

**X-ray structure analysis of a pathogenic
bacterial protease from *Stenotrophomonas
maltophilia* towards drug discovery**

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Universität Hamburg

vorgelegt von

Amr Negm (M.Sc.)

aus Mansoura, Ägypten

Hamburg 2011

Die vorliegende Arbeit wurde im Zeitraum von April 2007 bis June 2011 in der Arbeitsgruppe von Prof. Ch. Betzel am Institut für Biochemie und Molekularbiologie am Department Chemie der Universität Hamburg und am Universitäts-Klinikum Hamburg Eppendorf (UKE), Institute für Biochemie und Molekularbiologie (IBM I), im Labor von PD. Dr. Wolfgang Weber durchgeführt.

Gutachter:

Herr Prof. Dr. Christian Betzel

Herr Prof. Dr. Reinhard Bredehorst

Tag der Disputation: 22.12.2011

Abstract

Most pathogenic bacteria are known to produce extracellular proteases that can attack and degrade host tissues and therefore are mainly responsible for the proceeding pathogenesis caused by the bacteria. *Stenotrophomonas maltophilia* is one of those bacteria, which cause pulmonary inflammation. 4.5% of nosocomial pneumonia in patients in intensive care units and 6% of ventilator pneumonias are linked these days to these bacteria. The pathogen is multi-drug resistant, thus, evading conventional antibiotic therapy. Particularly, in immune-suppressed patients this bacteria cause severe infections associated with tissue lesions such as pulmonary hemorrhage. These observations strongly suggest that bacterial proteases are damaging the infected tissue area. Indeed, it was shown before that *S. maltophilia* produces two extracellular proteases with broad specificity. The associated major protease gene, termed StmPr1, codes for a 63 kDa precursor, which is processed to a mature protein of 47 kDa. The enzyme is an alkaline serine protease, which, by sequence homology and enzymatic properties, can be classified as a new member of the subtilisin family (subtilisin-like protease). However, the molecular size is substantial larger compared to so far known subtilisins, and also the 3D-structure is suggesting a new fold within the family of subtilisin proteases. The high resolution X-ray structure of the *Stenotrophomonas maltophilia* protease StmPr1 was determined and refined to a final R-factor of 15.4% and R-free of 16.7 %. The protein was crystallized by the hanging drop method using 1.8 M ammonium sulphate as a precipitant. Crystals diffracted to 1.4 Å resolution applying synchrotron radiation with unit-cell parameters of $a = 60.17$, $b = 86.10$ and $c = 131.40$ Å corresponding to the orthorhombic space group $C222_1$ with one molecule in the asymmetric unit. The overall folding of the enzyme is quite similar to that observed for subtilisins. The protein is rich in acidic amino acids and contains four cysteine residues forming two intra-chain disulfide bridges. In terms of structure-based drug discovery investigations co-crystallization of StmPr1 with the peptide aldehyde inhibitors, chymostatin and leupeptin, was performed and the high resolution structures of the complexes were analyzed. The peptide aldehydes react with the active site residues of StmPr1 forming a complex with a hemiacetal conformation between the C-terminal L-phenylalanine residue of chymostatin or the L-arginine residue of leupeptin and the hydroxyl group of the catalytic Ser-289 of StmPr1. Further, high-throughput screening (HTS) was applied using the compounds of the ENZO and ChemBioNet

libraries to identify further inhibitors. For all investigation the protein was produced in mg quantities in *E. coli*, and an enzyme assay was established to check inhibitory effects suitable for the robot-based screening procedure applied. Several additional potential inhibitors could be identified with IC₅₀ values < 10 μM. One of them is bortezomib, applied already in anticancer therapy, inhibiting effectively the proteasome and showing an IC₅₀ value of 0.3 μM, towards the StmPr1 protease. Bortezomib was also co-crystallized and the structure was analyzed in the presented thesis. The structure of the StmPr1 protease reveals some differences in the architecture of the active site compared to the classic subtilisins and other serine proteases. In principle these differences can be utilized for the development of specific drugs. The screening experiments performed combined with the structures analyzed and the results obtained will support future drug discovery investigations. Preliminary cell culture experiments showed already that the *S. maltophilia* protease, which is able to destroy human lung cells can be inhibited in presence of bortezomib. Beside the summarized screening approaches, a peptide from the *Agkistrodon bilineatus* venom showing inhibitory activity towards StmPr1 was analyzed in complex with StmPr1. The peptide was provided in terms of an internal collaboration.

Future biological experiments using cell cultures and animal models will have to show whether the inhibitors identified so far may serve as lead compounds for drug discovery.

Zusammenfassung

Einige pathogene Bakterien produzieren extrazelluläre Proteasen, welche humanes Wirtsgewebe angreifen und abbauen können und damit auch lebensbedrohliche Effekte verursachen. Das gramnegative und multiresistente Bakterium *Stenotrophomonas maltophilia* ist eines dieser Bakterien. *S. maltophilia* verursacht insbesondere bei Patienten deren Immunsystem durch andere Krankheiten oder Infektionen geschwächt ist schwere Lungenentzündungen mit Gewebeschäden die auch zu Lungenblutungen führen. Diese Beobachtungen lassen auf bakterielle Proteasen schliessen, die für die nachhaltige Beschädigung des infizierten Gewebes verantwortlich sind. Im Rahmen vorhergehender Arbeiten konnte für *S. maltophilia* gezeigt werden, dass zwei extrazelluläre und weitgehend unspezifische Proteasen diese Aktivität vermitteln. Ein entsprechendes Protease-Gen wurde *StmPr1* genannt. Es kodiert ein 63 kDa grosses Vorläuferprotein, welches in einem Enzym mit 47 kDa Molekulargewicht resultiert. Es handelt sich hierbei um eine alkalische Serinprotease, die anhand von Sequenzhomologie und enzymatischer Aktivität der Familie der Subtilasen (Subtilisin-ähnliche Proteasen) zugeordnet wurde. Das Molekulargewicht ist im Vergleich zu bisher bekannten Subtilisinen um ca. 30% grösser. Das Protein wurde in für Röntgenstrukturanalysen erforderlichen Mengen exprimiert, aufgereinigt und folgend über Dampfdiffusion nach der Methode des hängenden Tropfen kristallisiert. Die Zellparameter ergaben sich zu: $a = 60,17$, $b = 86,10$ und $c = 131,40$ Å. Die Raumgruppe wurde zu $C222_1$ mit einem Molekül in der asymmetrischen Einheit bestimmt. Die Röntgenstruktur der *Stenotrophomonas maltophilia* Protease *StmPr1* wurde zu 1.4 Å Auflösung ermittelt und zu einem R-Faktor von 15,4% und R_{free} von 16,7% verfeinert. Die 3D Struktur zeigt die für Subtilisine bekannte Grundfaltung mit einer zusätzlichen Loopstruktur an der Oberfläche des Enzyms. Das Protein ist reich an sauren Aminosäuren und enthält vier Cys-Reste die zwei Disulfidbrücken ausbilden.

Um potentielle Inhibitoren gegen diese *StmPr1* Protease zu identifizieren und strukturbasiertes Wirkstoffdesign zu unterstützen, wurde *StmPr1* mit dem Peptid-Aldehyd-Inhibitoren Chymostatin und Leupeptin co-kristallisiert und Diffraktionsdaten unter Anwendung von Synchrotronstrahlung zu hoher Auflösung gesammelt. Die Komplexstrukturen zeigten, dass das C-terminale L-Phenylalanin des Chymostatin

Zusammenfassung

und der L-Arginin-Rest des Leupeptin die Hydroxyl-Gruppe des aktiven Ser289 koordinieren und das Enzym inhibieren.

Um weitere Inhibitoren zu identifizieren wurde ein Hochdurchsatz-Screening (HTS) mit der ENZO- als auch der ChemBioNet Compound Bibliothek durchgeführt. Zu diesem Zweck wurde das Protein in ausreichenden Mengen in *E. coli* produziert und ein entsprechender Enzym-Assay etabliert. Mehrere potenzielle Inhibitoren mit IC₅₀-Werten <20 µM wurden identifiziert. Einer dieser Inhibitoren ist Bortezomib mit einem IC₅₀ Wert von 0,4 µM, ein Wirkstoff der bereits in der Krebstherapie eingesetzt wird und dort das Proteasome effektiv inhibiert. Um die strukturellen Details der Inhibierung zu analysieren, wurde der StmPr1 Komplex mit Bortezomib kristallisiert und die Struktur analysiert. Die erhaltenen Strukturdaten können für die Entwicklung von spezifischen Hemmstoffen genutzt werden.

Im Rahmen einer Zusammenarbeit innerhalb der Arbeitsgruppe wurde auch ein Peptid aus dem Schlangengift der Schlange *Agkistrodon bilineatus* als potentieller StmPr1 Inhibitor identifiziert und eine Röntgenstrukturanalyse des Komplexes durchgeführt und damit ein weiterer Beitrag zum strukturbasierten Inhibitor design erarbeitet.

Abschliessende Zellkultur-Experimente bestätigten, dass die Sekrete von *S. maltophilia* in der Lage sind menschliche Lungenzellen zu zerstören. Und erste Experimente in Gegenwart von Bortezomib zeigten, dass der Abbau der Lungenzellen substantiell gehemmt wird. Zukünftig geplante biologische Experimente mit Zellkulturen und Tiermodellen werden zeigen, ob die bisher identifizierten Inhibitoren als Leitstrukturen für weitere Wirkstoffentwicklung genutzt werden können, da für zukünftige Medikamentenentwicklung ausschliesslich Inhibitoren mit hoher Spezifität gegen die StmPr1 Protease eingesetzt werden können.

Table of contents

Abstract.....	II
Zusammenfassung	IV
Table of Contents.....	VI
List of Figures.....	IX
List of Tables	13
List of Abbreviations	XIV
Physical Units	XVI
Symbols for Amino Acids	XVII
1. Introduction	1
1.1 <i>Stenotrophomonas maltophilia</i>	1
1.1.1 Associations of <i>Stenotrophomonas maltophilia</i> with plants	2
1.1.2 Opportunistic Pathogenicity of <i>Stenotrophomonas maltophilia</i>	3
1.2 Proteases	5
1.2.1 Serine Proteases.....	6
1.2.1.1 Mechanism of action	9
2. Materials and Methods	13
2.1 Materials	13
2.1.1. Laboratory Equipments	13
2.1.2. Chemicals, Reagents and Kits	13
2.1.3. Molecular weight Markers	15
2.1.4. Buffers and Solutions	15
2.1.5. Reagents and Media for Cell Cultures	17
2.1.6 Bacteria Culture Media	17
2.2 Methods.....	18
2.2.1 Expression of the <i>Stenotrophomonas maltophilia</i> Protease.....	18
2.2.2 Purification of the <i>Stenotrophomonas maltophilia</i> Protease.....	18
2.2.2.1 Ammonium sulphate Precipitation	18
2.2.2.2 Fast Protein Liquid Chromatography (FPLC)	18
2.2.3 Photometric Determination of the StmPr1 protease activity.....	19
2.2.4 Determination of the IC ₅₀ value for competitive inhibitors of StmPr1 protease	19

Table of contents

2.2.5 SDS-polyacrylamide gel electrophoresis	19
2.2.6 Crystallization Experiments	20
2.2.6.1 Crystallization of StmPr1	20
2.2.6.2 Dynamic Light Scattering (DLS)	20
2.2.6.3 Pre-Crystallization Test	21
2.2.6.4 Robotic Screening	22
2.2.6.5 Optimization of Crystals.....	22
2.2.6.6 Data Collection	22
2.2.6.7 Matthews Coefficient (V_M).....	22
2.2.6.8 Cryogenic Techniques	23
2.2.6.9 Macromolecular Crystallography Beamline X13.....	24
2.2.6.10 Model Building and Refinement	24
2.2.6.10.1 Coot: Crystallographic Object-Oriented Toolkit	24
2.2.6.10.2 Refmac5	24
2.2.6.11 Ligand Binding Experiments.....	24
2.2.7 Test of StmPr1 in Cell Cultures	25
3. Results and Discussion.....	26
3.1 Expression and Purification of StmPr1	26
3.2 Crystallization of StmPr1.....	28
3.2.1 Data Collection.....	30
3.2.2 Structure Solution.....	30
3.2.3 Model Building and Refinement	31
3.2.4 Description of the Molecule	33
3.2.5 Overall Structure Comparison with related Enzymes	36
3.2.6 Calcium-binding Site.....	41
3.3 Inhibition of StmPr1 by commercially available Compounds.....	43
3.4 Co-crystallization of StmPr1 with different Peptide aldehyde Inhibitors.....	46
3.4.1 Structure of the Leupeptin StmPr1 Complex	47
3.4.1.1 Structure Analysis	49
3.4.2 Structure of the Chymostatin StmPr1 Complex	52

Table of contents

3.4.2.1 Structure Analysis	54
3.5 A new Inhibitor identified in a Snake Venom by Crystallography	59
3.5.1 Structure Analysis	61
3.6 Identification of StmPr1 Inhibitors by High-Throughput Screening	64
3.6.1 Effect of StmPr1 and Bortezomib in Cell Culture	67
3.6.2 Crystal structure the Bortezomib StmPr1 Complex	69
3.6.2.1 Structure Analysis	71
4. Summary and Outlook	75
5. References	79
Risk and Safety Statements for the Compounds used in the Study	89
Acknowledgement	91
Curriculum Vitae	93

List of Figures

Figure 1: Biotechnological applications of <i>Stenotrophomonas</i> spp.	1
Figure 2: Phylogenetic analysis of the eight validly described <i>Stenotrophomonas</i> species and related taxa.	2
Figure 3: Cartoon-plot showing the secondary and tertiary structure of subtilisin Carlsberg (PDB ID: 2SNI).....	9
Figure 4: Cartoon representation of the action mechanism showing substrate interactions.....	9
Figure 5: The nucleophilic attack of Ser OH on the peptide carbonyl.	10
Figure 6: Oxyanion hole formation.	10
Figure 7: The importance of the water molecule during the action.....	11
Figure 8: The liberation of the product.....	11
Figure 9: The path from genome to protein structure (blue arrows) and cycle of structure-based drug design.	12
Figure 10: Zinsser Pipetting Robot (Digilab Genomic Solution, Germany).	22
Figure 11: Cryogenic technique accessories.	23
Figure 12: SDS-polyacrylamide gel electrophoresis of purified StmPr1.	27
Figure 13: StmPr1 activity at different substrate concentrations.	28
Figure 14: DLS measurement showing a monodispersive protein solution.	29
Figure 15: StmPr1 Crystals.....	30
Figure 16: Sequence alignment between StmPr1 and <i>Dichelobacter nodosus</i> Subtilisin (PDB ID: 3IPA)..	31
Figure 17: Model of the StmPr1 crystal structure.	34
Figure 18: A topology diagram of the StmPr1 structure.	34
Figure 19: Electrostatic potential surface of StmPr1.....	35

List of Figures

Figure 20: Cartoon plot of the superposition of StmPr1 (blue) with the cold-adapted Vibrio proteinase (1SH7, green).....	36
Figure 21: Cartoon plot of the superposition of StmPr1 (blue) with proteinase K (red).	37
Figure 22: Cartoon plot of the superposition of StmPr1 (blue) with thermitase (yellow)	37
Figure 23: Cartoon plot of the superposition of StmPr1 (blue) with the subtilisin-like protease (3LPA, orange).....	38
Figure 24: C α -backbone plot of StmPr1 showing the loop insertions around the active site and the disulfide bridges providing the loop stability.	39
Figure 25: Fo-Fc electron density map depicting the disulfide bridge, the L2 loop of Stmpr1.....	40
Figure 26: StmPr1 structure colored according to B-factor value for the side chains.	40
Figure 27: StmPr1 ligplot showing the calcium-binding site.	41
Figure 28: Scheme representation of the commonly used nomenclature for the enzyme subsites and the corresponding ligand sites.	44
Figure 29: Inhibitory effects of the selected commercially available inhibitors.	45
Figure 30: Chemical structure of leupeptin.	47
Figure 31: Ligplot showing leupeptin in the binding site of StmPr1.	49
Figure 32: Active site of StmPr1. Catalytic triad residues (Asp-42, His-105, and Ser-289) are colored and labeled in purple. Main chains and side chains of StmPr1 residues forming polar contacts with Leupeptin (red dashed lines) are shown. Inhibitor is shown in gray.....	50
Figure 33: Surface view of the StmPr1-leupeptin complex, colored according to electrostatic potential (blue, positive; red, negative; white, neutral); leupeptin is shown in yellow.	51

List of Figures

Figure 34: Fo-Fc map contoured at 1.0 σ showing the electron density for the catalytic Ser-289 covalently bound to leupeptin through a hemiacetal bond.....	51
Figure 35: The chemical structure of chymostatin.	52
Figure 36: Ligplot of the chymostatin StmPr1 complex showing the hydrogen bond interactions.....	54
Figure 37: The chymostatin-StmPr1 complex structure.....	55
Figure 38: The interactions of the water molecules in the StmPr1 active site.	56
Figure 39: Surface view of StmPr1 showing chymostatin (in sticks) in the binding site of StmPr1.	56
Figure 40: Superposition of the C α -atoms of native StmPr1 structure (in cyan) and StmPr1-chymostatin Complex structure (in gray).	57
Figure 41: Fo-Fc map contoured at 1.0 σ , showing the electron density for the catalytic Ser-289 covalently bound to the chymostatin molecule through a hemiacetal bond.	58
Figure 42: Ligplot of the StmPr1- <i>Agkistrodon bilineatus</i> venom peptide inhibitor complex.....	61
Figure 43: Surface view of StmPr1 showing the peptide inhibitor (in sticks) in the binding site of StmPr1.	63
Figure 44: The Inhibitory effect of bortezomib on StmPr1.....	65
Figure 45: The chemical structure of bortezomib.	65
Figure 46: Inhibition of StmPr1 with bortezomib and chymostatin.....	66
Figure 47: Effect of bortezomib on cultures of human lung cells.....	68
Figure 48: Effect of bortezomib on cultures of human lung cells pre-treated with the <i>Stenotrophomonas maltophilia</i> culture medium.....	68
Figure 49: Ligplot of StmPr1 showing bortezomib in the binding site.	71

List of Figures

Figure 50: Active site of StmPr1 (top view). Catalytic triad residues (Asp-42, His-105, and Ser-289) are colored and labeled in purple. Main chains and side chains of StmPr1 residues forming polar contacts with Bortezomib (red dashed lines) are shown. Inhibitor is in gray.72

Figure 51: Covalent bond between Ser-289 and the boronic acid moiety of bortezomib.73

Figure 52: Surface representation of StmPr1 showing bortezomib (in yellow sticks) in the binding site of StmPr1.73

Figure 53: Five positive HTS hits represent non-peptidic compounds.74

List of Tables

Table 1: Purification Steps of StmPr1	26
Table 2: Data collection parameters and refinement statistics of the native StmPr1 ...	32
Table 3: Statistical values of the primary amino acid sequence	33
Table 4: Pairwise superposition of C α -atoms in StmPr1, the cold-adapted <i>Vibrio subtilisin</i> -like protease (PDB ID: 1SH7), proteinase K (PDB ID: 1IC6) and thermitase (PDB ID: 1THM).	38
Table 5: Pairwise superposition of C α -atoms in StmPr1, with <i>Dichelobacter nodosus subtilisin</i> 3LPA.....	39
Table 6: The ionic interactions in calcium binding site.....	42
Table 7: Sequences of the selected inhibitors.....	44
Table 8: Data collection parameters and refinement statistics of the leupeptin StmPr1 complex.....	48
Table 9: Interactions of the leupeptin StmPr1 complex	50
Table 10: Data collection parameters and refinement statistics of the chymostatin StmPr1 complex.....	53
Table 11: The Chymostatin StmPr1 complex interactions	55
Table 12: Statistics for <i>Agkistrodon bilineatus</i> snake venom peptide StmPr1 complex	60
Table 13: The <i>Agkistrodon bilineatus</i> venom peptide inhibitor-StmPr1 complex interactions.....	62
Table 14: Data collection parameters and refinement statistics of the bortezomib StmPr1 complex.....	70
Table 15: The interactions of StmPr1-bortezomib complex	71

List of Abbreviations

C222 ₁	C Centered Orthorhombic Space group
CCD	Charge-coupled Device
CCP4i	Collaborative Computational Project Number 4
DESY	Deutsches Elektronen-Synchrotron
DLS	Dynamic Light Scattering
EMBL	European Molecular Biology Laboratory
FDA	Food and Drug Administration
F_c	Calculated Structure-factor
F_o	Measured Structure-factor
FPLC	Fast Protein Liquid Chromatography
HCl	Hydrochloric acid
HTS	High-Throughput Screening
MR	Molecular Replacement
MWCO	Molecular Weight Cut Off
NCBI	National Center for Biotechnology Information
PAGE	Polyacrylamide Gel Electrophoresis
PCS	Photon Correlation Spectroscopy
PCT	Pre-Crystallization Test
PDB	Protein Data Bank
PDB ID	Identification Code for Protein Data Bank
PEG	Polyethylene Glycol
R_{free}	R -factor, based on selection of reflections not considered for structure solution
R_{merge}	Reliability factor of all symmetry-equivalent reflexes
SDS	Sodium Dodecyl Sulphate

List of Abbreviations

StmPr1	<i>Stenotrophomonas maltophilia</i> Protease 1
T	Temperature
Tris	tris (hydroxymethyl) aminomethane
UK	United Kingdom
USA	United States of America
UV	Ultraviolet

Physical Units

°	Degree
°C	Degree centigrade
Å	Angstrom
eV	Electron volt
g	Gram
k	Boltzmann's Constant
K	Kelvin
kDa	Kilo Dalton
M	Molar
mg	Milligram
min	Minute
ml	Milliliter
mm	Millimeter
mM	Milli-molar
nl	Nano-liter
nm	Nanometer
R_H	Hydrodynamic Radius
rpm	Revolutions per minute
V_M	Matthews Coefficient
η	Viscosity
λ	Wavelength
μl	Micro-liter
μs	Micro-second
μM	Micro-molar

Symbols for Amino Acids

A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartate
C	Cys	Cysteine
E	Glu	Glutamate
Q	Gln	Glutamine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine

1. Introduction

1.1 *Stenotrophomonas maltophilia*

The species *Stenotrophomonas maltophilia* was originally named *Pseudomonas maltophilia*, but was later transferred to the genus *Xanthomonas* [1] before it obtained its own genus. The genus name (from the Greek 'stenos', meaning narrow, 'trophus', meaning one who feeds and 'monas', meaning unit) was intended to highlight the limited nutritional range of the bacterium [2]. However, several studies subsequently demonstrated that the genus is capable of great metabolic versatility and intraspecific heterogeneity [3], as shown in (Figure 1). The genus currently comprises eight species including *Stenotrophomonas maltophilia* [4]. Phenotypic and genotypic studies as well as analysis of the ecological and metabolic diversity of these bacteria have revealed further differentiation at the species level.

