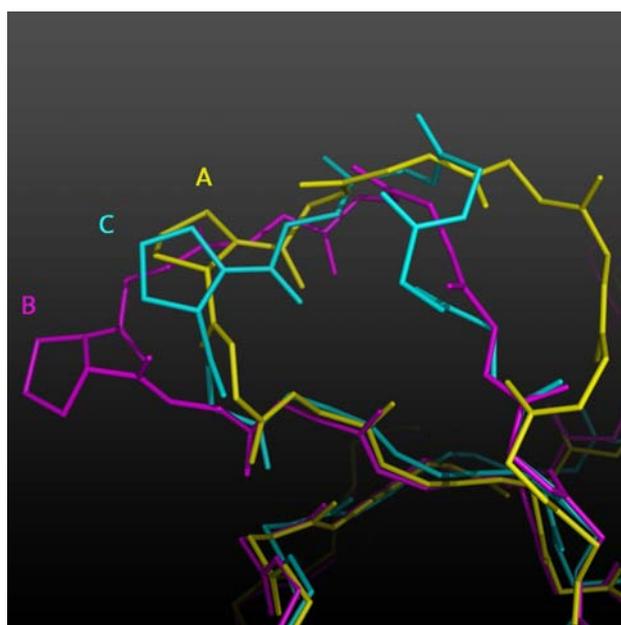
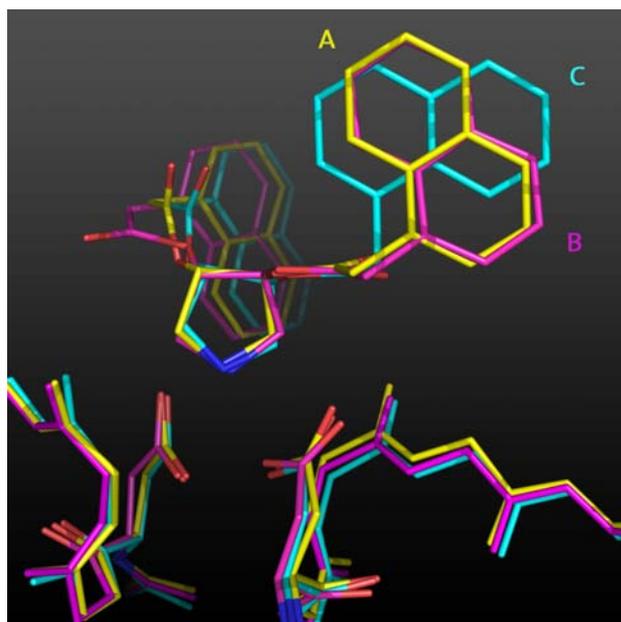


Figure 22: Stick model of superimposed chains A (yellow), B (magenta), and C (cyan) of ST47 showing in the upper image the three inhibitors and the catalytic dyad. The chain C naphthyl ring near the flap pocket is coplanarly flipped and stays in the same region as the other two. The ester bond of the inhibitor bound to chain B is turned away and oriented towards a H₂O (not drawn). The C-domain near naphthyl ring in the background of the image is tightly bound to S1' pocket. The lower image shows the not congruent flexible loop in C-domain (His276 to Leu284) mentioned in the text.

of the two inhibitors ST43 and ST47 displayed the protonated nitrogen of the pyrrolidine ring nearly at same position, indicating a strongly bound inhibitor to the catalytic dyad (not shown). In an overlay, the backbone showed no significant structural differences. Congruent chains display a mean-square displacement of 0.454 Å (without H₂O; all atoms compared). Greater displacements up to 1.3 Å were found in the α -helix like structure directly beside the flap of chain C (Phe111C-Phe120C).

The almost identical 3D structures of PMII complexes indicate that ST43 and ST47 bind to PMII in a very similar manner. Both occupy with one naphthyl ring to the entrance of the flap pocket (N-domain-near). The second aromatic ring interacts with the hydrophobic pocket S1' of the C-domain. This groove serves as an anchor for the first C-terminal residue of the substrate being hydrolyzed by PMII.

The structural analysis showed no significant differences in the binding geometry. According to the binding mode, there is no obvious reason why ST47 binds 10 times better than ST43. We assume that ST47 forms a more relaxed structure in solution and therefore binds with a higher probability to the active site. Another explanation might be the naphthyl ring interacting with the S1' pocket. If α -connected (ST47), this ring extends deeper into the hydrophobic groove, serving thereby as better anchor for the inhibitor.