Stenotrophomonas species have many qualities, which could be used in different biotechnological processes. Some *Stenotrophomonas* spp. can produce antimicrobial compounds that protect plants, as well as generate factors that can promote plant growth. Also, many *Stenotrophomonas* spp. have a high level of intrinsic resistance to heavy metals and antibiotics and have been shown to degrade a wide range of compounds, including pollutants, and could potentially be used in bioremediation and phytoremediation.

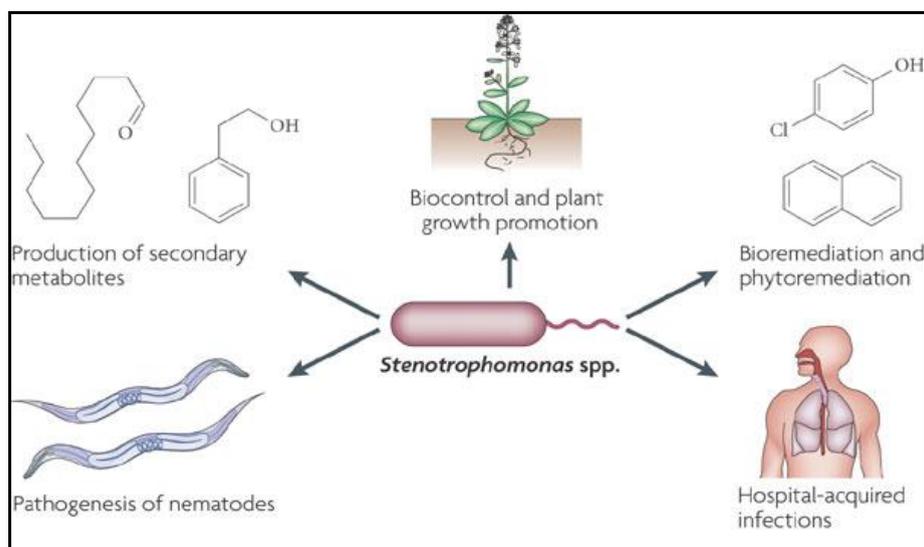


Figure 1: Biotechnological applications of *Stenotrophomonas* spp.

Stenotrophomonas maltophilia is also known to cause human diseases as a result of its ability to colonize immune-compromised patients, and has been shown to be virulent in a nematode model [5]. The following tree illustrates the maximum likelihood of *Stenotrophomonas* species and related taxa [based on 16S rDNA sequences].

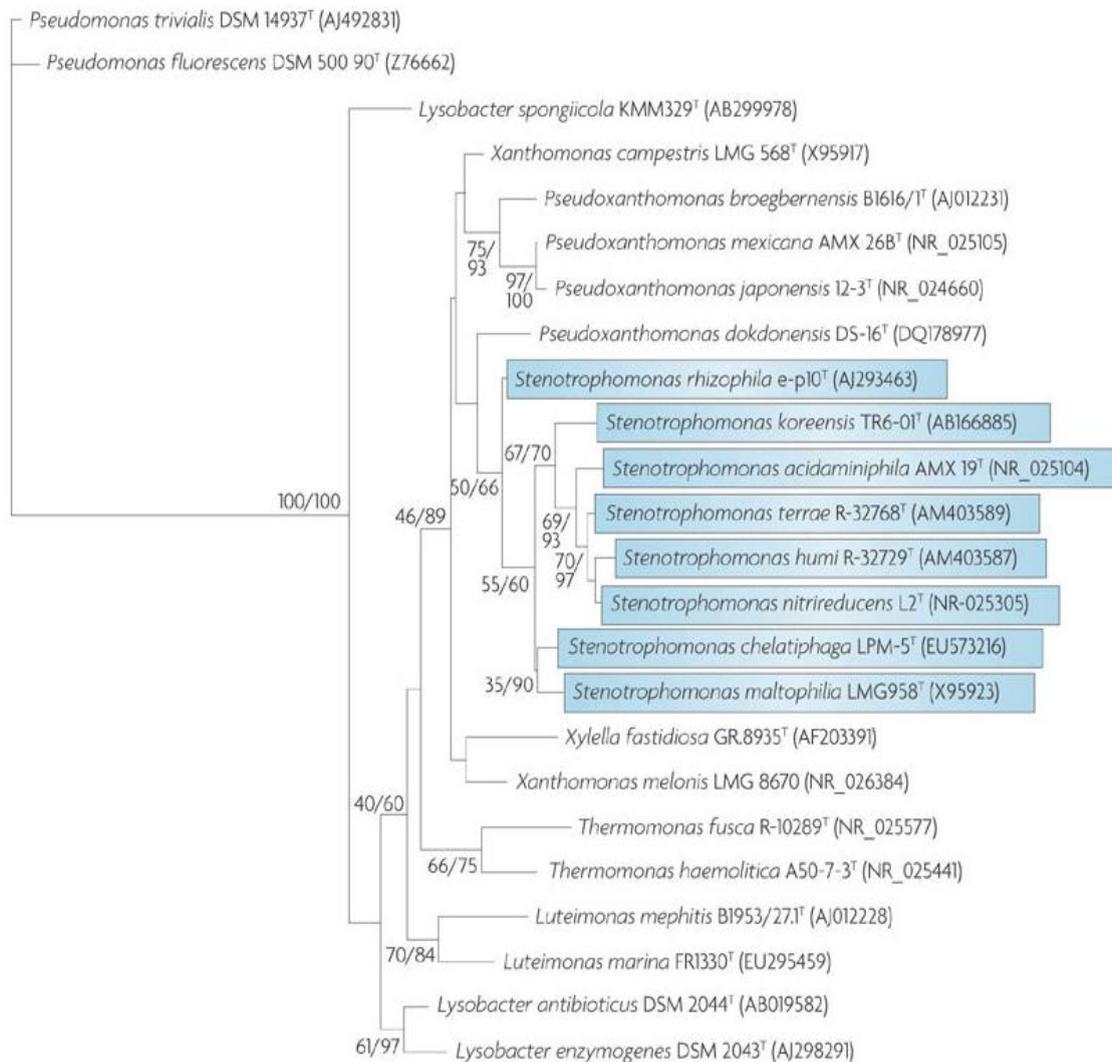


Figure 2: Phylogenetic analysis of the eight validly described *Stenotrophomonas* species and related taxa.

1.1.1 Associations of *Stenotrophomonas maltophilia* with plants

Stenotrophomonas species, especially *S. maltophilia* are often found in association with plants. These bacteria can be isolated from the rhizosphere [6] or from internal plant tissues, particularly from the vascular tissues of the root and stem. Endophytic strains of *S. maltophilia* have been isolated from the roots of many plant

species, including cucumber (*Cucumis sativus*) [7], potato (*Solanum tuberosum*), strawberry (*Fragaria x ananassa*) [8], sunflower (*Helianthus annuus*) [9], maize (*Zea mays*) [10], rice (*Oryza sativa*) [11], wheat (*Triticum aestivum*) [12].

S. maltophilia strains have an extraordinarily high hydrolytic potential; they produce diverse proteases, chitinases, glucanases, DNases, RNases, lipases and laccases [13]. Both chitinolytic and proteolytic activities contribute to the biocontrol activity of *S. maltophilia* [14]. Chitinases might protect plants against fungal pathogens through fungal cell wall lysis but might also have a role in triggering plant defence mechanisms [15].

1.1.2 Opportunistic pathogenicity of *Stenotrophomonas maltophilia*

Stenotrophomonas species colonizes a varied biodiversity efficiently as plants, humans and marine environments. *S. maltophilia* is the only species of *Stenotrophomonas* that is known to cause human disease [3]. Most infections reflect contact with separate environmental sources. Indeed, there are few instances of outbreaks of *S. maltophilia*, and those that occur are caused by a single contaminated source, such as a water source [16]. It is a hospital-acquired pathogen and has been associated with bacteraemic infections and pneumonia, both with a high rate of mortality in immune-compromised patients [17]. This reflects a requirement for three major risk factors for infection: severe debilitation and/or neutropenia; the presence of indwelling devices such as ventilator tubes and/or intravenous catheters for prolonged periods; and multiple and/or prolonged courses of broad-spectrum antimicrobial drug treatment [18].

The bacteraemia isolate *S. maltophilia* carries several genes that encode factors that could allow this strain to adhere to surfaces and to form biofilms, which are both key factors in the colonization of indwelling devices. Studies have shown that *S. maltophilia* isolates have cytotoxic effects in vitro against some cell lines after 24 hours [19], and it can kill almost all of the N2 *Caenorhabditis elegans* in the assay within 24 hours [20].

A recent study tried to address the lack of documentation of the potential of *S. maltophilia* for virulence by investigating the immune-stimulatory properties of 24 *S. maltophilia* clinical respiratory and non-respiratory isolates (from blood, skin and soft

tissue). In this study, which involved a neonatal mouse model of pneumonia and macrophage cell lines, they determined the rates of pneumonia, bacteraemia and mortality, as well as the inflammatory response that is elicited by *S. maltophilia* infection. They demonstrated that the respiratory and non-respiratory *S. maltophilia* isolates were highly immune-stimulatory but weakly invasive, indicating that these bacteria have contributions in airway inflammation [21]. Another study showed that 4.5% of nosocomial pneumonia in patients in intensive care units and 6% of ventilator-associated pneumonias are caused by *S. maltophilia* [22].

The primary reason for the increase in *S. maltophilia* infections is the intrinsic resistance of this species to many front-line antimicrobials, such as β -lactams, including carbapenems [23], aminoglycosides (except gentamicin) [24], macrolides, tetracycline, chloramphenicol and older quinolones [25]. Furthermore, *S. maltophilia* isolates can rapidly develop resistance to newer fluoroquinolones, gentamicin and minocycline through mutation; the underlying mechanisms are not certain, but are likely to be the result of the overproduction of intrinsic efflux pumps [26]. Typically, empiric therapy for *S. maltophilia* is trimethoprim–sulphamethoxazole (TMP–SMX), to which >95% of isolates are sensitive [27]. However, resistance is increasing as a result of the spread of acquired mobile resistance determinants [28], and, in many patients, TMP–SMX therapy is contra-indicated [29].

In immune-competent individuals, *S. maltophilia* can cause bacteremia [30–32], endocarditis [33], pneumonia [34], mastoiditis [35], peritonitis and exit sign infections in patients undergoing peritoneal dialysis, meningitis [36], or infections of the eye [37], bones and joints [38], urinary tract [38], soft tissue, and wounds [32]. *S. maltophilia* infections are of increasing importance, especially in intensive care units [39], and are a growing source of pulmonary infections and bacteremia [40]. Especially susceptible to *S. maltophilia* are immune-compromised patients, for example, patients with cancer, as well as those receiving mechanical ventilation, broad spectrum antimicrobial therapy, and those subjected to invasive procedures [41]. In addition, *S. maltophilia* is considered as an important emerging pathogen in patients with cystic fibrosis [42]. This pathogen, as many other non-fermenting gram-negative bacteria, such as *Pseudomonas*, *Burkholderia*, or *Acinetobacter*, is frequently multidrug-resistant, thus, causing difficulties in its treatment [43].

Infection due to multidrug-resistant *S. maltophilia* has emerged as an important nosocomial infection in many hospitals [44-46]. Risk factors, such as old age, poor physiological score, and bacteraemia, have been found to increase mortality among patients with *S. maltophilia* infection [47-49]. There are further reports demonstrating involvement of this bacterium in massive haemorrhagic processes of the small intestine and of the subclavian artery accompanied by severe lesions of the tissue [50, 51]. These observations strongly suggest participation of proteolytic activity, produced by the bacteria, which may damage the infected tissue. Indeed, it is known that members of the Pseudomonaceae express and secrete a variety of proteases [52]. Those extracellular proteases have been considered for many years to be potential virulence factors [53].

Stenotrophomonas maltophilia had been shown to produce two extracellular proteases. The genes have been cloned, termed StmPr1 and StmPr2, and the gene products have been characterized [54]. Both proteases are highly homologous and do not differ significantly in enzymatic properties. Since StmPr2 seemed to be expressed – if at all – only in trace amounts, this work is focused on the StmPr1 protease as the major secretory protease of *S. maltophilia*. It is an alkaline serine protease, and according to sequence homology can be assigned to the superfamily of subtilases (subtilisin-like proteases). However, StmPr1 differs from the classic subtilisins in molecular size and substrate specificity. Whereas known bacterial subtilisins, like proteinase K, are enzymes with a molecular weight of approximately 27 kDa. The mature *Stenotrophomonas* enzyme, because of its C-terminal extension, is significantly larger: The gene codes for a 63-kDa precursor that is processed to the mature protein of 47 kDa. Also, the amino acid sequence predicted differences in the architecture of the active site, as indicated by inserts adjacent to the catalytic His and Ser residues. The StmPr1 protease is able to degrade several human proteins from serum and connective tissues [55].

1.2 Proteases

Proteases occur naturally in all organisms. These enzymes are involved in the physiological reactions from simple digestion of food proteins to highly-regulated cascades (e.g., the blood-clotting cascade, the complement system, apoptosis pathways, and the invertebrate prophenol oxidase-activating cascade). Bacteria also

secrete proteases to hydrolyse the peptide bonds in proteins and therefore break the proteins down into their constituent monomers [56]. These proteases act as an exotoxin, and are an example of virulence factor in bacterial pathogenesis. Those bacterial exotoxic proteases destroy extracellular structures.

Proteases belong to the class of enzymes known as hydrolases, which catalyse the reaction of hydrolysis of various bonds with the participation of a water molecule. They are currently classified into six broad groups: Serine proteases, Threonine proteases, Cysteine proteases, Aspartate proteases, Metalloproteases and Glutamic acid protease. StmPr1 Proteinase belongs to the subtilase family which basically belongs to serine proteases group [57].

1.2.1 Serine proteases

Serine endo- and exo-peptidases are of extremely widespread occurrence and diverse function. They have been grouped into six clans, of which the two largest are the chymotrypsin-like and subtilisin-like clans. These two clans exploit the same catalytic triad, although the residues occur in a different order (His-Asp-Ser) in case of chymotrypsin and (Asp-His-Ser) in case of subtilisin, but the structures show no other similarity [58]. Several members contain C-terminal extensions, relative to the subtilisins, which display additional properties such as sequence repeats, Cysteine-rich domains, or trans-membrane segments. Subtilases are members of the clan of subtilisin-like serine proteases. Proteins of the subtilase superfamily have been identified in a diverse range of organisms including achaea, bacteria, fungi, yeasts, higher eukaryotes, and even viruses [52, 59]. Structures have been determined for several members of the subtilase superfamily. Subtilases were further subdivided into six subgroups (families):

Subtilisin family

Subtilisins, one clan of the subtilase superfamily, play an important role in the environmental and agricultural fields, and have received extensive interest for their proteolytic properties, which make them useful in various industrial applications. The subtilisin family is considered to be the second largest characterised serine protease family. Over 200 subtilises are presently known, more than 170 of which with their

complete amino acid sequence. They are widespread, being found in eubacteria, archaeobacteria, eukaryotes and viruses [58]. The vast majority of the family are endopeptidases, although there is an exopeptidase, tri-peptidyl peptidase [60].

It includes mainly enzymes from *Bacillus*. Currently, based on the amino acid sequences, subtilisin family has been divided into six clans, i.e., true-subtilisins, high-alkaline proteases, intracellular proteases, oxidatively stable proteases, high-molecular-mass subtilisins and phylogenetically intermediate subtilisins [61]. Several 3D structures of these group members are known.

Thermitase family

It is a group of enzymes found only in micro-organisms, including some thermophiles (>55% identity) and halophiles. The characteristic N-terminal sequence was also found in several other *Bacillus* proteases [52]. Only one 3D structure is known which is called thermitase (PDB ID: 1THM).

Proteinase K family

It is a large family of secreted endo-peptidases found only in fungi, yeasts, and gram-negative bacteria as yet; the bacterial subgroup has >55% sequence identity. This family is characterized by a high degree of sequence similarity (>37% identity), only minor insertions and deletions and the absence of the Ca²⁺-binding loop residues 76-81. Only a few of these enzymes have a significant C-terminal extension beyond the catalytic domain. One 3D structure is known which is called proteinase K (PDB ID: 1IC6). In the proteinase K structure an unusual, short hydrogen bond between aspartic acid and histidine in the catalytic triad was reported which occurred as a part of an elaborate hydrogen bond network, involving Asp of the catalytic triad [62].

Lantibiotic peptidase family

It is a small number of highly specialized enzymes for cleavage of leader peptides from precursors of lantibiotics, a unique group of post-transnationally modified, antimicrobial peptides [63]. These endopeptidases have only been found in gram positive bacteria, and several are intracellular. They are characterized by low sequence similarity with each other and other subtilases, and by numerous insertions

and deletions. The most recently reported protein bspara from *Bacillus subtilis* is described as a putative protease required for plasmid stability. A few 3D structures have been predicted by homology modeling [64, 65]

Kexin family

It is a large group of proprotein convertases (PCs). Nearly all of them are involved in activation of peptide hormones, growth factors, viral proteins, etc. [66, 67]. High specificity is reported for cleavage after dibasic (Lys-Arg or Arg-Arg) or multiple basic residues. Nearly all are eukaryotic and have high sequence homology (>40% identity), while two more distant members from *Aeromonas* and *Anabaena* provide links to other subtilase families. A subgroup of yeast enzymes is evident, as are subgroups of PC1 (55 % identity), PC2 (>73% identity), and furin (>55% identity) [58]. Several 3D structures have been predicted by modeling.