3.5.3 Active Site and Flap Rearrangement

Both discussed inhibitors are chiral non-peptidomimetic hydrophobic compounds. These compounds interact directly with the catalytic dyad without need of a conserved catalytic H₂O located between the two aspartic acids. The H-bonding partner of the catalytic dyad is 3,4-(s,s)-dihydroxy-pyrrolidine, a heterocyclic amine

ring. Unsubstituted pyrrolidine has a pK_a of 11.27^[20]. The two hydrophobic naphthyl rings are connected to the pyrrolidine by an acetic ester. The naphthyl ring is connected at α (ST47) or β position (ST43) to the acetic acid.

Peptidomimetic Inhibitors

During the first part of the discussion about the active site we will reference to a well known PMII-pepstatin complex published by Binkert *et al.*^[21](PDB code: 1XDH), shown in Fig. 23. Pepstatin is a classical peptidomimetic compound which binds to PMs with not less than 10 H-bonds (not all are visible in Figures 23 and 24). Beside amino acids located at the surface of the binding cleft the tip of the flap provides interactions at the far side of the catalytic dyad. In all complexes with peptidomimetic inhibitors the flap (N-domain) is therefore lowered down, orienting and stabilizing the substrate during cleavage.

Due to the lowered orientation of the flap the space underneath it is completely filled by hydrophobic residues, whereby the bottom and one side of the pocket are occupied by Trp41 and Phe111. The entrance to this pocket is blocked by an inward-turned Tyr77 at the tip of the flap (see Figure 23).

As the two inhibitors ST47 and ST43 provide no H-bonding partners at the far side of the pyrrolidine ring, the flap does not interact with them and stays therefore in a rather open conformation. Due to missing interactions in the proximity of the binding cleft, the tip of the flap (Val78) moves relative to the pepstatin structure 7-8 Å away from the active site. This open flap arrangement entails a new binding mode for non-peptidomimetic inhibitors.

The flap opening has consequences for mainly three amino acids: Trp41, Tyr77, and Phe111. Because of the lost steric inhibition, Trp41 is

able to turn outside the flap pocket along χ_1 by around -120° (looking from C_α to C_β). Due to the flap drift, Tyr77 is getting repressed from its location in front of the pocket by one naphthyl ring. At this position the naphthyl ring near the N-domain side of the inhibitor blocks the entrance of the flap pocket. This allows Phe111 from the backside to slip into the enlarged pocket. The movement of Phe111 inward the pocket draws the backbone about 2 \AA relative to the pepstatin structure. Fig. 24 shows an overlay of the pepstatin and ST47 structure.

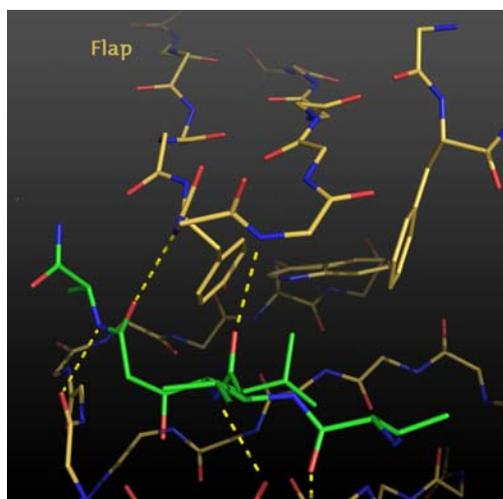


Figure 23: Stick model of PMII-pepstatin complex. The flap coming from behind in the upper half of the image is lowered down and interacts with two hydrogen bonds with pepstatin (green). At the tip of the flap Tyr77 is inward-turned and blocks together with Phe111 at the right side the entrance of the flap pocket. In the background Trp41 lying on the bottom of the pocket is also visible. (PDB code: 1XDH)

For the above-mentioned reasons ST43 and ST47 interact in a different way with PMII compared to peptidomimetic inhibitors, whose compound class is associated with relatively low bioavailability and high production costs^[22]. Especially the latter reason makes non-peptidomimetic compounds more interesting in drug design, as malaria stays a main problem in developing countries.

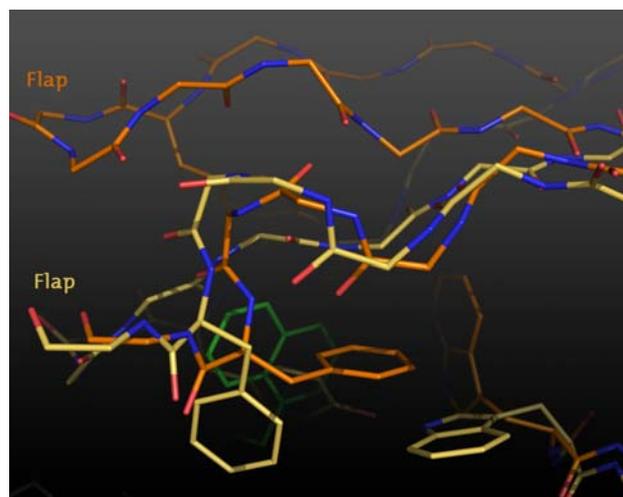
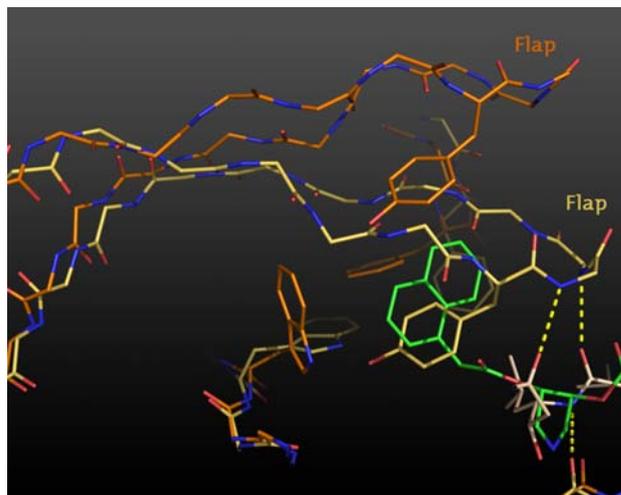


Figure 24: Overlay of stick models of PMII-pepstatin (yellow with pale yellow pepstatin) and PMII-ST47 (orange with green ST47) complex. Due to missing H-bonds with the inhibitor the tip of the flap moves by $7-8 \text{ \AA}$ upwards away from the active site. Thereby Trp41 turns by 120° around χ_1 outside the pocket (under half of the upper image). The position of Tyr77 at the tip of the flap gets thereby occupied by the N-domain-near naphthyl ring of ST47. This position closes the flap pocket which allows the Phe111 in the background of the image to slip inside this pocket and thereby draws the backbone of Phe111 about 2 \AA toward the flap (rearrangement of Phe111 is shown in lower 180° turned image). (PDB code for pepstatin: 1XDH)

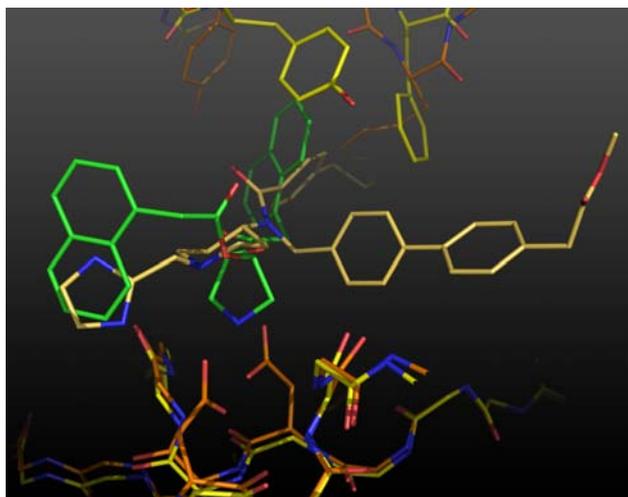
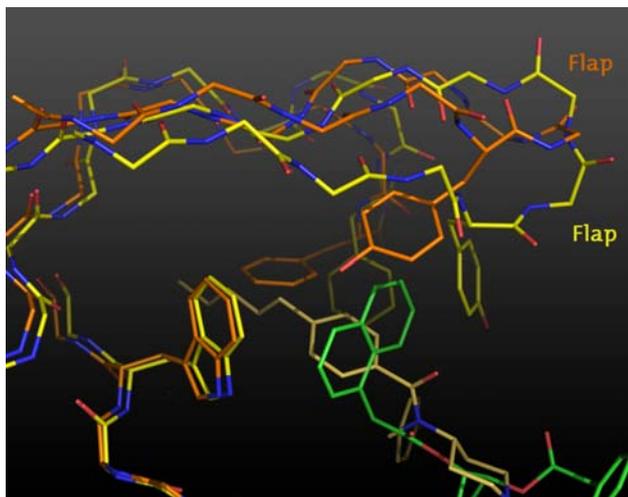