Pyrolysin family

It is considered to be a heterogeneous group of enzymes of varied origin and low sequence conservation (most <37% identity). It is characterized by large insertions and/or long C-terminal extensions, many with sequence homology suggesting common ancestors. The most extreme example is llspO9 from the plant *Lilium* with insertions totaling more than 260 residues compared to subtilisin, almost doubling the size of the catalytic domain. Subgroups of tripeptidyl peptidases and plant subtilases (>37% identity) are distinguished; the former are of higher eukaryotic origin, but only the human and mouse enzymes have actually been identified biochemically as tripeptidyl peptidases [52]. Several 3D structures have been predicted by modeling.

More than 200 proteases have been assigned to the superfamily of subtilases (subtilisin-like serine proteases), with representatives in both of microorganisms (archaea, bacteria, fungi and yeast) and in higher eukaryotes [52]. All the enzymes belonging to this superfamily have in common a core structure, the catalytic domain, characterized by the presence of structurally conserved regions, which correspond to common secondary structure elements. The conserved catalytic residues Asp-32, His-64, and Ser-221 are highlighted in (Figure 3). The coordinates of subtilisin BPN', Subtilisin Carlsberg, Thermitase, Savinase, Esperase and Proteinase K were used previously by Siezen 1991 [68] to determine the core of "structurally conserved

regions” and the common secondary structure elements. This core of about 190 residues contains virtually all of the common α -helix and β -strand elements, including the active site residues D-32, H-64 and S-221 [68].

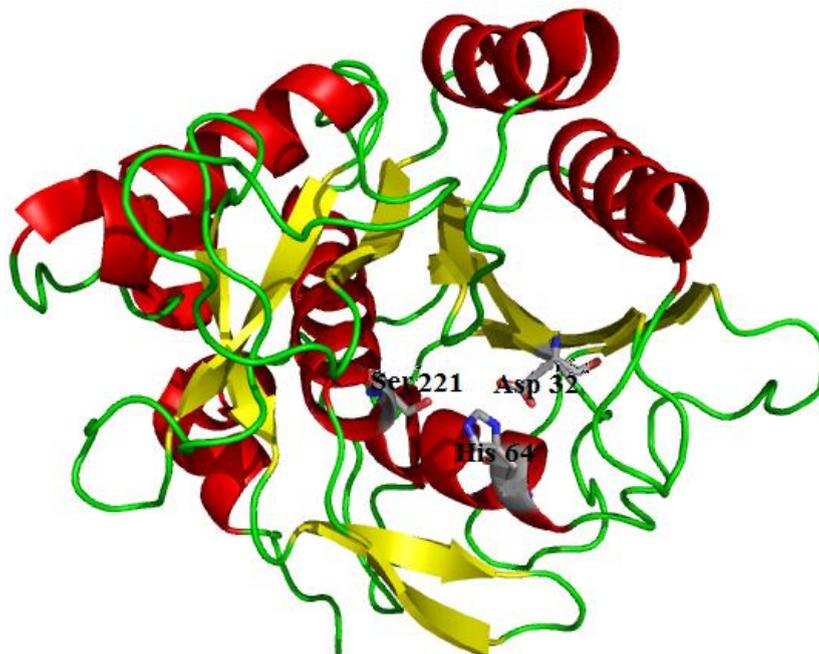


Figure 3: Cartoon-plot showing the secondary and tertiary structure of subtilisin Carlsberg (PDB ID: 2SNI).

This cartoon shows the typical subtilisin folding; α -helices are shown in red, β -sheets in yellow, loops in green. The side chains of the catalytic residues are shown in stick representation.

1.2.1.1 Mechanism of action

The mechanism of action of chymotrypsin, outlined below, applies to all of the serine proteases, with small variations. The hydrolysis of a peptide substrate occurs in several steps [69]:

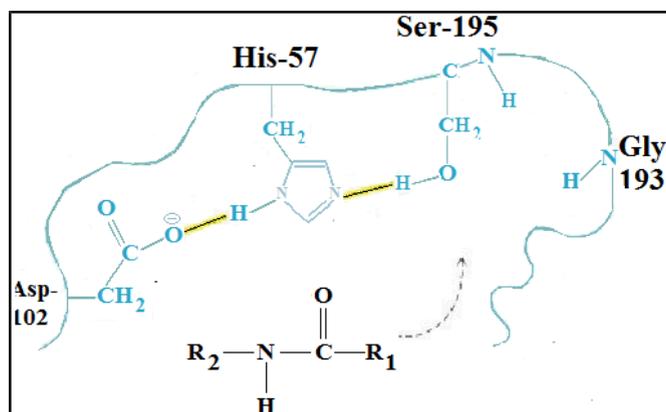


Figure 4: Cartoon representation of the action mechanism showing substrate interactions.

Firstly, the nucleophilic attack of Ser-195 on the carbonyl group of the substrate. At the same time serine protonates the imidazole ring of the neighboring His-57. Asp-102 plays an important role in stabilizing the positive charge of histidine.

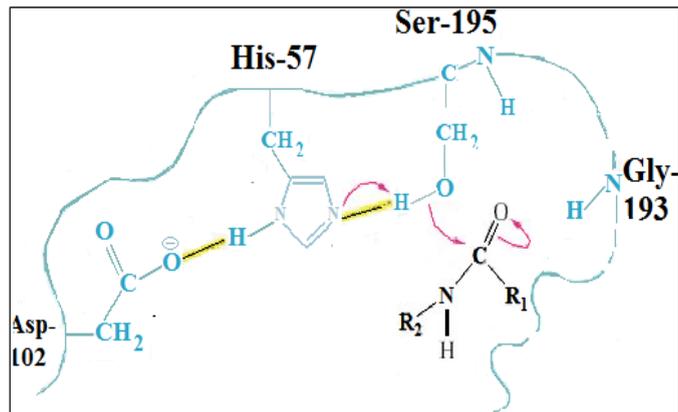


Figure 5: The nucleophilic attack of Ser OH on the peptide carbonyl.

The shift of the negative charge onto the carbonyl oxygen of the peptide is facilitated by hydrogen bonding of the oxygen to the backbone NH groups of Ser-195 and Gly-193. These form "the oxyanion hole" (Figure 6).

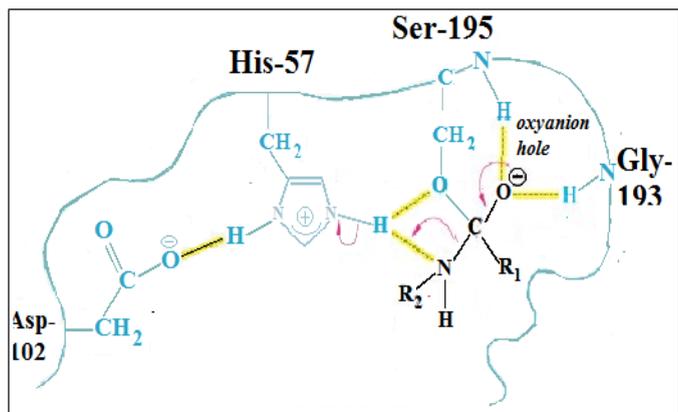


Figure 6: Oxyanion hole formation.

The next step is the reconstruction of the carbonyl double bond; with expulsion of the leaving group - in this case- the rest of the protein (Figure 6). This is the stage, at which the protein chain actually is cleaved, and it produces an "acyl enzyme", this needs a molecule of water (Figure 7).

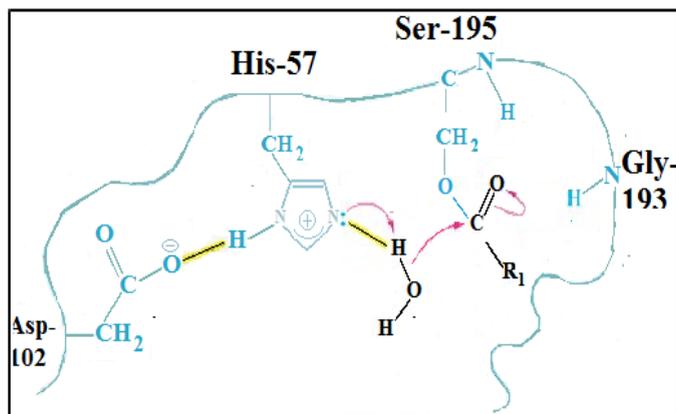


Figure 7: The importance of the water molecule during the action.

Enzymes are catalysts, and are not permanently altered in the reaction, so the next step will be to restore the carbonyl double bond. This releases the other end of the original protein, and restores the catalytic triad to its beginning state. The dissociation of the second protein fragment leaves the enzyme ready to attack another substrate molecule (Figure 8).

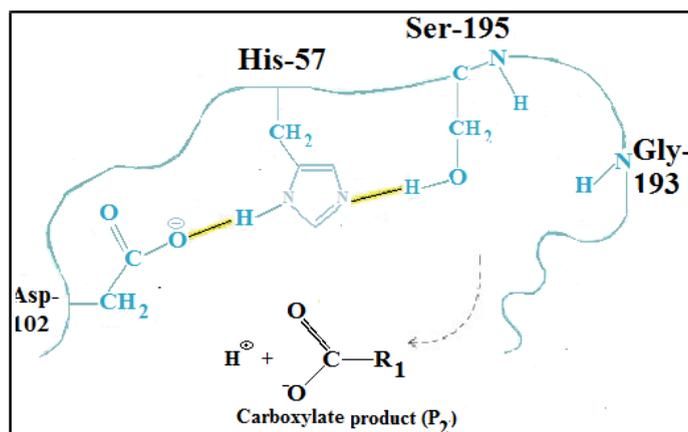


Figure 8: The liberation of the product.

In fact there are many pathogenic bacteria which produce extracellular proteases that can degrade the extracellular matrix of the host, therefore they are involved in disease pathogenesis and considered as main virulence factors, such as *Dichelobacter nodosus*, which is a causative agent of ovine foot rot, and which is also considered to be a highly contagious disease that is characterized by the separation of the hoof from the underlying tissue. It secretes three subtilisin-like proteases, which have been postulated to play a role in virulence [70]. As proteases are involved in many pathological processes, so it is meaningful to inhibit these enzymes. Selective inhibitors have the potential to support the discovery of new drugs.

Today several commercially serine protease inhibitors are available, however, they are unstable chemically, degrading in solution, difficult to synthesize, and they are not specific. Therefore, in order to aid in the process of developing better inhibitors with more specific activity, it is important to determine the structure of the *Stenotrophomonas maltophilia* Protease StmPr1. The determination of the 3-dimensional structures of proteases and the knowledge of the atomic structure of the proteins can lead to a better understanding of the chemical reactions which take place and can help to provide a basis for drug design. The determination of a protein structure by X-ray crystallographic analysis involves several steps. Of course the first step is the protein purification. For that purpose StmPr1 was over expressed in *E.coli* to produce enough amounts, because it is not possible to produce sufficient amounts of the protein directly through *Stenotrophomonas maltophilia* [54]. The path from genome to 3D protein structure is outlined in (Figure 9).

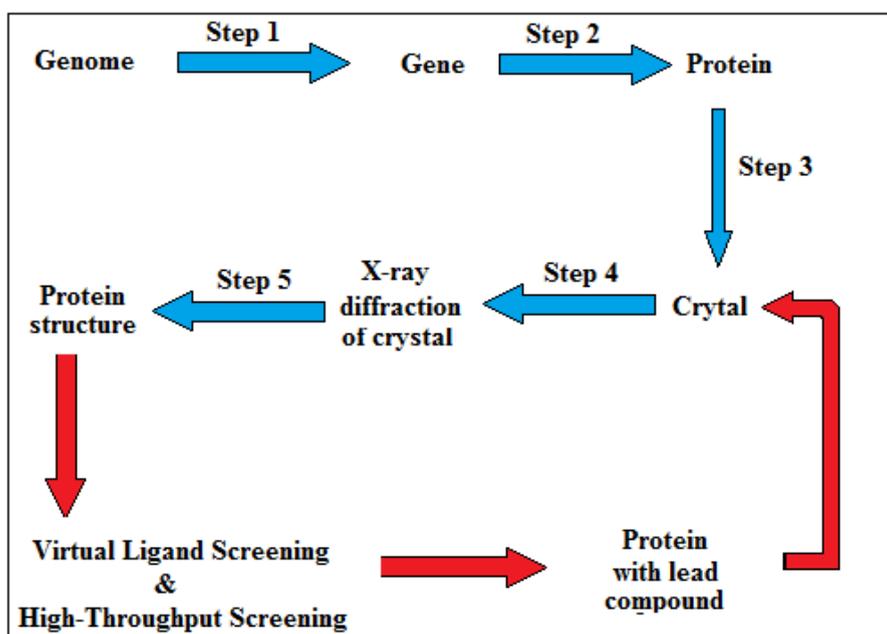


Figure 9: The path from genome to protein structure (blue arrows) and cycle of structure-based drug design.

The steps of screening and incubation of protein with lead compounds are shown to illustrate the use of crystallography for lead optimization (red arrows).

2. Materials and Methods

2.1 Materials

2.1.1. Laboratory equipments

Provider	Apparatus
Mettler-Toledo GmbH, Giessen, Germany	Analytical balance P1200
Johanna Otto GmbH, Hechingen, Germany	Analytical balance AE160
	Edmund Bühler shaker model KM-2
	Edmund Bühler shaker model SM-30
Bacterial incubator	Heraeus
Electrophoresis chambers (horizontal)	MWG-Biotech
Electrophoresis chambers (vertically)	Protean [®] IIXiCell, Bio-Rad
Lyophilisator: SpeedVac Concentrator	Savant
pH meter CG 820	Schott-Geräte GmbH
stir plate	Heidolph
Shaker	HAT Infor AG
Spectrophotometer DU-8	Beckman
PCR device	MWG Biotech
Thermo mixer 5436	Eppendorf
Vacuum oven	Heraeus
Water bath1086	GFL
Centrifuges:	
Centrifuge 5415C	Eppendorf
Centrifuge 5403	Eppendorf
Sorvall-centrifuge RC-5B	DuPont Instruments

2.1.2. Chemicals, reagents and kits

Provider	Substance
Sigma-Aldrich Chemie GmbH,	β -Mercaptoethanol
Deisenhofen, Germany	2-propanol
	Boric acid
	Bovine serum albumin (BSA)
	Bromphenol blue
	Di Methyl Sulfoxide (DMSO)
	Ethanol absolute

Ethylene diamine tetracetic acid (EDTA)

Agarose

Glucose

Glycine

Magnesium chloride (MgCl₂)

Potassium chloride (KCl)

Di-sodium hydrogen phosphate (Na₂HPO₄)

Sodium dihydrogen phosphate (NaH₂PO₄)

Sodium chloride (NaCl)

Potassium dihydrogen phosphate (KH₂PO₄)

Sodium Dodecyl Sulfate (SDS)

Sucrose

Coomassie brilliant blue R250

Coomassie brilliant blue G250

Bromophenol Blue

Dithiothreitol (DTT)

Qiagen GmbH, Hilden, Germany

Qiagen plasmid Mini kit

Qiagen plasmid Maxi kit

Merck Biosciences GmbH, Bad,
Soden, Germany

Glycerol

2.1.3. Molecular weight markers

- **Protein markers**
 - Bench Mark prestained protein ladder (Invitrogen, Karlsruhe, Germany)
 - Magic Mark Western standard (Invitrogen, Karlsruhe, Germany)
- **DNA marker**
 - 1kb DNA ladder, Invitrogen, Karlsruhe, Germany

2.1.4. Buffers and solutions

For all buffer preparations double distilled water was used.

- **Phosphate buffer:** 0.5 M K_2HPO_4 with 0.5 M KH_2PO_4 and pH was adjusted to 7.
- **PBS:** 20 mM phosphate buffer pH 7.2 with 120 mM NaCl
- **TBS:** 20 mM Tris / HCl pH 7.5, 120 mM NaCl
- **Coomassie-blue staining solution**

2 mM	Coomassie brilliant blue R250
0.6 mM	Coomassie brilliant blue G250
42.5%	Ethanol
10%	Acetic acid
- **Destaining solution**

13%	Methanol
10%	Acetic acid
- **DNA loading buffer**

30%	Glycerol
0.25%	Bromophenol Blue
0.25%	Xylene Cyanol

Materials and Methods

- **10x Tris-glycine buffer (TG-buffer)**

1.92 M glycine

0.25 M Trizma base

- **Electrophoresis buffer (1x)**

100 ml/l 10x TG-buffer

10 ml/l 10% SDS

- **PBS (phosphate buffered saline)**

140 mM NaCl

3 mM KCl

8 mM Na₂HPO₄

1.5 mM KH₂PO₄

- **PBST (0.05% Tween 20)**

0.5 ml Tween 20

995 ml PBS

- **5x protein loading buffer**

250 mM Tris-HCl; pH6.8

500 mM DTT

10% SDS

0.5% Bromophenol blue

50% Glycerol

- **10x TBE buffer**

1.8M Tris-base

1.8M Boric acid

20mM EDTA

2.1.5. Reagents and media for cell cultures (Invitrogen GmbH, Karlsruhe, Germany)

- Geneticin Selective Antibiotic G41 8 Sulphate
- Preservation solution 10% DMSO in FCS
- Foetal calf serum (FCS)
- DME Medium
- Penicillin-Streptomycin 10,000 U/ml Penicillin, 10,000µg/ml Streptomycin

2.1.6 Bacteria culture media

S. maltophilia medium

5 mM MnSO₄, 0.36 mM CaCl₂, 0.5 mM L-methionine, 0.8 mM MgSO₄, 2.2 mM K₂HPO₄, 3.7 mM KH₂PO₄, 6 mM (NH₄)₂HPO₄, 50 mM di-Sodium succinate, 2 g / L yeast extract were dissolved in 1 L H₂O, the pH value was adjusted to 7.2 and the medium was autoclaved. Subsequently, the addition of the sterile filtered antibiotics in the cooled medium: 40 mg / L Refobacin, 50 mg / L Claforan and 100 mg / L ampicillin was done.

LB medium

In one liter of H₂O, 10 g tryptone, 5 g yeast extract, 10 g NaCl were dissolved and the pH was adjusted to 7.2, and then the medium was autoclaved.

LB / Amp plates: 7.5 g agar in 1 L LB medium, autoclaved and then after cooling to 50 °C add 1 ml ampicillin (50 mg/ml).

2.2 Methods

2.2.1 Expression of *Stenotrophomonas maltophilia* protease

StmPr1 was expressed in *E.coli* BL21 (DE3). Bacteria were grown up to an optical density (OD_{600 nm}) of 1.0 in LB-Medium, and protein expression was induced by adding 0.05 mM IPTG (final concentration) to the culture medium. After 48 hours of incubation at room temperature, the medium was harvested.

2.2.2 Purification of *Stenotrophomonas maltophilia* protease

2.2.2.1 Ammonium sulphate precipitation

The protein solution was concentrated by adding 80% ammonium sulphate, and then the suspension was stirred for overnight at 4° C. Thereafter, for 30 minutes at 4° C and 15,000 rpm (Sorvall SS-40 rotor), the supernatant discarded and the protein sediment then dissolved and suspended in TBS buffer (20 mM Tris/HCl pH 7.5; 120 mM NaCl). Excess salt was removed by dialysis.

2.2.2.2 Fast Protein Liquid Chromatography (FPLC)

The StmP1 protein was purified by size exclusion chromatography and in a next step by anion exchange chromatography. The concentrated medium was centrifuged, the supernatant supplied to a superdex column (Pharmacia), and the proteins were eluted with TBS buffer. Enzymatically active fractions were pooled and supplied to the EMD® DEAE-650(s) (Merck) column to remove minor impurities and to obtain crystallization grade purity. Proteins were separated by a flow rate of 1ml/minute in buffer A (20 mM Tris-HCl pH 7.5), and eluted with a linear gradient from 0 to 100% of buffer B (20 mM Tris-HCl pH 7.5/0.8M NaCl) [55]. Measurement of enzymatic activity revealed that StmPr1 eluted at 400 mM NaCl. The enzymatically active fractions were analyzed by SDS page.

2.2.3 Photometric determination of protease activity of StmPr1

Serine protease assay:

For determining the serine protease activity the chromophore tetrapeptide Suc-Ala-Ala-Pro-Phe-pNA was used, which is considered to be a non-specific serine protease substrate [71]. After incubation of a serine protease with the substrate nitroaniline is released which absorbs at 405 nm. The amount of released nitroanilines is a measure of the activity of the serine protease. To determine the protease activity, 10 µl of a diluted protein solution mixed with 80 µl reaction buffer and the reaction was started by addition of 10 µl of the diluted substrate solution (10 mM). The absorbance measurements were made at intervals of 1 minute at 405 nm in the microplate photometer. The enzyme activity values were off a linear dependence of OD at 405 nm against time.