Figure 25: Overlay of stick models of 2IGX.pdb (yellow with pale yellow inhibitor) and PMII-ST47 (orange with green ST47) complex. Upper image shows the flap pocket which gets occupied by the *n*-pentyl chain of the 2IGX-inhibitor, repressing thereby Phe111 from the pocket (background of the image). As ST47 only covers the pocket, Phe111 is able to slip in to the flap-pocket (orange). The lower image shows the flap pocket from the right side and the catalytic dyad (side chains of ST47 model). The C-domain-near naphthyl ring is bound to the same pocket (S1') as the 2-methylene-imidazole unit.

Non-peptidomimetic Inhibitors

There are several non-peptidomimetic structures reported^[22, 23] and deposited (PDB codes: 2BJU, 2IGX and 2IGY). This class is characterized by absent interactions with the tip of the flap. The complete absence of hydrogen bonds between the tightly bound inhibitor and the flap, as well as diminished attractive hydrophobic interactions lead to a opened flap conformation^[22].

The movement of the tip of the flap away from the catalytic dyad opens a cylinder-shaped lipophilic pocket, which is not filled but covered by the naphthyl rings of ST43/47. According to the electron density we assume that this naphthyl ring is not exactly positioned and may reveal two different orientations. D. Bur, L. Prade *et al.* (as well as the third solved PMII-mzd_1 complex) have shown that the flap pocket can be opened completely by repressing Phe111.

Fig. 25 shows one of these (achiral) non-peptidomimetic inhibitors (PDB code: 2IGX) in overlay with ST47. By comparing the binding of the two inhibitors one can see that

1. the 2IGX-inhibitor binds tightly with an *n*-pentyl chain into the pocket, whereas the naphthyl ring of ST47 only covers the entrance of the pocket,
2. the C-domain-near naphthyl ring is bound to the same pocket (S1') as the 2-methylene-imidazole moiety of the 2IGX-inhibitor. This outcome was rather unexpected because this naphthyl ring could reach parts of the S1 and S3 pockets, in which the 4'-substituted biphenyl-methylene unit binds (this unit is visible at the right side in the lower image of Fig. 25),
3. ST43/47 repressed a very strong hydrogen bonded catalytic water molecule (not drawn). This water molecule was not ex-

pelled from its position between the two carboxylates of the catalytic aspartates^[22].

The third solved complex with the inhibitor *mzd_1* binds to the flap pocket in a similar manner as the above mentioned inhibitor 2IGX does (see Fig. 26). The structure of *mzd_1* is shown in Table 1 and reveals also an *n*-pentyl moiety which interacts with the lipophilic flap pocket. Compared to 2IGX, this hydrophobic interaction is weaker because the pentyl chain fills not the whole pocket (visible in an overlay of 2IGX and *mzd_1* inhibitor in lower image of Fig. 26). The second moiety (naphthyl ring) does neither interact with S1' (as ST47 and ST43) nor with parts of the S1 and S3 pockets (as the biphenyl-methylene unit of 2IGX) but remains completely unbound and was therefore not visible in the electron density.

The electron density of the tip of the flap (especially the above mentioned Tyr77 at the entrance of the pocket) is missing (not shown). In contrast to ST43 and ST47, *mzd_1* revealed a completely different packing in which no H-bond between the tip of the flap and the hydrophobic loop was observable. This missing H-bond as well as a weak interaction between the *n*-pentyl unit with Tyr77 may explain the lack of electron density at the tip of the flap.

4 Conclusion and Look-out

4.1 Purification

The protocol was optimized for PMII purification. The modifications targeted on maximal yield in rapid refolding and concentration steps as well as most efficient binding during incubation. As the elution profile in Fig. 10 of a PMIV sample shows, the protocol has to be optimized for PMIV but it works for this protease as well.