2.2.4 Determination of the IC₅₀ value for competitive inhibitors of StmPr1 Protease

The affinity of competitive inhibitors is described by the IC₅₀ value. This Value determines which inhibitor concentration is necessary to inhibit 50% of the enzyme activity. The IC₅₀ value was determined in this work for the peptide aldehyde inhibitors, chymostatin and leupeptin, and bortezomib by a non-linear regression. The protease inhibitors were pre-incubated with StmPr1 for 20 minutes at RT and then the reaction was started by addition of the substrate Suc-Ala-Ala-Pro-Phe-PNA. The enzyme activity values were obtained from the mean of three independent measurements. The analysis of the determination of the IC₅₀ value was performed by using Solver module in the Excel programme.

2.2.5 SDS-polyacrylamide gel electrophoresis (PAGE) [72]

In this work, discontinuous SDS-polyacrylamide gels were used consisting of a stacking gel and a separating gel. In this method, proteins accumulate in the stacking gel with larger pores and lower pH dependent manner under these conditions only on the charge of their mobility, however, the molecular size is affected. After admission to the separating gel with the higher pH the net charge of proteins increases and the mobility will be dependent on the molecular weight. By boiling with SDS, the proteins filled with negative charges and lose their original conformation and they will be

arranged according to their molecular weights. The proteins migrate according to the applied voltage with a mass proportional to their Velocity through the pores of the separating gel to the anode. The pore size of the gel is determined by the concentration of acrylamide and bisacrylamide. The SDS-gels were 104 mm X 98 mm X 1 mm. Before applying, the samples were 1:1 diluted in reducing Sample buffer and heated 4 min at 95 °C. 15 µl of the protein molecular weight marker (dilution 1:20 in reducing sample buffer) were applied.

2.2.6 Crystallization experiments

For the crystallization experiments the vapour diffusion method was applied. In a closed system containing a solution of the so-called precipitants, a drop with a highly concentrated protein solution is placed. In the precipitants are substances which are suitable for protein precipitation (salts, organic solvents, polyethylene glycol) are defined by buffer additives pH values adjusted. In the vapour diffusion effect, a slow concentration of the protein takes place under optimal conditions till the formation of protein crystals occurs [73, 74]. 1 ml Precipitants solutions (1.8 mM ammonium sulphate and 100 mM tris buffer at pH 8.0) was added to the wells, pipette equal parts of protein and precipitants solution (2 µl) to siliconized cover glass and close the cover by sealing the edges of the wells with silicone grease.

2.2.6.1 Crystallization of StmPr1

After completion of the first step of protein purification the protein was immediately subjected to the different stages of crystallization protocols.

2.2.6.2 Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) Spectroscatter 201 (Molecular Dimensions, UK) is an instrument which can be used to determine the size distribution profile of small particles in suspension or polymers in solution. Analysis of the autocorrelation function in terms of particle size distribution is done by the program CONTIN [75]. A truly monodisperse sample would give rise to a single exponential decay to which fitting of a calculated particle size distribution is relatively straightforward. Dynamic light scattering (also known as Photon Correlation Spectroscopy (PCS)) is the general

designation for a method to determine the size of small particles in the submicron range. These particles are found in Brownian motion in suspension or emulsion. The diffusion coefficient **D** for this is inversely proportional to the radius (hydrodynamic radius) **r_h** of the particles.

$$D = \frac{k \times T}{3 \times \pi \times \eta \times r_k}$$

(k = Boltzmann's constant)

Temperature **T** and viscosity **η** of the solutions are important parameters, which must be accurately known. Since particles in solutions have Brownian motion, the back scattered light will have a frequency shift (Doppler-Shift). The resulting intensity variation is recorded by a highly sensitive detector.

2.2.6.3 Pre-Crystallization Test

The PCT Pre-Crystallization Test was used to determine the optimal protein concentration to start crystallization screening. Highly concentrated samples resulted in amorphous precipitate, while diluted samples had produced transparent drops. Amorphous precipitate and clear drop production was avoided by changing the protein concentration accordingly.

The four reagents of the PCT kit, used to evaluate protein concentration for crystallization screening, are:

- 1) Reagent A1: 0.1M Tris Hydrochloride pH 8.5, 2.0M ammonium sulphate
- 2) Reagent B1: 0.1M Tris Hydrochloride pH 8.5, 1.0M ammonium sulphate
- 3) Reagent A2: 0.1M Tris Hydrochloride pH 8.5, 0.2M Magnesium Chloride Hexahydrate, 30% w/v Polyethylene Glycol 4,000
- 4) Reagent B2: 0.1M Tris Hydrochloride pH 8.5, 0.2M Magnesium Chloride Hexahydrate, 15% w/v Polyethylene Glycol 4,000

2.2.6.4 Robotic Screening

In structural genomics/proteomics, major advances in protein crystallization have been made using robotics, which has automated the crystallization experiment and reduced the amount of protein required by an order of magnitude, improving the reproducibility of the experiments and allowing a large number of set ups [76].



Figure 10: Zinsser Pipetting Robot (Digilab Genomic Solution, Germany).

2.2.6.5 Optimization of Crystals

The hanging drop vapour diffusion method has been applied to improve initially obtained crystals using the Linbro Plates. A 2 μ l protein solution was mixed with 2 μ l reservoir solution and equilibrated against 1.0 ml reservoir solution. X-ray diffracting crystals, possessing approx. dimensions of $0.35 \times 0.20 \times 0.05$ mm, were obtained by incubating plates at 20° C for three weeks. Optimal size crystals were produced using 1.8 M ammonium sulphate and 100mM Tris-HCl buffer at pH 8.0.

2.2.6.6 Data Collection

Native X-ray diffraction data from single crystals were collected by exposing them at the synchrotron Consortium-Beamline X13 DESY, Hamburg. A crystal was mounted on a nylon loop and flash-cooled in cold nitrogen-gas stream at 100 K. All intensity data were indexed, integrated and scaled with the HKL-2000 package [77]. Phase information was generated using the Molecular Replacement (MR) strategy. Therefore, the program Molrep [78] from the CCP4i suite [79] was used.

2.2.6.7 Matthews Coefficient (V_M)

One of the most important crystal parameters is the fraction of the unit cell (or asymmetric unit) that is considered as solvent. The most often used way to calculate

Materials and Methods

the *solvent content* is via the Matthews parameter [80]. The Matthews coefficient provides this number. The Matthews coefficient (V_M) is calculated as:

$$\frac{\text{volume of unit cell}}{\text{the molecular weight of macromolecule} \times Z \times X}$$

Z is the number of asymmetric units in the unit cell (i.e. the number of symmetry operators in space group). The unknown variable, X , is the number of molecules in the asymmetric unit.

2.2.6.8 Cryogenic Techniques

After successful data collection, a small Dewar was filled with liquid nitrogen to keep the internal environment in freezing temperature while crystal was still on the goniometer. The head of the cryo-tongs was placed into the nitrogen along with cryo-vial and forceps until they stopped boiling.



Figure 11: Cryogenic technique accessories.

For cryo-protecting and to keep the crystals, the tong was clamped around the crystal and quickly returned with the crystal to the liquid nitrogen in the Dewar. With another pair of forceps clamped the base of the cryo-pin, and the crystal was released from the cryo-tongs.

The crystal was transferred to the cryo-vial. During the whole procedure of crystal preservation, the crystal was kept strictly under the liquid nitrogen surface as little change in temperature can damage the crystal order and ultimately the data quality. The cryo-pin was screwed onto the cryo-vial using the magnetic wand and was transferred into the liquid nitrogen filled Dewar.

2.2.6.9 Macromolecular Crystallography Beamline X13

Beamline X13 is a bending magnet macromolecular crystallography beamline located in the HASYLAB Hall 5. It is a monochromatic fixed wavelength beamline operating at a wavelength of 0.8123 Å (15.3 keV). The beamline is equipped with a MARCCD detector (165 mm). The station is optimized for rapid data collection, high resolution studies.

2.2.6.10 Model Building and Refinement

After solving the phase problem, the three dimensional model of the protein and complexes were build using the programme Coot [81] and refined by the programme Refmac5 [82].

2.2.6.10.1 Coot: Crystallographic Object-Oriented Toolkit [81]

Coot is designed for macromolecular model building, model completion and validation, particularly suitable for protein modelling using X-ray data. Coot displays maps and models and allows model manipulations such as idealization, real space refinement, manual rotation/translation, rigid-body fitting, ligand search, solvation, mutations, rotamers, demonstration of Ramachandran plots, skeletonization, display of non-crystallographic symmetry and much more.

2.2.6.10.2 Refmac5

Refmac5 is a refinement program for macromolecular structures. The Refmac5 program can carry out rigid body, restrained or unrestrained refinement against X-ray data. Refmac5 will refine an atomic model by adjusting the model parameters (coordinates, B-factors, etc.) in order to obtain the model which best explains the experimental data (i.e. maximizes the likelihood). Progress is monitored by the R-factor and Free R-factor [83], as well as by the likelihood scores themselves [82].

2.2.6.11 Ligand Binding Experiments

The main use of protein structures in medicine and biotechnology is of course guiding medicinal chemists in their effort to synthesize better compounds to bind with selected target proteins. To do this, it is necessary to analyze the structures of protein-ligand complexes. Many of drug discovery projects are supported by X-ray

Materials and Methods

crystallographic efforts and two common ways to obtain these ligand protein complexes are co-crystallization and crystal soaking. Both co-crystallization and crystal soaking experiments with different concentrations of inhibitors were performed.

2.2.7 Test StmPr1 in cell cultures

The human lung epithelial tumour cells were cultured in Dulbecco's Modified Eagle Medium (DMEM), 10% fetal calf serum and 1% penicillin / streptomycin at 37°C under 5% CO₂ gassing at 80% confluence drawn. The media was exchanged with DMEM without fetal calf serum and the tested inhibitors were added in different concentrations in the presence and absence of both pure StmPr1 enzyme and wild type *Stenotrophomonas maltophilia*.

3. Results and Discussion

3.1 Expression and Purification of StmPr1

The product of the StmPr1 gene of *Stenotrophomonas maltophilia* is the major secretory protease and considered to be a pathogenic factor. In order to elucidate its potential use as a drug target, the protein has to be purified, and if possible, its three-dimensional structure has to be analysed by X-ray crystallography. Such projects normally require milligram amounts of a highly purified protein. Protein expression and purification are often rate limiting steps, but the availability of a standardized protocol has streamlined this process. Although the StmPr1 protease had been directly purified from cultures of *Stenotrophomonas maltophilia* [54], but the protein amounts obtained by this means would hardly meet the needs of crystallization experiments. Therefore, a bacterial expression system was adopted which had been developed in the working group [55]. The protein purification protocol involves several steps as described in the materials and methods, with some modifications which will be mentioned later.

StmPr1 protease was expressed in *E. coli*, as a secretory protein which was processed from a precursor to the active form, accumulating in the culture medium. The cell-free supernatant was concentrated by ammonium sulphate precipitation, and the dissolved protein fraction was applied to a gel filtration column. Fractions with proteolytic activity were then subjected to anion exchange chromatography, and the active fractions were pooled and concentrated.

Table 1: Purification Steps of StmPr1

Purification process	Volume (mL)	Total protein (mg)	Total activity (U)	Specific activity (U/mg)	Purification (fold)
Culture medium	1000	6195	30,5	0,0049	1
Amm. sulfate	25	1194	22,6	0,0189	4
EMD BioSEC 650	48	368,5	12,3	0,0334	7
DEAE-Sephadex	14	4,01	2,26	0,565	115

A typical purification procedure (Table 1) yields about 4 mg of protease from one litre of expression medium. Only 3 steps were necessary to obtain pure protein as confirmed by SDS-PAGE, which demonstrated a single band (Figure 12) indicating a homogeneous preparation.

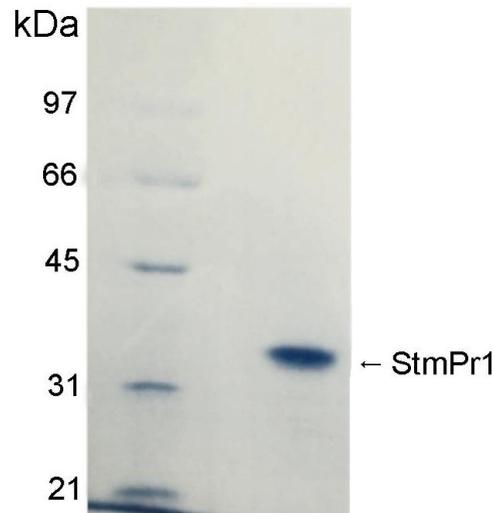


Figure 12: SDS-polyacrylamide gel electrophoresis of purified StmPr1.

The molecular weight of the protein under denaturing conditions was estimated to be 36 kDa. This is in contrast to the previously reported protein molecular weight of 47 kDa for the mature protease, which is processed from a 63 kDa precursor (see Introduction). Obviously a further C-terminal truncation has occurred. This may have resulted from the modified purification protocol as the harvested culture medium was concentrated by overnight ultrafiltration. During this period, the protein undergoes an auto-processing step ending up with the new 36 kDa form.

Enzyme activity and substrate affinity of this truncated protease were comparable with the native enzyme produced by *S. maltophilia* [54]. When assayed with increasing concentrations of substrate (Figure 13), a specific enzyme activity of 12 U/mg could be calculated and K_m of approximately 0.5 mM.

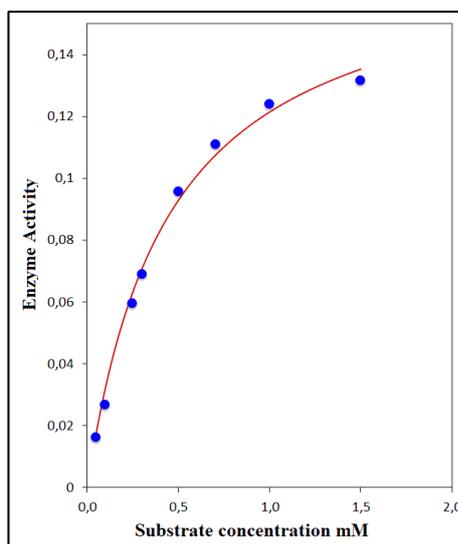


Figure 13: StmPr1 activity at different substrate concentrations.

The specific activity of StmPr1 was calculated using different concentrations of N-succinyl-Ala-Ala-Pro-Phe-4-nitroanilide substrate 0.05-1.5 mM.

It is known that the truncated proteins reduced to a stable core domain are often less flexible thereby facilitating the crystal growth. The former attempts to crystallize the 47 kDa protein had not been successful but the 36 kDa core protein motivated us to try the crystallization approach for the structure analysis.

3.2 Crystallization of StmPr1

The first step of the 3-dimensional structure determination of a protein is to obtain high quality protein crystals. It is common known that this is the bottleneck in many projects. Protein crystal growth is not predictable, and numerous experimental trials are necessary to find the proper conditions. To get regular X-ray-suitable crystals, a lot of optimization steps are required.

Besides the purity of the protein, there are many factors which have to be optimized to facilitate the crystal growth such as protein concentration, precipitant, pH, temperature, etc. Crystallization depends on protein solution conditions, the molecular structure may be affected along with a subsequent change in the particle size. Thus monitoring the size of a protein molecule is one way of observing stability in the protein solution under its native conditions. Proteins can aggregate to heterogeneous complexes thus preventing crystallization, so a fast, accurate, and simple technique which is called Dynamic Light Scattering (DLS) was performed. This technique is used to determine the size distribution profile of the protein particles in solution, and it

is shown to be ideal for protein characterization and to check the monodisperse nature of the protein solution.

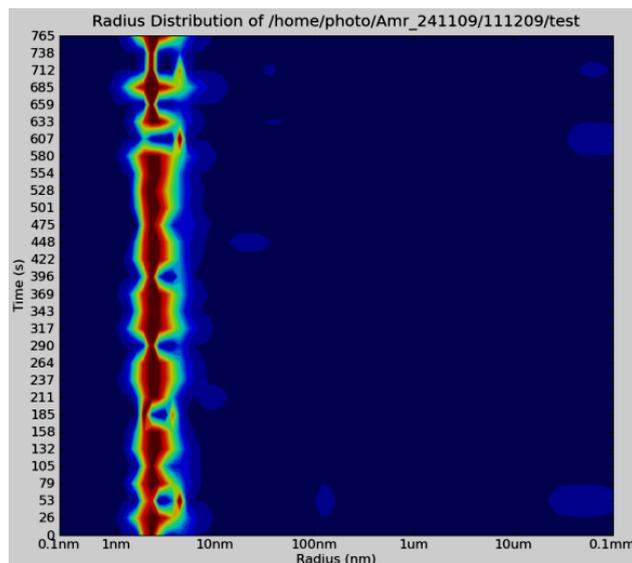


Figure 14: DLS measurement showing a monodisperse protein solution.

When applied to the StmPr1 preparation, a strong single signal was obtained indicating that the protein existed in a homogeneous non-aggregated form (Figure 14). The purified samples of StmP1 indicated a molecular mass of 36 kDa on SDS-PAGE and the results of dynamic light-scattering (DLS) showed that the hydrodynamic radius (R_H) of StmPr1 of a protein concentration 10 mg/ml was 2.7 nm, assuming a globular shape of the protein and the molecular mass corresponding to this hydrodynamic radius can be estimated to be about 36 kDa. This indicates that all StmPr1 molecules existed in monomeric form in solution.

After having demonstrated the homogeneity of the protein solution, the next important step was to determine the precipitant which will be used in crystallization. Since it is known from statistical studies that ammonium sulphate is the precipitant which can be used with more than 60 % proteins. The crystal growth was optimized using the hanging drop method varying the concentration of ammonium sulphate (the precipitant) from 1.2 – 2 M and also varying the pH from 5-9. Another important factor which affects crystallization is protein concentration. Too concentrated samples result in amorphous precipitation, while too diluted samples will result in clear drops. To address this point, several protein concentrations ranging from 7-15 mg/ml have been checked with the previously mentioned ammonium sulphate concentrations rang. This test showed that the protein concentration 11 mg/ml was the appropriate

concentration to start crystallization. Regular uniform crystals were obtained as shown in (Figure 15) using the following optimized crystallization condition (1.8M ammonium sulphate, 0.1M Tris at pH 8.0).

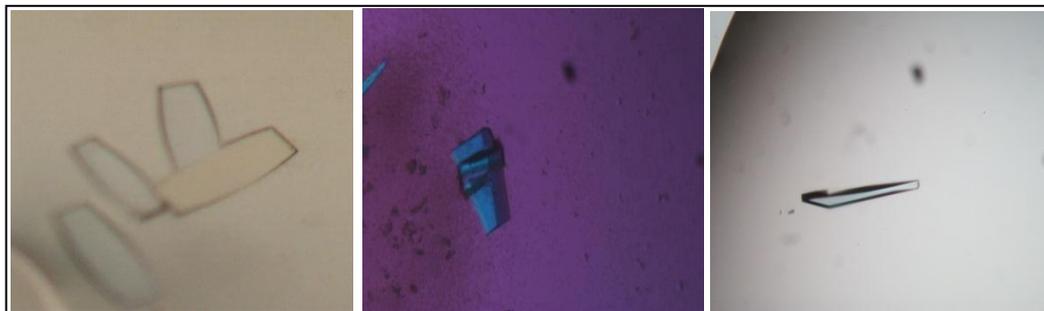


Figure 15: StmPr1 Crystals.

2 μ l StmPr1 mixed with 2 μ l precipitant solution (1.8M ammonium sulphate, 0.1M Tris at pH 8.0), and incubated at 20 °C temperature. After three weeks good shaped crystals with dimensions of around $0.30 \times 0.20 \times 0.05$ mm were obtained.

3.2.1 Data Collection

The next key step in the structure determination process is the collection of X-ray diffraction data. Native diffraction data were collected at the Consortium-Beam line X13, DESY Hamburg. The detector type used was a Mar CCD 165 mm and the wavelength was 0.8123 Å. Data collection was performed at 100 K and ice production was prevented by soaking the crystals in the same mother liquor with 15% glycerol just before mounting. Sufficient data with a maximum resolution of 1.32 Å were collected. The data were processed using the HKL-2000 program package [77]. The space group was identified to be $C222_1$ with unit cell dimensions of $a = 60.17$ Å, $b = 86.10$ Å and $c = 131.40$ Å. Matthews coefficient [80] ($V_M = 2.12$ Å³ / Da) suggested one molecule in the asymmetric unit with a solvent content of 42 %.

3.2.2 Structure Solution

The last part of the structure determination process is the construction of a 3D model. To calculate the electron density, phase information was obtained by Molecular Replacement (MR) using the program Molrep [78] from CCP4i suite [79]. A search model was constructed using the homologous structure of *Dichelobacter nodosus* (PDB ID: 3LPA) as the reported primary amino acid sequence of StmPr1 showed 48 % sequence identity. Different strategies were applied on this model using the original

Results and Discussion

model as it and the residual replacement with polyalanines. Fortunately, statistical values were satisfying and phase information was obtained successfully (Table 2).

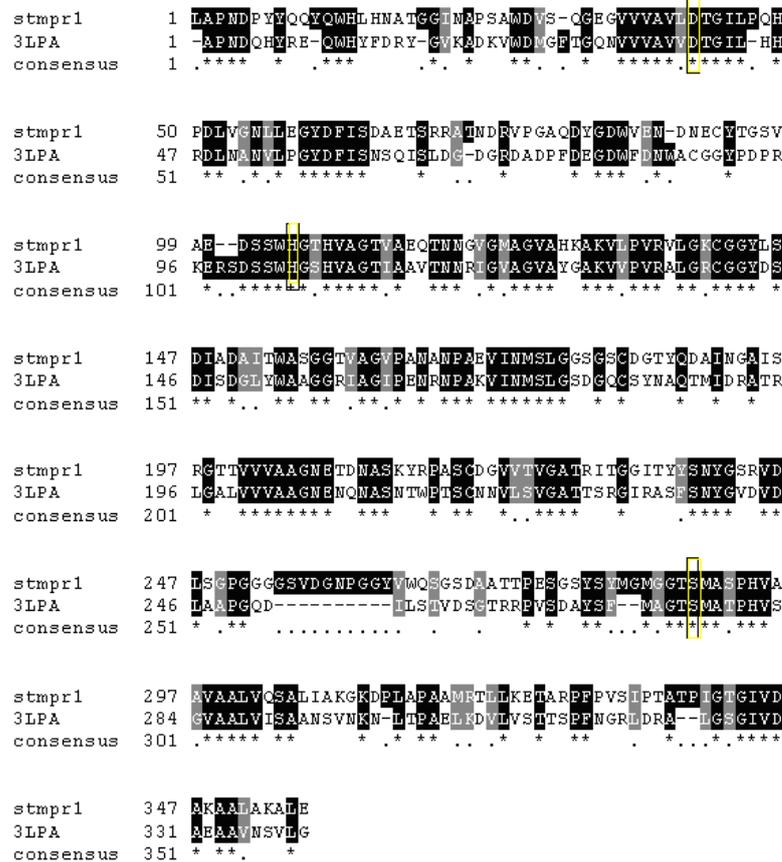


Figure 16: Sequence alignment between StmPr1 and *Dichelobacter nodosus* Subtilisin (PDB ID: 3IPA). The figure was prepared using the BOXSHADE server [84].

The active site residues are marked with open boxes and the conserved residues that are shared with *Dichelobacter nodosus* subtilisin are marked with dark boxes.

Figure 16 shows the alignment of the StmPr1 sequence against *Dichelobacter nodosus* subtilisin, which indicates that both proteins have the putative catalytic triad characteristic for serine proteases.

3.2.3 Model Building and Refinement

Substantial phase information was generated mainly with good statistical values. Model building and refinement was performed using the programs Coot and Refmac5. The crystallized 36 kDa catalytic domain of StmPr1 proteinase encompasses amino acids 151–506 of the 617 amino acid of the prepro-enzyme. The model was refined at a resolution of 1.4 Å with R_{crystal} of 15.4 % and R_{free} of 16.7%. The data collection and refinement parameters are shown in Table 2.

Table 3: Data collection parameters and refinement statistics of the native StmPr1

Parameters	
Space group	C222 ₁
a (Å)	60.165
b (Å)	86.099
c (Å)	131.398
V _M (Å ³ / Da)	2.12
Solvent content (%)	42%
Completeness of data (%)	99.6 (100)
No. of total reflections	
No. of unique reflections	63714 (4715)
Average I/sigma intensity	19.77 (3.20)
Resolution (Å)	46.17-1.40 (1.436-1.40)
Redundancy	5.1(4.6)
Rmerge (%)	6.6 (48.5)
No. of reflections used in refinement	63714
Rcrystal (%)	15.36 (17.4)
No. of reflections used in Rfree	3185
Rfree (%)	16.67 (18.8)
Protein atoms	2734
Amino Acids:	356
Calcium Ions	1
Sulphate Ions	5
Glycerol molecules	2
Water molecules	352
R.m.s.d in bond lengths (Å)	0.004
R.m.s.d in bond angles (°)	0.929
R.m.s.d in torsion angles (°)	6.2
B-factor estimated from Wilson plot (2)	19.6
Mean B-factor for side chain atoms and waters (Å ²)	15.7
Mean B-factor for all atoms (Å ²)	13.9
Residues in the most allowed regions (%)	89.5
Residues in the additionally allowed regions (%)	10.5
Residues in the generously allowed regions (%)	0.0

3.2.4 Description of the Molecule

The full length amino acid sequence for StmPr1 has already been reported [55]. Some of the primary sequence statistics is summarized in Table 3. The statistics has been prepared applying the ExPASy Proteomics Server [85].

Table 4: Statistical values of the primary amino acid sequence

Chain ID	Chain A
Formula	C ₁₅₇₃ H ₂₄₄₆ N ₄₄₂ O ₅₂₁ S ₁₀
No. of Amino acids	356
Molecular Weight (Da)	36205.9
Theoretical pI	4.73
Total No. of atoms	4992
No. of negative residues	34
No. of positive residues	19
No. of Cysteines	4

A total of 356 amino acids have been detected in the structure which corresponds to the protein molecular weight of 36 kDa. Also from the amino acid sequence, it is clear that the processing is done from the C-terminal to prove the idea of C-terminal truncation of this protein. The protein has two disulphide bonds.

The three-dimensional structure of the StmPr1 proteinase shows the scaffold characteristic of subtilisin-like serine proteinases. It consists of six helices, one 3/10 helix, a β sheet made of seven parallel strands and two β sheets made of two antiparallel strands (Figure 17). Determination of the structure confirms the presence of two disulfide bonds, Cys93–Cys141 and Cys183–Cys220. One calcium-binding site is found in StmPr1.

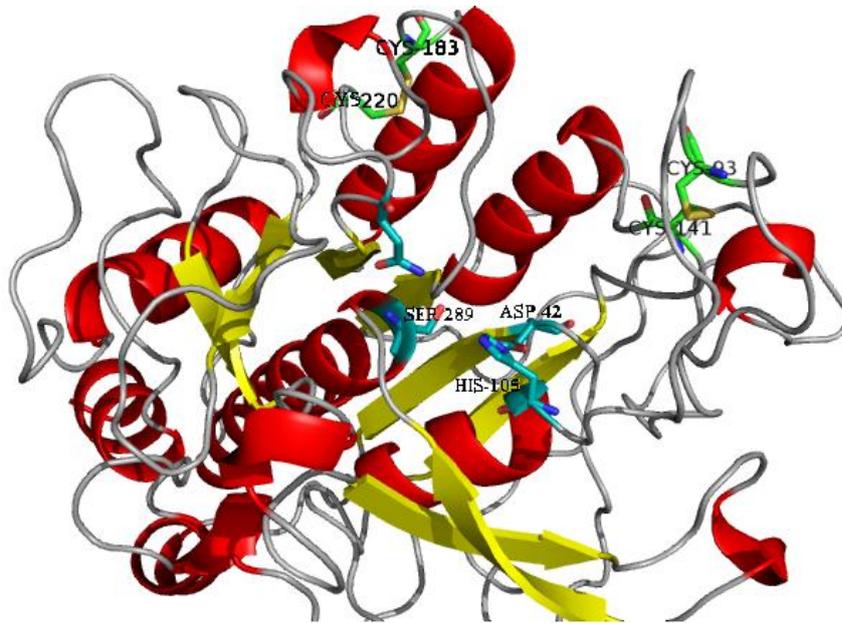


Figure 17: Model of the StmPr1 crystal structure.

The residues of the catalytic triad D-42, H-105 and S-289, are shown in cyan sticks, and the disulfide in yellow bridges between cysteine residues in green sticks.

The active site of StmPr1 consists of the catalytic triad Asp-42, His-105 and Ser-289, and substrate recognition and binding sites that are well conserved among subtilases [38]. Although it differs from other subtilisins in that, the catalytic serine is far away from the other catalytic triad residues in sequence. The substrate-binding site in StmPr1 appears on the surface as a relatively distinct cleft in which the substrate is accommodated by forming a triple-stranded antiparallel β sheet with residues of the S4- and S3- binding sites.

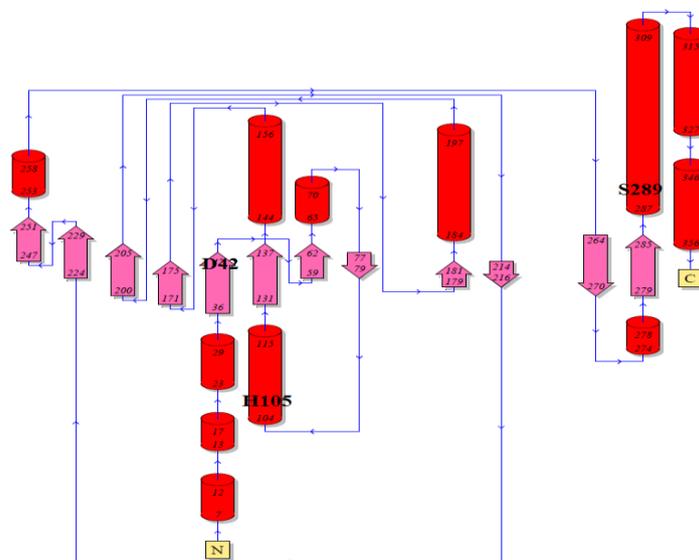


Figure 18: A topology diagram of the StmPr1 structure.

α -helices are shown as cylinders, β -sheet strands as arrows, and the locations of the catalytic triad residues D-42, H-105 and S-289 are indicated as well.

The topology of a protein structure is a highly simplified description of its fold including only the sequence of secondary structure elements, and their relative spatial positions and approximate orientations. This topology cartoons are useful to understand particular folds and making comparisons between similar subtilisins fold. From this topology diagram, it is observed that the StmPr1 protease has the α/β fold catalytic centre containing the typical seven-stranded parallel β -sheet, which was observed with other subtilisins [86].

The substrate-binding site in StmPr1 appears on the surface as a relatively distinct cleft (Figure 19) in which the substrate can be accommodated by forming a triple-stranded antiparallel β -sheet with residues of the S4- and S3-binding sites (nomenclature of subsites, S4–S2', is according to Schechter and Berger [87]). The bottom of the S1 substrate-binding pocket is made up of residues Ser-176 – Ser-289 and the oxyanion hole residue Asn-207. The substrate-binding cleft appears to be relatively opened. The S4 subsite is occupied by a larger residue, typically a tyrosine, which is assumed to form a flexible lid on the S4 pocket [88].

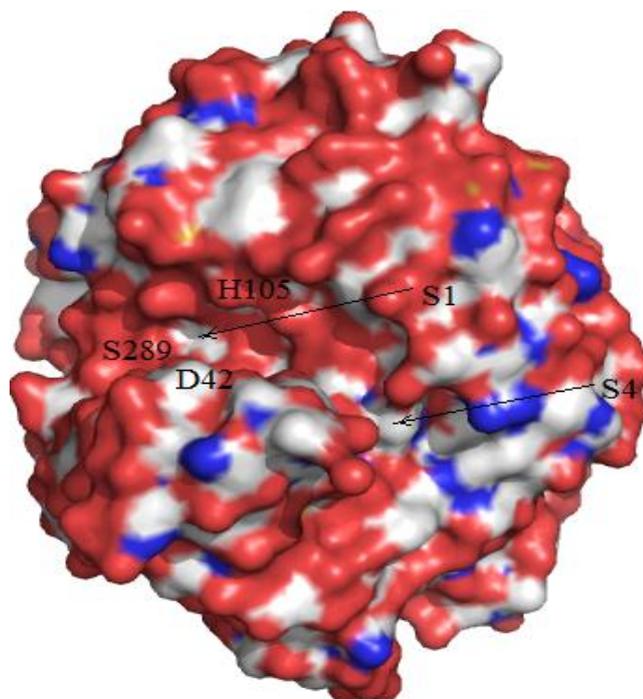


Figure 19: Electrostatic potential surface of StmPr1.

Surface is colored according to electrostatic potential, the positively charged electrostatic potential is coloured blue and negatively charged electrostatic potential is coloured red. The location of the active site residues is indicated and the location of the S1 binding pocket is indicated as well.

The chemical properties of the groups comprising the protein surfaces are important, as these groups determine the important interactions of the protein with

water. It was suggested that large polar surfaces of proteins contribute to the increased stability. StmPr1 proteinase exhibits some structural features, like the high exposure area of Asp residues on the surface. It has a high content of Glu and Asp residues, and a low content of Arg and Lys residues on the surface this may give the ability to the enzyme to withstand the alkaline conditions. This is in agreement with the finding of StmPr1 to be an alkaline protease with an optimum pH of 9 [54].

3.2.5 Overall structure comparison with related enzymes

StmPr1 showed some distinct features, and it is worth now to compare the StmPr1 proteinase with the classical subtilisins with respect to folding and loop conformation. Searching the protein data bank for subtilisins which have highest sequence homology with StmPr1 proteinase revealed several examples. Next, The comparison of StmPr1 proteinase with classical subtilisins such as the cold-adapted *Vibrio* subtilisin-like protease (PDB ID: 1SH7), proteinase K (PDB ID: 1IC6) and thermitase (PDB ID: 1THM) will be shown.

Comparing StmPr1 in blue with the cold-adapted *Vibrio* subtilisin-like protease (PDB ID: 1SH7) in green is shown in Figure 20. 356 residues from StmPr1 against 281 residues from 1SH7 are aligned. 231 atoms have been aligned with RMS deviation for all $C\alpha$ -atoms of 0.75 Å. Some deletions and insertions have been observed as shown in Figure 20.

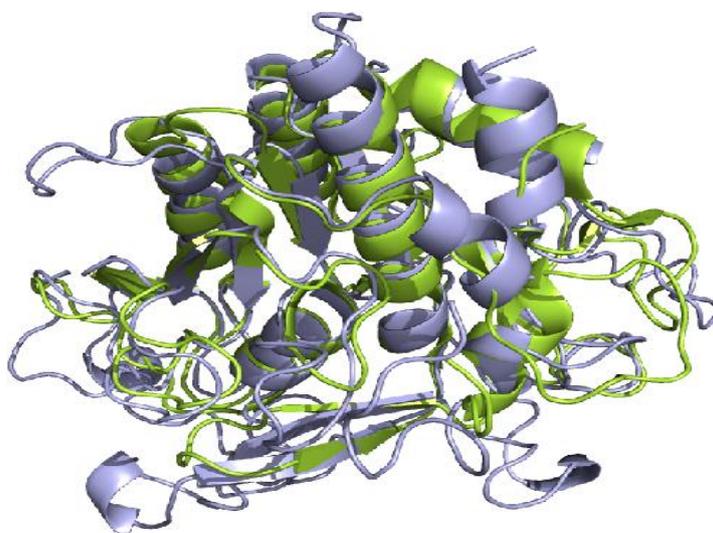


Figure 20: Cartoon plot of the superposition of StmPr1 (blue) with the cold-adapted *Vibrio* proteinase (1SH7, green), the structure was oriented to show the places of homology.

StmPr1 is then compared with another example from subtilisin family which is proteinase K. 356 residues from StmPr1 were aligned against 279 residues from Proteinase K. 246 atoms have been aligned with RMS deviation for all C α -atoms of 0.91 Å. This higher RMS value is in consistence with Figure 21, which showed a lot of deletions and insertions of loops and helices.

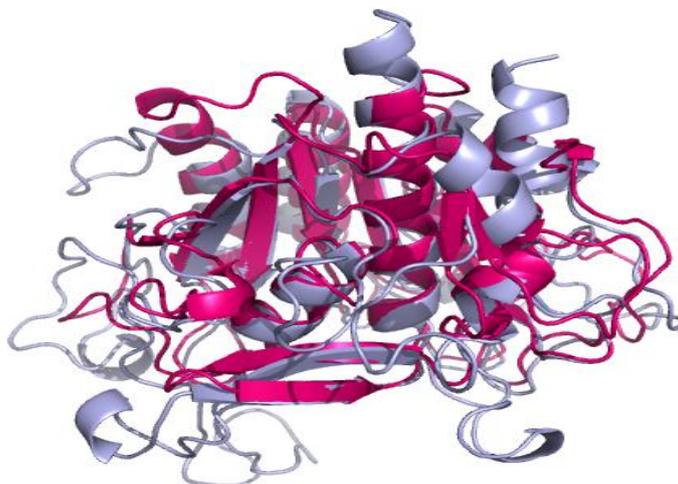


Figure 21: Cartoon plot of the superposition of StmPr1 (blue) with proteinase K (red), the structure was oriented to show the places of homology.

The comparison of the secondary structure of StmPr1 in blue with thermitase in yellow as a known member from the classical subtilisins family is shown in Figure 22. 356 residues from StmPr1 against 279 residues from thermitase were aligned. 266 atoms have been aligned with RMS deviation for all C α -atoms of 0.66 Å. Some insertions have been observed.

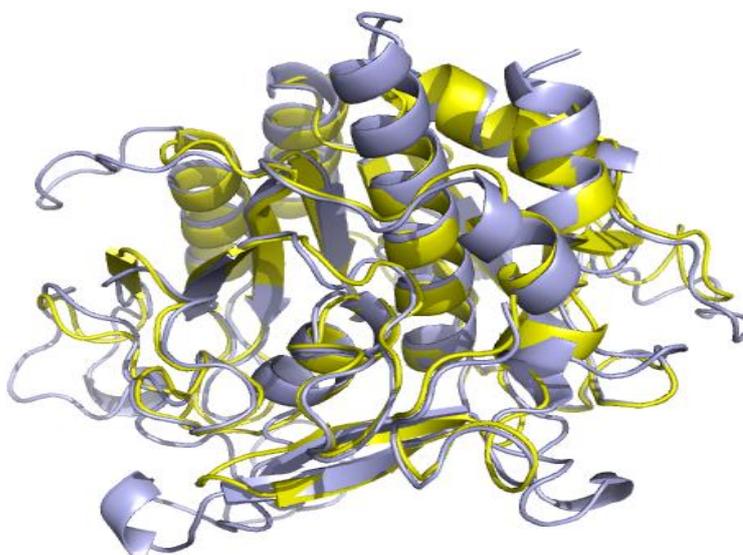


Figure 22: Cartoon plot of the superposition of StmPr1 (blue) with thermitase (yellow), the structure was oriented to show the places of homology.

The high resolution of all four structures allows reasonable comparison with respect to the quality of the models. Pairwise least square superposition of the four structures showed that 63–76% of the C α -atoms lie at common positions and gave a root mean square deviation of 0.66–0.91 Å (Table 4, Figures 20-22). The structural resemblance with regard to root mean square deviation, fraction of common C α -atoms and the amino acid sequence identity, is in the following order StmPr1–1THM > StmPr1–1SH7 > StmPr1–1IC6.

Table 5: Pairwise superposition of C α -atoms in StmPr1, the cold-adapted *Vibrio subtilisin*-like protease (PDB ID: 1SH7), proteinase K (PDB ID: 1IC6) and thermitase (PDB ID: 1THM).