The purification protocol had been optimized during the whole master's thesis and therefore

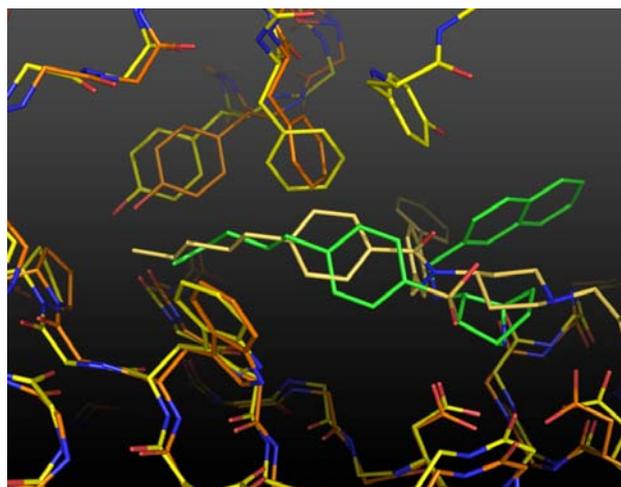
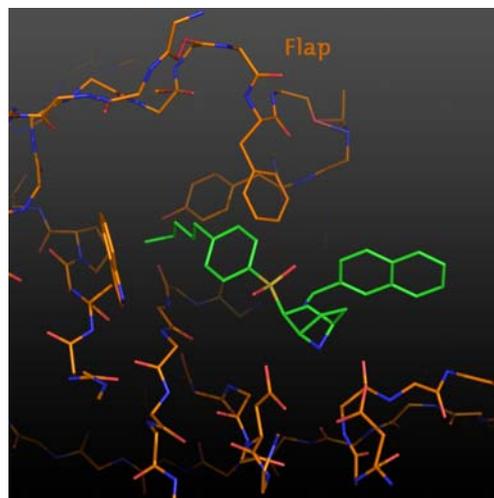


Figure 26: Stick model of PMII-*mzd_1* complex. Compared to ST43 and ST47, *mzd_1* (green) contains an *n*-pentyl moiety which interacts with the flap pocket, thereby repressing Phe111 from the pocket. Its binding mode to the flap pocket is comparable to 2IGX-inhibitor (shown in lower image: PMII yellow with pale-yellow inhibitor). The electron density of the naphthyl ring is not visible, suggesting that this ring does not interact with parts of the S1 and S3 pocket and stays unbound in the vicinity of the binding cleft.

some points still remain unanswered. During writing the report we found several issues that would have been of interest.

We never checked whether the PMs are properly folded within the inclusion bodies. It would be interesting to dilute a sample in renaturation buffer D after french press. Sadly I always forgot to branch off a small sample after centrifugation of buffer B.

For denaturation in buffer C it has to be taken into account that the yield of PMII expressed from *E. coli* grown in different media can differ by factors (yield of Tm is around five times higher than in LB). The diametrical differences between the expression media may be caused by an insufficiently denaturated proPMII or due to too high final urea concentrations in buffer D. Measuring the circular dichroism of denaturated proPMII in buffer C should clarify whether the protein got completely linearized.

As the binding of the inhibitor involves the formation of only two H-bonds this process seems to be mainly entropically driven. Incubation of the inhibitors may be even more efficient when run at higher temperatures.

It would be also worth trying to reduce the NaCl in SEC buffer and increasing the PEG 4K concentration during crystallization instead. This may lead to complexes which grow in other packing than P1. By measuring only 180°, P1 led during data reduction to a rather low redundancy which created noisy electron density maps. During NCS-based refinement the packing caused a lot of work because of 3 enzymes in the unit cell. It would be much easier for processing if the PMII-complexes grew in other conditions and hopefully different crystal packings.

4.2 Inhibitors

The pyrrolidine ring of ST43 and ST47 binds tightly to the catalytic dyad. The structure re-

veals within the unit cell different conformations of the ester bonds of ST47.

Thereby the six ester bonds (3 PMII per unit cell, two different inhibitors) of the C-domain-near naphthyl rings (S1'-pocket) display two different conformations. Four of them show a cis-conformation along the methylene-carbonyl bond (carbonyl double bond is oriented in nearly 90° to Thr217-alcohol), two show gauche conformations (carbonyl oriented toward Phe241, an amino acid of the hydrophobic loop of the neighboring chain). To provide a better binding the ester group might be exchanged by an amide bond, forcing the ester to stay in gauche conformation. The proton of the amide bond will thereby be oriented directly to Asp214 of the catalytic dyad in about 3 Å distance.