	StmPr1–1THM	StmPr1–1SH7	StmPr1–1IC6
Number of residues	356 - 279	356 - 274	356 - 279
Aligned residues	266 (76%)	229 (72%)	246 (63%)
Root mean square deviation (Å)	0.66	0.75	0.91

The comparison of StmPr1 in blue with the *Dichelobacter nodosus* subtilisin-like protease (PDB ID: 3LPA) in orange is shown in Figure 23. 356 residues from StmPr1 against 340 residues from subtilisin-like protease 3LPA were aligned. 329 atoms have been aligned with RMS deviation for all C α -atoms of 0.55 Å.

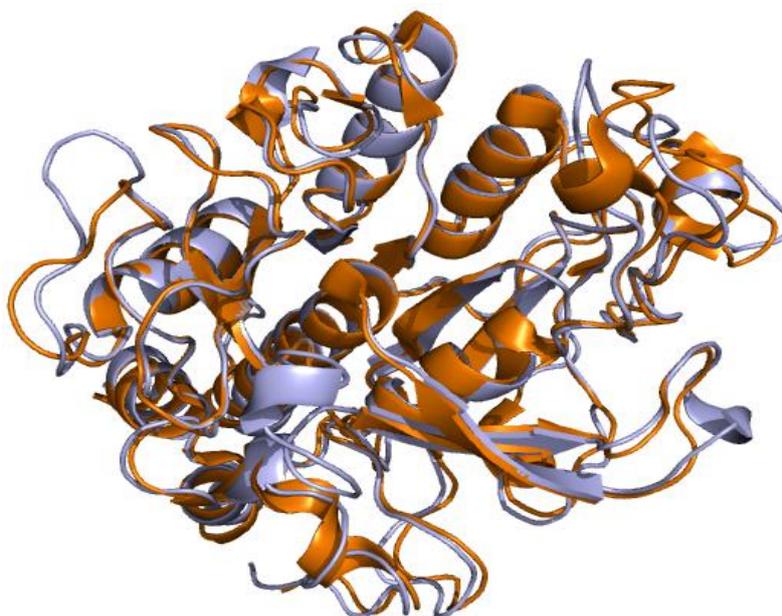


Figure 23: Cartoon plot of the superposition of StmPr1 (blue) with the subtilisin-like protease (3LPA, orange), the structure was oriented to show the places of homology.

As opposed to the classical subtilisin which showed about 63%-76% structure homology. The pairwise least square superposition of the structure non-classical subtilisin from *Dichelobacter nodosus* (PDB ID: 3LPA) against StmPr1 showed that 80% of the C α -atoms lie at common positions and gave a root mean square deviation of 0,55 Å (Table 5, Figure 23).

Table 6: Pairwise superposition of C α -atoms in StmPr1, with *Dichelobacter nodosus* subtilisin 3LPA

	StmPr1 - 3LPA
Number of residues	356 - 340
Aligned residues	329 (80%)
Root mean square deviation (Å)	0.55

The Comparison of Stmpr1 proteinase with subtilisin-like proteases reveals several major insertions (termed L1-L5) in the loops that surround the active site cleft (Figure 24). Most notable is the large well-ordered L2 loop (residues 81–102) that is tethered to the subtilisin-like fold by a disulphide bond between Cys-141 and Cys-93.

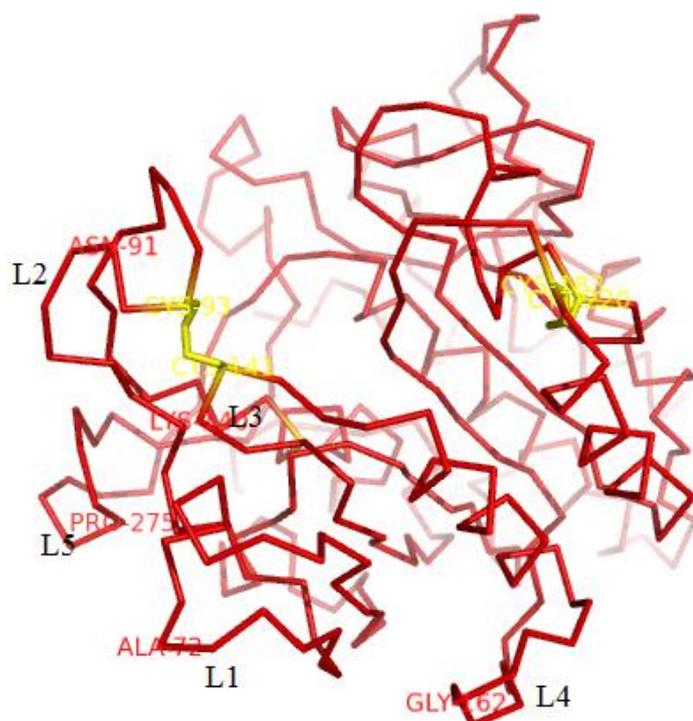


Figure 24: C α -backbone plot of StmPr1 showing the loop insertions around the active site and the disulfide bridges providing the loop stability.

The L2 loop is well defined in the electron density as shown in Figure 25, with moderately high B-factors (Figure 26), suggesting some flexibility and mobility. However, the apparent stability of the L2 loop is likely arises from crystal packing but it is uncertain whether this conformation would be favored in solution or not.

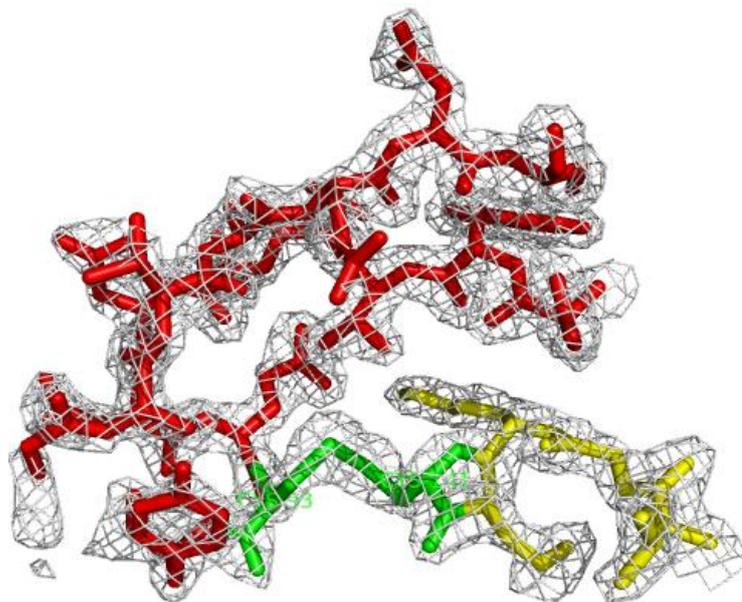


Figure 25: (Fo-Fc) electron density map depicting the disulfide bridge and the L2 loop of the StmPr1.

The map is contoured at 1.2σ . The disulfide bridge is shown in green, while the L2 loop in red. Water molecules have been removed for clarity.

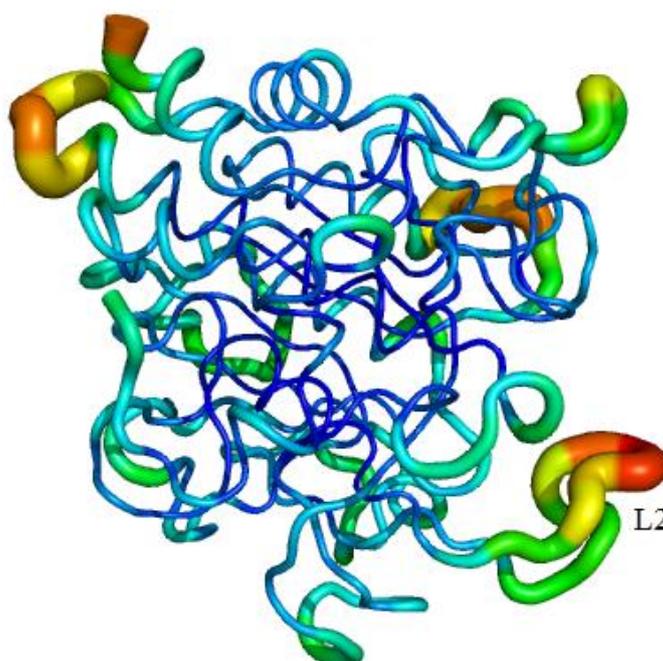


Figure 26: StmPr1 structure colored according to B-factor value for the side chains. The L2 loop shows a moderately high B-factor.

In conclusion, the comparison of StmPr1 structure with different subtilisins makes the StmPr1 protease different from the classical subtilisins and may allow defining a new sub-family of non-classical subtilisins.

3.2.6 Calcium-binding site

The presence of bound calcium ions is a feature shared by members of the subtilisin superfamily, where calcium binding has been shown to be essential for correct folding and structural stability [52, 89]. Considering the stabilizing effect of binding metal ions in many proteins, it would be expected that increased affinity and the number of bound metal ions should correlate with the thermostability of proteins. It is known that the integrity of the Ca^{2+} -binding site is important. Hence, its modifications or mutation may abolish the catalytic activity of the enzyme, as this leads to the absence of autoprocessing and inactivation of the enzyme.

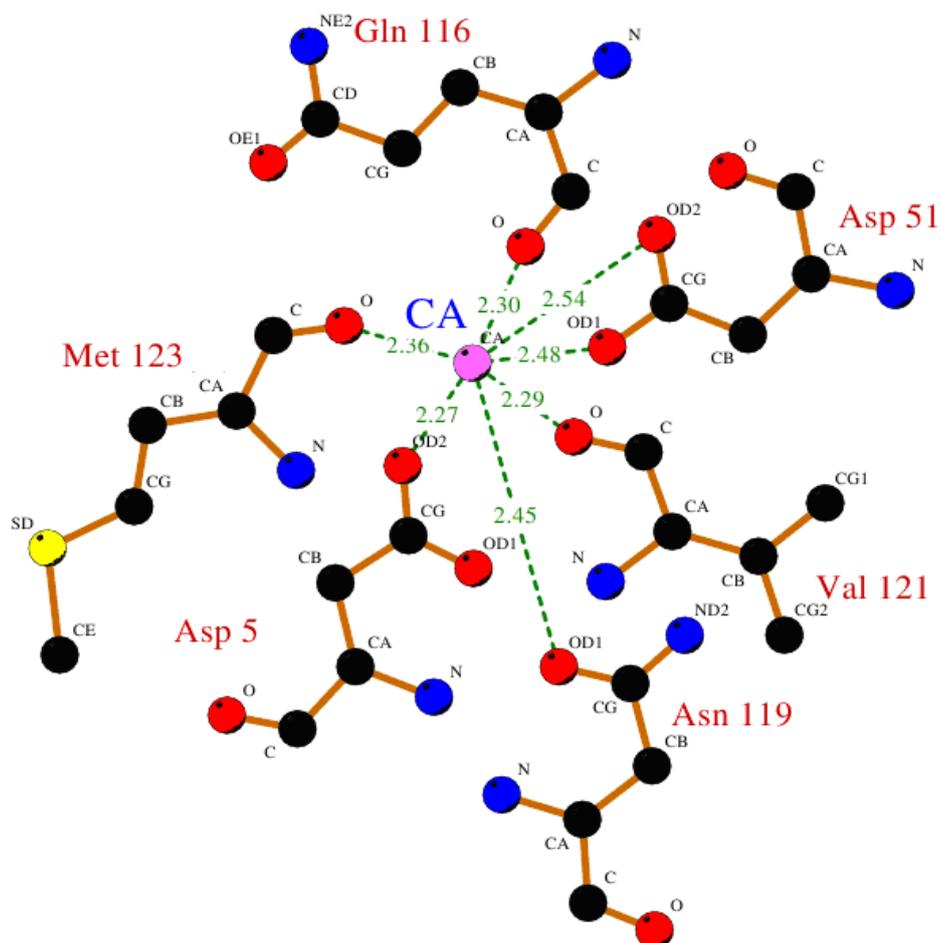


Figure 27: StmPr1 ligplot showing the calcium-binding site.

The calcium ion is coordinated in a pentagonal bipyramidal manner (Figure 27) by the carboxyl groups of Gln-116, Val-121 and Met-123, and the side chain oxygen atoms of Asp-5, Asp-51, and Asn-119. This Ca^{2+} ion (shaded with pink) in the StmPr1 proteinase is well conserved among the most related enzymes of the subtilisin family.

Table 7: The ionic interactions in calcium binding site

Nr.	Residue name	Atom name		Atom name	Distance (Å)
1.	Asp-5	O	\leftrightarrow	Ca^{2+}	2.27
2.	Asp-51	OD1	\leftrightarrow	Ca^{2+}	2.48
3.	Asp-51	OD2	\leftrightarrow	Ca^{2+}	2.54
4.	Gln-116	O	\leftrightarrow	Ca^{2+}	2.30
5.	Asn-119	OD1	\leftrightarrow	Ca^{2+}	2.45
6.	Val-121	O	\leftrightarrow	Ca^{2+}	2.29
7.	Met-123	O	\leftrightarrow	Ca^{2+}	2.36

Differences in stability and kinetic properties have been related to fewer or weaker metal ion binding sites [90]. In case of thermitase, differences in calcium binding were considered as one of the major reasons for the enhanced stability of the enzyme compared to its mesophilic counterparts [91]. One calcium ion is found associated with the structure of StmPr1, whereas each of 1IC6 and 1THM has two each and 1SH7 has three. In comparing the calcium-binding site in StmPr1 with the classical subtilisins, it was found that it is analogous to the known strong calcium-binding site in proteinase K [92]. According to sequence alignments, this site is well conserved among members of the proteinase K family including enzymes of thermo- and mesophilic origin most related to the StmPr1 proteinase. It is well known that Ca^{2+} ions exhibit significant effects on the stability and the folding kinetics in subtilisin [93, 94]. This is in accordance with the observed stability of StmPr1 when preparations could be used for long time periods without loss of enzyme activity and of crystallizability.

The structure of the StmPr1 protease reveals some differences in the architecture of the active site and the overall secondary structure compared to the

classic subtilisins and other serine proteases. These differences can be utilized for the development of specific drugs. The active site has to be characterized as well and studying the binding site, the sub-sites and amino acid residues involved in the binding. Therefore, it may be possible to take advantage of the structural particularities of StmPr1 and to find inhibitors with high specificity.

3.3 Inhibition of StmPr1 by commercially available compounds

Peptide aldehydes have been reported as inhibitors of all of the four major classes of proteolytic enzymes (serine proteases, cysteine proteases, aspartyl proteases, and metalloproteases). They have been used extensively as inhibitors of serine proteases. They may be promising for use as lead compounds in drug discovery (although their therapeutic use is limited to parenteral application). They are mostly composed of amino acids and therefore have few metabolic degradation products that are toxic. In case of serine proteases, they form a hemiacetal linkage with the active site serine residue. Such complexes are similar in structure to the proposed tetrahedral intermediate formed during peptide hydrolysis so that peptide aldehydes are regarded as transition-state analogue inhibitors of peptide hydrolases.

The concept of the active site was introduced in 1967 [87, 95], and the amino acid residues of an interacting ligand towards the N-terminus from the peptide bond to be cleaved are designated as P1, P2, P3, etc. Most subtilases are non-specific peptidases with a preference for an aromatic amino acid residue at the P1 position of the substrate/inhibitor [96] such as chymostatin which has a phenyl alanine residue at the P1 position. It is questionable whether the StmPr1 protease has also this preference and whether different amino acids in this position would change the inhibitory potential. Therefore, several protease inhibitors, namely chymostatin, leupeptin, calpain inhibitor, *z-leu-leu-leu-al*, and *z-leu-leu-phe-al* were compared (Table 7) and were tested for inhibition of StmPr1 (Figure 29).

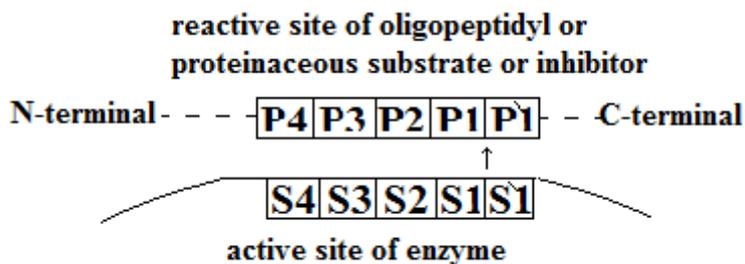


Figure 28: Scheme representation of the commonly used nomenclature for the enzyme subsites and the corresponding ligand sites.

This figure shows the amino acid residues of a substrate or inhibitor (P1 etc.) and the corresponding subsites (S1 etc.) of the active site of the enzyme. The arrow indicates the peptide bond to be hydrolysed by the enzyme.

Table 8: Sequences of the selected inhibitors

The inhibitor z-leu-leu-leu-*al* is abbreviated as L-L-L, while z-leu-leu-phe-*al* is abbreviated as L-L-F

	chymostatin	leupeptin	calpain inhibitor	L-L-L	L-L-F
P1	Phe- <i>al</i>	Arg- <i>al</i>	Met- <i>al</i>	Leu- <i>al</i>	Phe- <i>al</i>
P2	Leu	Leu	Leu	Leu	Leu
P3	<i>cyclic-Arg</i>	Leu	Leu	Leu	Leu
P4	Phe <i>deriv.</i>	–	–	–	–

All inhibitors shown in Table 7 contain an aldehyde group at the C-terminus (P1 position). In case of chymostatin and L-L-F, this position has aromatic amino acid. All five compounds were able to inhibit the StmPr1 enzyme activity, but to a different extent. This was demonstrated by dose response curve (Figure 29) and the calculation of IC₅₀ values.

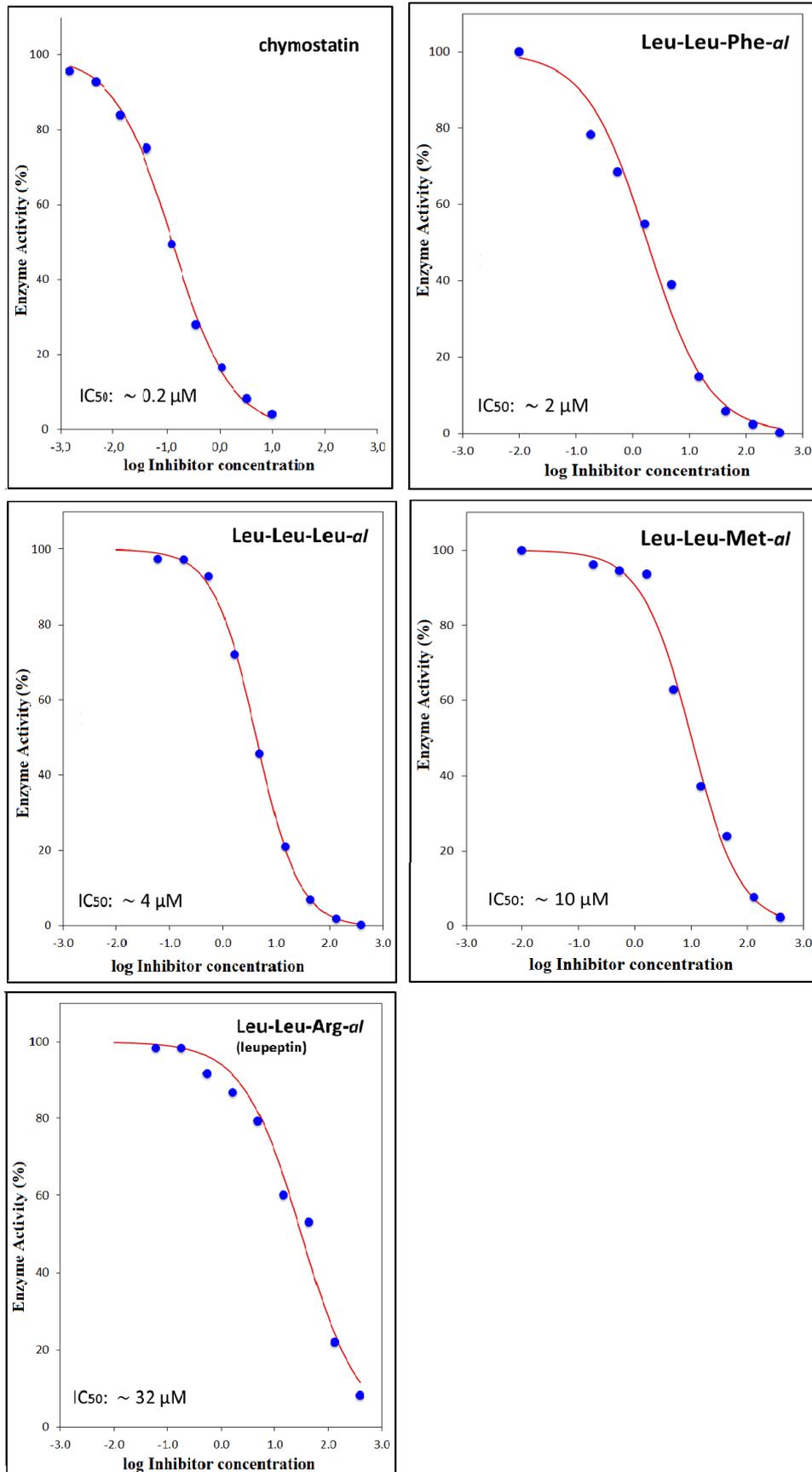


Figure 29: Inhibitory effects of the selected commercially available inhibitors.

The StmPr1 protease was incubated with different concentrations of each inhibitor for 20 minutes prior to addition of the chromogenic substrate Suc-Ala-Ala-Pro-Phe-PNA. The IC₅₀ value was calculated for each inhibitor.

The IC₅₀ values of the selected peptide aldehyde inhibitors reflected a difference in the affinity towards StmPr1 protease. The order from stronger to weaker inhibition is found as follows: chymostatin > LLF > LLL > calpain inhibitor II > leupeptin. Since the strongest inhibitors, chymostatin and LLF, both have Phe at the P1 position, the observation mentioned above that subtilisins prefer an aromatic amino acid in P1, holds also for the StmPr1 protease. The lower IC₅₀ of chymostatin as opposed to LLF may be explained by a stronger interaction of a tetrapeptide versus the tripeptide LLF.

In order to study the mode of binding, and in more details on the molecular level, protein crystallography is the appropriate technique.

3.4 Co-crystallization of StmPr1 with different peptide aldehyde inhibitors

To study of enzyme-inhibitor interactions, complexes have to be formed. There are several ways to address this point either using the co-crystallization or the crystal soaking technique. In co-crystallization, the inhibitor is combined with the protein prior to crystallization and the complex is crystallized. Whereas in the soaking technique, the preformed crystals of the native protein were incubated with the potential ligands – expecting these small molecules to penetrate into the crystal lattice to move through the crystal solvent channels until finding their binding sites – thus using the crystal as an analytical tool. If this technique works, it is advantageous because it is fast, convenient, and reproducible. The success of the ligand soaking experiment relies on the existence of wide solvent channels running through the lattice providing access to all protein molecules and all active sites. The soaking experiments took some time to determine the optimal conditions because the solubility of the ligand is considered as an obstacle, as those peptide aldehyde inhibitors are more soluble in dimethyl sulfoxide (DMSO), and sometimes precipitate in the aqueous solution, and on the other hand protein crystals tend to crack or dissolve in DMSO containing solutions. Another important factor is the applied ligand concentration. Several concentrations have been tested to reach the optimum concentration which avoids the ligand precipitation and the cracking of the protein crystals.

After optimizing the soaking conditions, StmPr1 protease crystals were soaked in all inhibitor solutions. The soaking experiments were successful only with leupeptin

(the weakest inhibitor) and chymostatin (the strongest inhibitor). Consequently, co-crystallization was applied with the other three inhibitors, but no crystals were obtained.

3.4.1 Structure of the leupeptin StmPr1 complex

Leupeptin, also known as N-acetyl-L-leucyl-L-leucyl-L-argininal, is a naturally occurring organic compound produced by actinomycetes, which inhibits serine proteinases. Leupeptin is a reversible inhibitor. It acts as a transition state analogue. Transition-state analogues are good inhibitors because they are bound to the enzyme more tightly than the substrates [97].

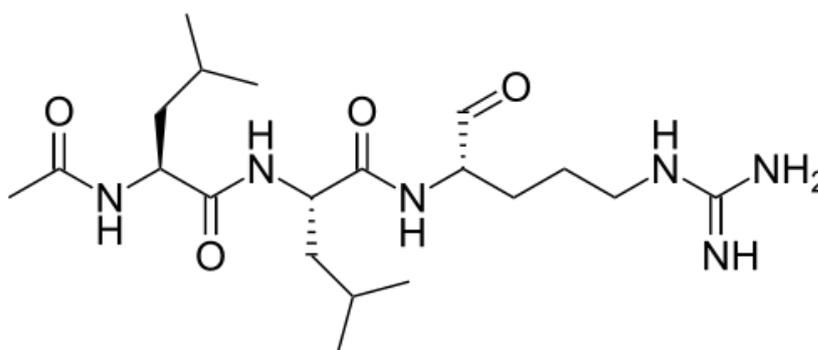


Figure 30: Chemical structure of leupeptin.

The diffraction data of the complex crystals were collected at the Consortium-Beam Line X13, DESY Hamburg. The detector type used was a Mar CCD 165 mm and the wavelength was 0.8123 Å. Data of up to 2.0 Å resolution were collected. The unit cell was orthorhombic and the space group was identified to be C222₁, as the native structure. The model was refined at a resolution of 2.0 Å and an R-factor of 17.3% and R_{free} value of 21.0%. The data collection parameters and refinement statistics are shown in Table 8.

Table 9: Data collection parameters and refinement statistics of the leupeptin StmPr1 complex

Parameters	
Space group	C222 ₁
a (Å)	60.78
b (Å)	86.41
c (Å)	132.35
VM (Å ³ / Da)	2.12
Solvent content (%)	42.%
Completeness of data (%)	98.2 (88.5)
No. of total reflections	102137 (12424)
No. of unique reflections	21042 (2742)
Average I/sigma intensity	8.9 (2.7)
Resolution (Å)	30.0-2.08 (2.19-2.08)
Redundancy	4.9 (4.5)
*Rmerge (%)	0.164 (0.569)
No. of reflections used in refinement	19915
Rcrystal (%)	17.3
No. of reflections used in Rfree	996
Rfree (%)	21.0
Protein atoms	2734
Amino Acids:	356
Calcium Ions	1
Sulphate Ions	5
Glycerol molecules	2
Water molecules	242
Root mean square deviation	
Bonds (Å)	0.0221
Bond angles (°)	0.940
Residues in regions of the Ramachandran plot (%)	
Most favored	90.5
Allowed	9.5
Generally allowed	0
Disallowed	0

3.4.1.1 Structure analysis

The Leupeptin-StmPr1 proteinase complex structure has been analyzed and the main interactions between the StmPr1 proteinase and the inhibitor have been identified. A covalent hemiacetal bond is formed between the active-site Ser-289 and the C-terminal aldehyde moiety of leupeptin.

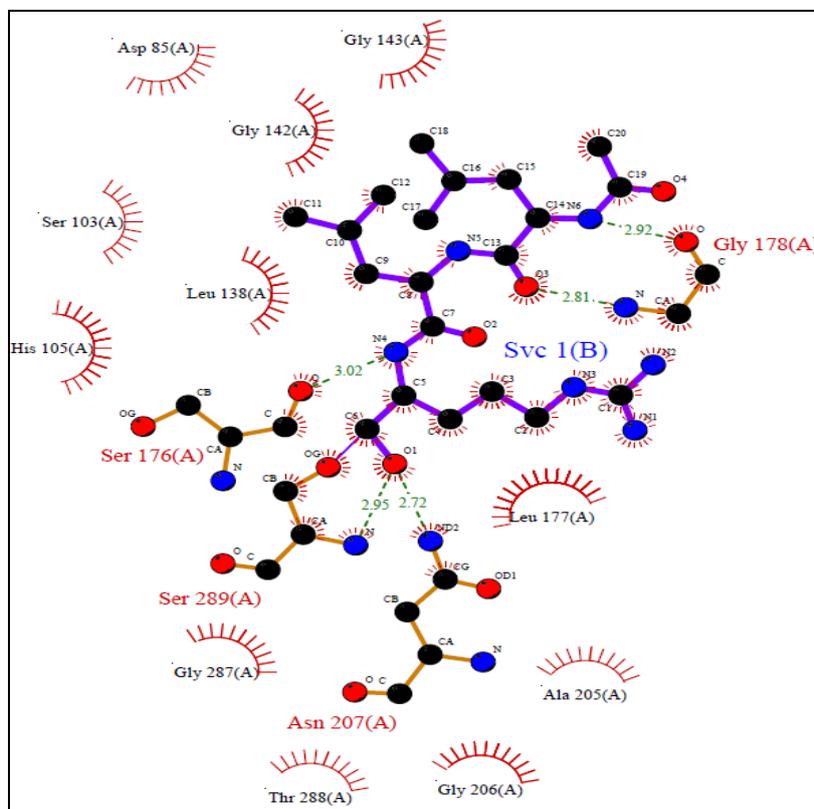


Figure 31: Ligplot showing leupeptin in the binding site of StmPr1.

The interactions between leupeptin and StmPr1 active site residues (Ser-176, Gly-178, Asn-207, and Ser-289) are summarized in Table 9. Several specific contacts are made to the peptide backbone of leupeptin.

Table 10: Interactions of the leupeptin StmPr1 complex

Nr.	Residue name	Atom name		Atom name	Distance (Å)	Bond type
1.	Ser-289	OG	↔	C6	1.42	Hemiacetal
2.	Ser-289	N	↔	O1	2.95	H-bond
3.	Asn-207	ND2	↔	O1	2.72	H-bond
4.	Ser-176	O	↔	N4	3.02	H-bond
5.	Gly-178	N	↔	O3	2.81	H-bond
6.	Gly-178	O	↔	N6	2.92	H-bond

It was observed that there is no significant movement of the active site residues of StmPr1 upon complexation with the inhibitor. The C-terminal arginine residue of the inhibitor is bound to the specificity pocket of the enzyme, which indicates that the S1 subsite can accept the positively charged amino acids. There are hydrophobic interactions of the penultimate leucine to the small S2 pocket. Finally, H-bonding between the main chain of the inhibitor and a part of the main chain of the enzyme stabilizes the interaction.

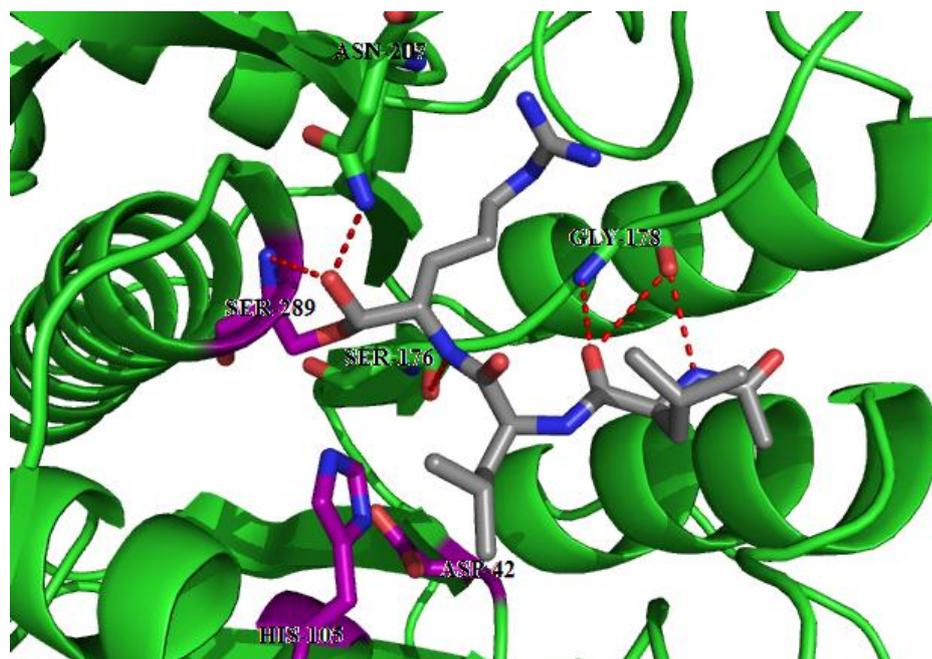


Figure 32: Active site of StmPr1. Catalytic triad residues (Asp-42, His-105, and Ser-289) are colored and labeled in purple. Main chains and side chains of StmPr1 residues forming polar contacts with Leupeptin (red dashed lines) are shown. Inhibitor is shown in gray.

The inhibitor is almost completely buried in the active site cleft, as shown in Figure 33.

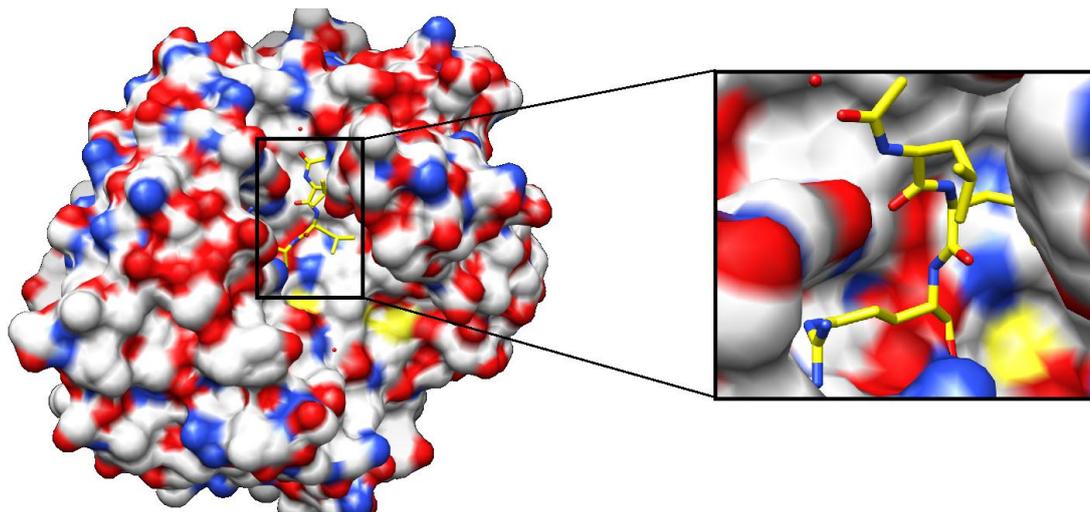


Figure 33: Surface view of the StmPr1-leupeptin complex, colored according to electrostatic potential (blue, positive; red, negative; white, neutral); leupeptin is shown in yellow.

The initial electron density maps which are calculated for leupeptin using 2.0 Å data (Figure 34) showed that the three amino acid residues (leucine, leucine, and arginine) forming the tripeptide backbone of the inhibitor were clearly located in the enzyme active site, and the electron density of the P1 to P3 side chains matched to the corresponding amino acid sequence. The maps indicated that the P1 residues lie at the bottom of the active site. The electron density bridging to the essential Ser289 forming the hemiacetal bond (Figure 34), and the remaining peptide extends outwards, making contacts with the active site cavity.

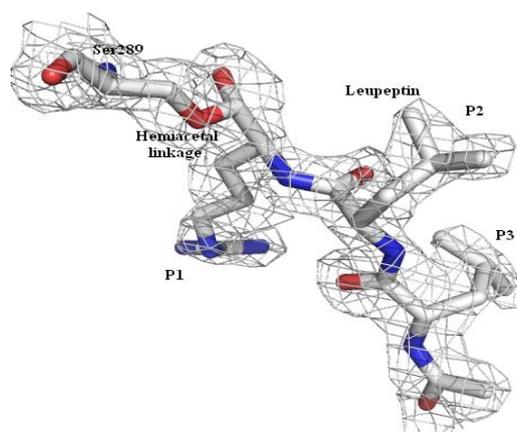
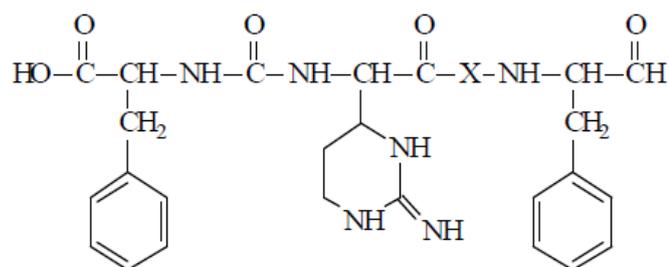


Figure 34: Fo-Fc map contoured at 1.0 σ showing the electron density for the catalytic Ser-289 covalently bound to leupeptin through a hemiacetal bond.

After having demonstrated this structure with a tripeptide inhibitor bound to StmPr1, the complex with the tetrapeptide inhibitor chymostatin – the strongest inhibitor for StmPr1 tested so far – was investigated.

3.4.2 Structure of the chymostatin StmPr1 complex

Chymostatin is a tetrapeptide analogue from culture filtrates of *Streptomyces hygroscopicus* and *Streptomyces lavendulae*. It is a powerful inhibitor of a number of serine and cysteine proteases. The chemical structure determined for chymostatin [98] is shown in Figure 35. It contains a C-terminal aldehyde group which is essential for inhibitory activity via formation of a hemiacetal or hemithioacetal adduct with the nucleophilic hydroxyl or thiol group of serine and cysteine proteinases, respectively. When compared to normal peptide inhibitors, chymostatin has some unusual structural features: A capreomycidine residue occupies the P3 position and an ureido linkage (-NH-CO-NH-) which connects the phenylalanine to the three subsequent residues (P3 through P1) and reverses the chain direction.



Chymostatin A: X=L-Leu
Chymostatin B: X=L-Val
Chymostatin C: X=L-Ile

Figure 35: The chemical structure of chymostatin.

The diffraction data were collected at the Consortium-Beam Line X13, DESY Hamburg. Data with a maximum resolution of 2.0 Å were collected. The data collection and refinement parameters are summarized in Table 10.

Table 11: Data collection parameters and refinement statistics of the chymostatin StmPr1 complex

Parameters	
Space group	C2221
a (Å)	60.43
b (Å)	86.51
c (Å)	132.20
V _M (Å ³ /Da)	2.17
Solvent content (%)	43.34%
Completeness of data (%)	98.8(92.2)
No. of total reflections	144330(17280)
No. of unique reflections	24332(3272)
Average I/sigma intensity	11.3(3.5)
Resolution (Å)	30.0-1.98(2.09-1.98)
Redundancy	5.9(5.3)
*Rmerge (%)	0.129 (0.457)
No. of reflections used in refinement	22646
Rcrystal (%)	15.6
No. of reflections used in Rfree	1132
Rfree (%)	19.0
Protein atoms	2734
Amino Acids:	356
Calcium Ions	1
Sulphate Ions	5
Glycerol molecules	2
Water molecules	252
Root mean square deviation	
Bonds (Å)	0.023
Bond angles (o)	1.73
Residues in regions of the Ramachandran plot (%)	
Most favored	90.5
Allowed	9.5
Generally allowed	0
Disallowed	0

3.4.2.1 Structure analysis

The chymostatin- StmPr1 proteinase complex structure has been analyzed by X-ray crystallography and the main interactions between the proteinase and the inhibitor have been identified. A covalent hemiacetal bond is formed between the active-site Ser- 289 and the C-terminal aldehyde moiety, which contributes to the binding by its resemblance to the tetrahedral intermediate in proteolytic catalysis [99].