Most of the inhibitors contained electron withdrawing groups (sulfuric acid). If these groups were exchanged, the secondary amine serving as binding partner of catalytic dyad might be even better protonated at incubation pH. An exchange of these groups may lead to crystals growing in higher pH and different packing.

Acknowledgment

Many thanks go to Dr. Lars Prade for the scientific steering of the master's thesis, as well as to Daniel Bur for helpful discussions. I gratefully acknowledge the excellent technical assistance of Alain Chambovey and Antoinette Amrein and Silvia Rogers for critically reading the manuscript. Finally I would like to thank Prof. Tilman Schirmer for supervising my master's thesis.

References

- [1] R. L. Krauth-Siegel, H. Bauer, and H. Schirmer, *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION* **44**, 690 (2005).
- [2] J. Wiesner and M. Schlitzer, *Angewandte Chemie* **115**, 5432 – 5451 (2003).
- [3] M. J. Blackman, *CELLULAR MICROBIOLOGY* **10**, 1925 (2008).
- [4] M. Madigan, J. Marinko, and J. Parker, *Brock Mikrobiologie*, volume 9, Spektrum, Akad. Verlag, 2000.
- [5] C. Berry et al., *FEBS LETTERS* **447**, 149 (1999).
- [6] R. Banerjee et al., *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA* **99**, 990 (2002).
- [7] J. HILL et al., *FEBS LETTERS* **352**, 155 (1994).
- [8] Swiss Light Source, *go.com script to auto-index (LABELIT)*.
- [9] A. G. W. Leslie and H. R. Powell, *Processing diffraction data with mosflm*, 2007.
- [10] E. J. Dodson, M. Winn, and A. Ralph, *MACROMOLECULAR CRYSTALLOGRAPHY, PT B* **277**, 620 (1997).
- [11] W. KABSCH, *JOURNAL OF APPLIED CRYSTALLOGRAPHY* **26**, 795 (1993).
- [12] E. Potterton, P. Briggs, M. Turkenburg, and E. Dodson, *ACTA CRYSTALLOGRAPHICA SECTION D-BIOLOGICAL CRYSTALLOGRAPHY* **59**, 1131 (2003).
- [13] G. N. Murshudov, A. A. Vagin, and E. J. Dodson, *ACTA CRYSTALLOGRAPHICA SECTION D-BIOLOGICAL CRYSTALLOGRAPHY* **53**, 240 (1997).
- [14] P. Emsley and K. Cowtan, *ACTA CRYSTALLOGRAPHICA SECTION D-BIOLOGICAL CRYSTALLOGRAPHY* **60**, 2126 (2004).
- [15] P. Johannesson, *Moloc: using prolog for conceptual modelling*, 1991, <http://www.moloc.ch>.
- [16] W. L. DeLano and J. W. Lam, *Pymol: A communications tool for computational models*, 2005, <http://www.pymol.org>.
- [17] E. SUBRAMANIAN et al., *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA* **74**, 556 (1977).
- [18] O. A. Asojo et al., *Journal of Molecular Biology* **327**, 173 (2003).
- [19] E. Krissinel and K. Henrick, *JOURNAL OF MOLECULAR BIOLOGY* **372**, 774 (2007).
- [20] P. Y. Bruice, *Organic Chemistry*, Pearson Education International, 4th edition, 2004.
- [21] C. Binkert et al., *CHEMBIOCHEM* **7**, 181 (2006).
- [22] L. Prade et al., *JOURNAL OF BIOLOGICAL CHEMISTRY* **280**, 23837 (2005).
- [23] C. Boss et al., *CHEMMEDCHEM* **1**, 1341 (2006).
- [24] O. A. Asojo et al., *ACTA CRYSTALLOGRAPHICA SECTION D-BIOLOGICAL CRYSTALLOGRAPHY* **58**, 2001 (2002).
- [25] O. A. Asojo et al., *JOURNAL OF MOLECULAR BIOLOGY* **327**, 173 (2003).
- [26] F. Hof et al., *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION* **45**, 2138 (2006).