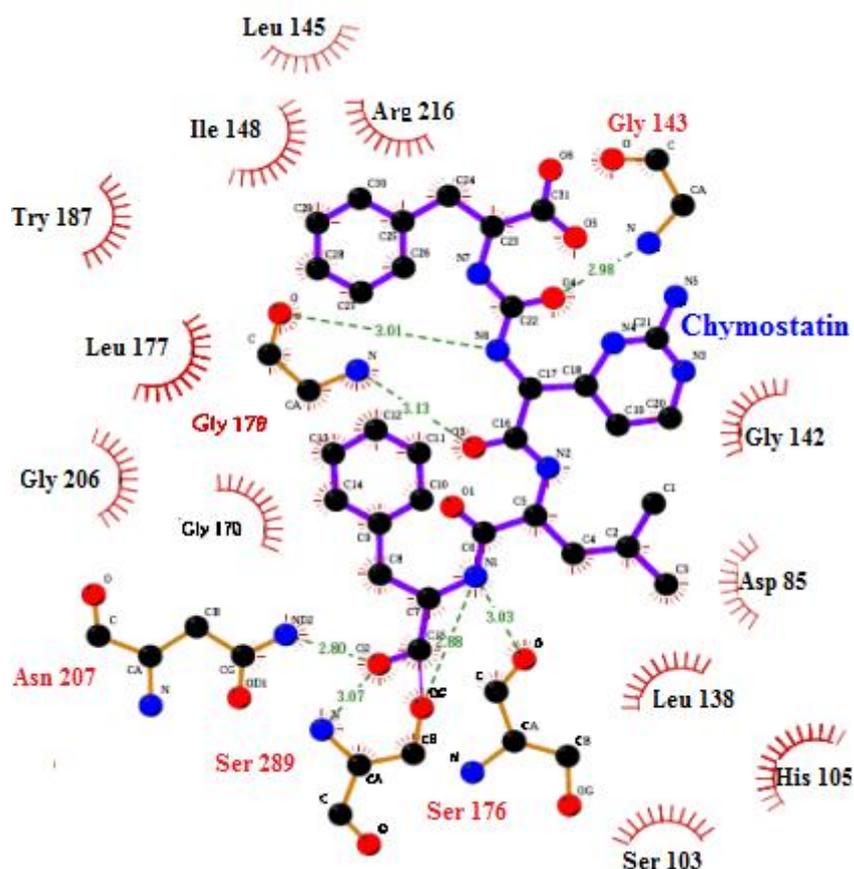


Figure 36: Ligplot of the chymostatin StmPr1 complex showing the hydrogen bond interactions.

The C-terminal phenylalanine residue of the inhibitor is bound to the primary specificity pocket of the enzyme. Furthermore, a hydrophobic binding of the penultimate leucine to the small P2 pocket is observed. The H-bonds between the main chain of the inhibitor and a part of the main chain of the enzyme stabilize the interaction. The number of hydrogen bonds formed with the active site residues is higher compared with the number of hydrogen bonds in case of leupeptin. This explains why chymostatin showed a significant stronger inhibition effect.

Table 12: The Chymostatin StmPr1 complex interactions

Nr.	Residue name	Atom name		Atom name	Distance (Å)	Bond type
1.	Ser-289	OG	↔	C15	1.42	Hemiacetal
2.	Ser-289	OG	↔	N1	2.88	H-bond
3.	Ser-289	N	↔	O2	3.07	H-bond
4.	Asn-207	ND2	↔	O2	2.80	H-bond
5.	Ser-176	O	↔	N1	3.03	H-bond
6.	Gly-178	N	↔	O3	3.13	H-bond
7.	Gly-178	O	↔	N6	3.01	H-bond
8.	Gly-143	N	↔	O4	2.98	H-bond

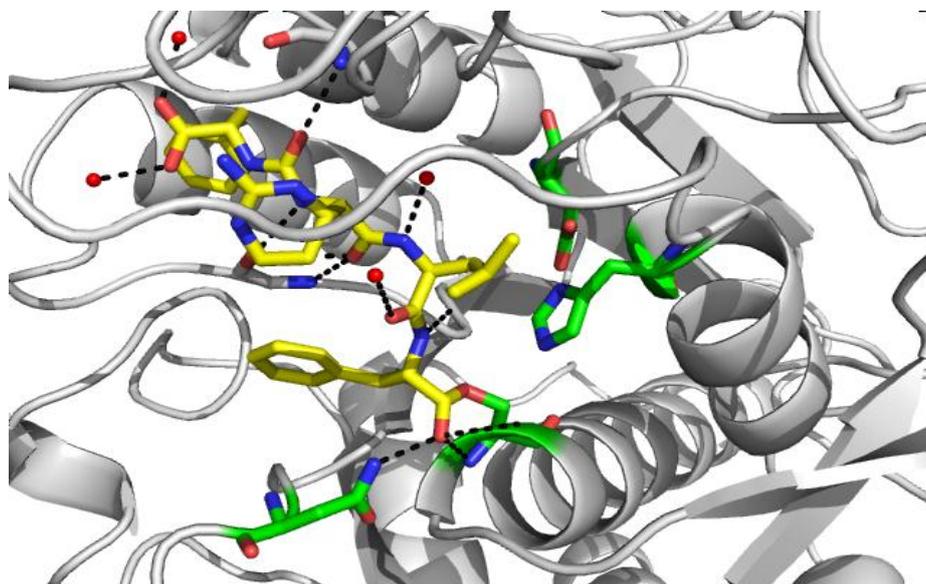


Figure 37: The chymostatin-StmPr1 complex structure.

Catalytic triad residues (Asp-42, His-105, and Ser-289) are colored in green. Main chain and side chain of StmPr1 residues forming polar contacts with chymostatin (black dashed lines) are shown. Inhibitor is shown in yellow. The structure was oriented to show the inhibitor clearly.

The Chymostatin-StmPr1 model was refined to a R-factor of 16%, with good geometry at 2.0 Å resolution. All four residues of chymostatin are well-localized in the putative S1 through S4 sites. The S1 site of StmPr1 forms a hydrophobic binding surface for the P1 phenylalanyl side-chain of chymostatin. The hydrogen bonds to the peptide backbone are made primarily with the side-

chains of the enzyme. The hemiacetal linkage to the essential Ser-289 are of a single stereoisomer with tetrahedral geometry and with an oxygen atom occupying the “oxyanion hole” region of the enzyme. The oxygen atom accepts three hydrogen bonds; one from the polypeptide backbone, one from a water molecule and one from the positively charged amino group of Asn 207. Thus, the binding forms an excellent approximation to the oxyanion intermediate formed during peptide hydrolysis. Water molecules play an important role in stabilizing the inhibitor in the active site cleft. It was found that four water molecules hydrogen bond to chymostatin back bone (Figure 38).

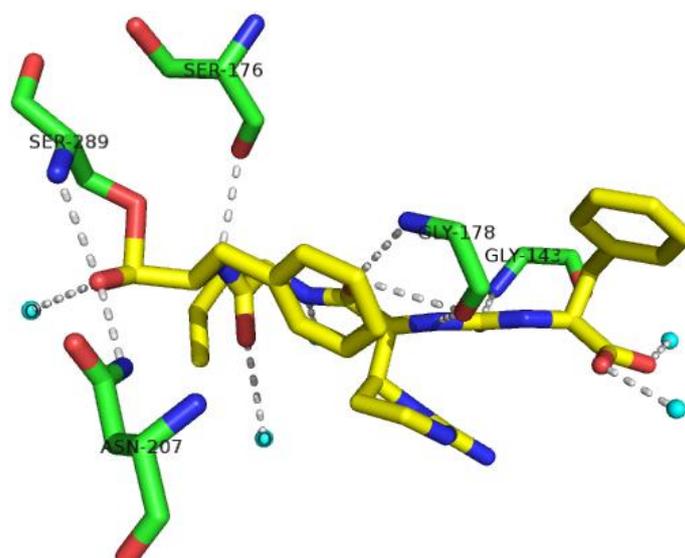


Figure 38: The interactions of the water molecules in the StmPr1 active site.

Catalytic triad residues, main chains and side chains residues of StmPr1 are colored in green and labeled in black. Polar contacts with chymostatin are shown in white dashed lines. Inhibitor is shown in yellow.

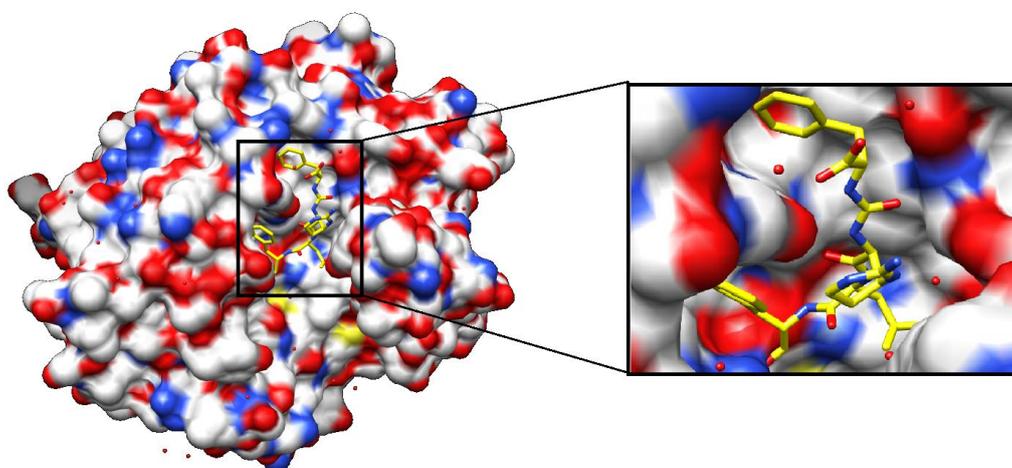


Figure 39: Surface view of StmPr1 showing chymostatin (in sticks) in the binding site of StmPr1.

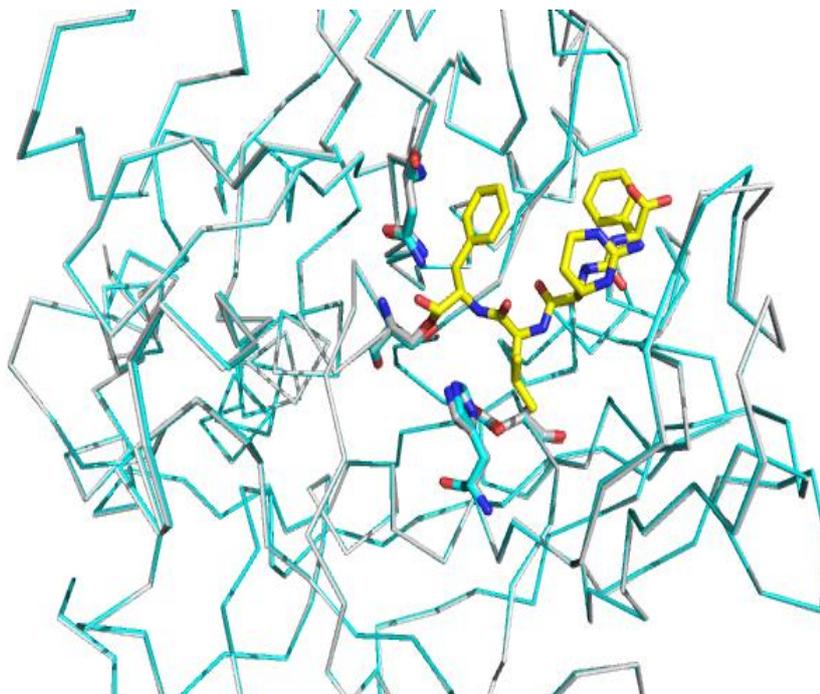


Figure 40: Superposition of the C α -atoms of native StmPr1 structure (in cyan) and StmP1-chymostatin Complex structure (in gray).

The StmPr1 chymostatin complex structure has been superimposed to the native StmPr1 structure (Figure 40). This indicates that there were no significant movements of the active site residues of StmPr1 upon complexation with the inhibitor, because the chymostatin molecule is small to accommodate itself in the relatively big active site of the enzyme.

Initial electron density maps, calculated for chymostatin, showed that the peptide inhibitor was clearly located in the enzyme active site. The electron density for all four residues of chymostatin was consistent with the peptide sequence, including a cyclized arginine side-chain at P3 and the connecting ureido bond to P4. The electron density at P4 is weaker than for other parts of the molecule, which may indicate that this residue has a larger degree of thermal mobility in accordance with its location at the free end of the peptide and surface of the protein. The maps indicated that the P1 residue lies at the bottom of the active site, with a continuous electron density bridging to the essential Ser289. The remainder of each peptide extends outward, making contacts with one wall of the active site cavity. At the P2 position, the electron density matched leucine in shape and size, indicating that chymostatin A is the major form of the inhibitor bound to the enzyme.

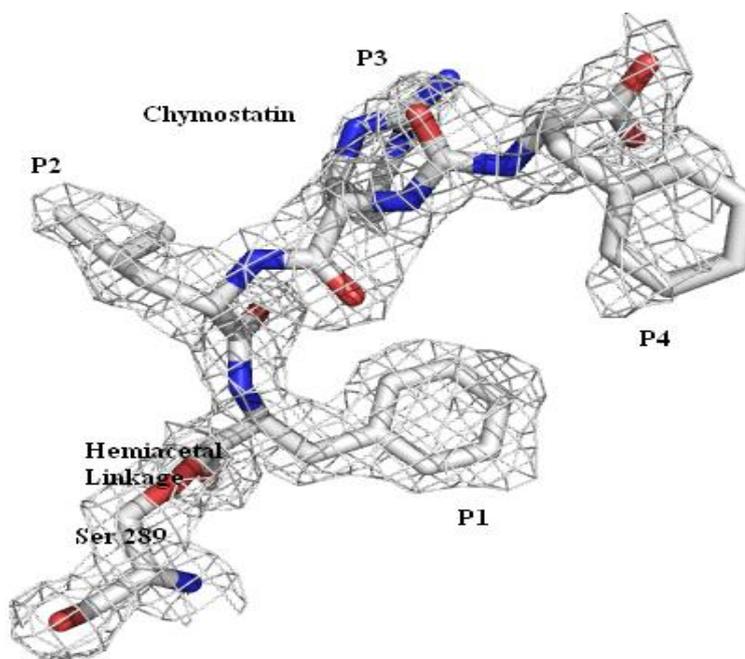


Figure 41: Fo-Fc map contoured at 1.0 σ , showing the electron density for the catalytic Ser-289 covalently bound to the chymostatin molecule through a hemiacetal bond.

The structures of the complexes formed between either chymostatin or leupeptin and StmPr1 protease are reported here for the first time. It was observed that there were no significant movements of the active site residues of StmPr1 upon complexation with both inhibitors. This is not surprising because they are relative small molecules and the active site, as already observed, is big enough to accommodate the inhibitor without inducing any conformational changes. The S1 subsite of StmPr1 protease accepted the positively charged and the aromatic amino acids (arginine and phenylalanine) respectively.

It is observed that the S3 subsite of StmPr1 is located at the surface of the protein, and this means that it can accept a variety of different residues e.g. big aromatic side chains at the P3 position of the inhibitor. In case of chymostatin, the aromatic group of Phe in P4 is highly positioned at S4 subsite. The analysis of the complex structure with chymostatin showed that it formed more hydrogen bonds than leupeptin, conforming what have been shown in the kinetic experiments (Figure 29) that chymostatin is a potent inhibitor of StmPr1 more than leupeptin.

On searching the protein data bank for complex structures of chymostatin, 14 crystal structures were found. 13 complex structures came from bacterial

origin and only one reported structure belongs to an eukaryotic protease (the wheat carboxy-peptidase II chymostatin complex structure). Comparing the two structures revealed that chymostatin showed nearly the same mode of binding especially the hemiacetal linkage. This supports the idea that chymostatin is non-specific. Although chymostatin revealed to be a good inhibitor for StmPr1, its broad specificity is considered to be a limitation for its use as a drug.

3.5 A new inhibitor identified in a snake venom by crystallography

Snake venoms are a rich source of several bioactive compounds that possess therapeutic potentials. The high target specificity, small size, structural stability and relative ease of chemical synthesis make the snake venom peptides promising alternatives to the contemporary drugs.

A project of the Hamburg working group deals with the characterization of *Agkistrodon bilineatus* snake venom peptides. Therefore, it was intriguing to test such a peptide fraction for inhibitory activity towards StmPr1. Indeed, a preparation containing numerous low molecular weight peptides showed inhibition (data not shown; only one preliminary experiment). Since only a very limited amount of this crude fraction was available, this was directly - without further purification – used for co-crystallization experiments. Purified StmPr1 was mixed with the venom fraction and subjected to the crystallization conditions, as applied for the native protease. Interestingly, crystals grew in spite of the heterogeneity of the solution. Next the question was, whether components of the peptide mixture could be found in the crystal structure and, if so, whether a particular peptide specifically is bound in the active site.

Table 13: Statistics for *Agkistrodon bilineatus* snake venom peptide StmPr1 complex

Parameters	
Space group	C222 ₁
a (Å)	60.56
b (Å)	86.78
c (Å)	132.11
V _M (Å ³ /Da)	2.17
Solvent content (%)	43.30%
Completeness of data (%)	97.3 (85.1)
No. of total reflections	132119(15673)
No. of unique reflections	32679(4123)
Average I/sigma intensity	14.8(5.5)
Resolution (Å)	26.28-1.78(1.88-1.78)
Redundancy	4.0(3.8)
*Rmerge (%)	6.1(21.3)
No. of reflections used in refinement	30966
R _{crystal} (%)	14.9
No. of reflections used in R _{free}	1548
R _{free} (%)	18.77
Protein atoms	2734
Amino Acids:	356
Calcium Ions	1
Sulphate Ions	5
Glycerol molecules	2
Water molecules	344
Root mean square deviation	
Bonds (Å)	0.028
Bond angles (°)	2.06
Residues in regions of the Ramachandran plot (%)	
Most favored	90.0
Allowed	10.0
Generally allowed	0
Disallowed	0

The diffraction data were collected at the Consortium-Beam Line X13, DESY Hamburg. The data collection and refinement statistics are summarised in Table 12.

3.5.1 Structure analysis

While interpreting the data it became obvious that one particular peptide out of the heterogeneous *Agkistrodon bilineatus* venom preparation was found in the active site of StmPr1 (Figure 42): the tetrapeptide Ala-Ser-Pro-Ser (now named ASPS).

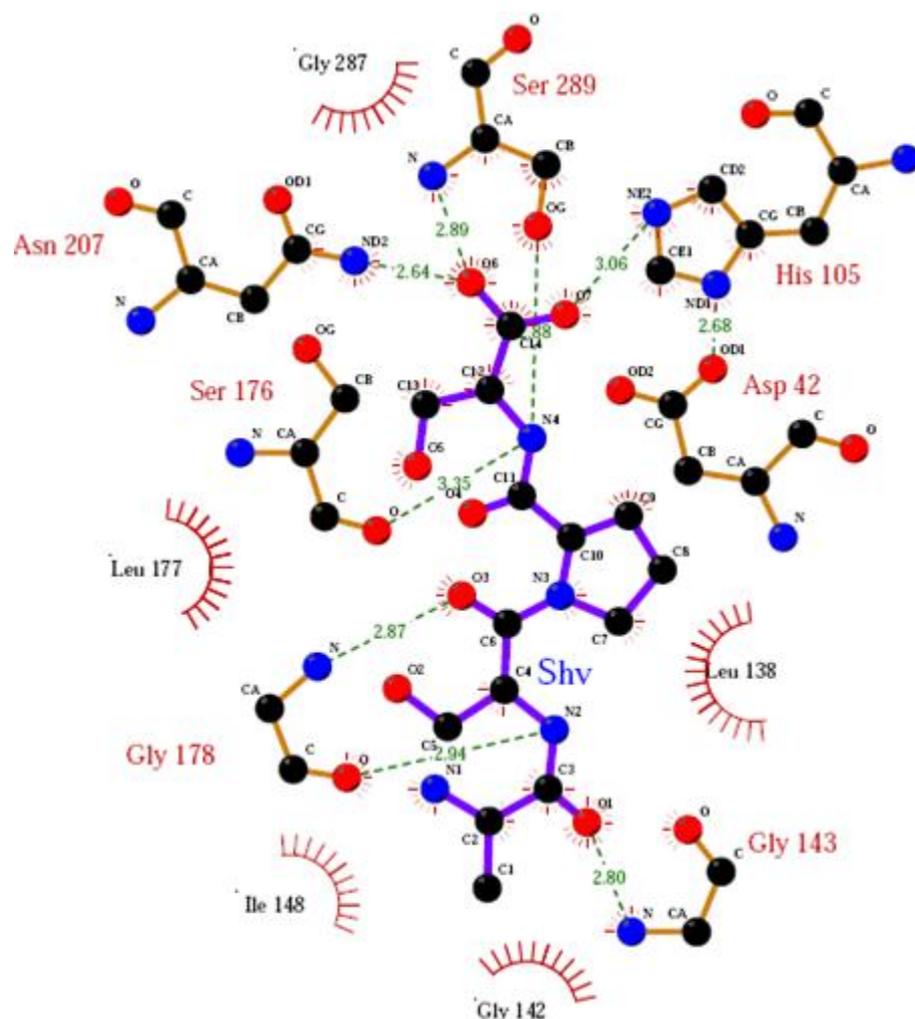


Figure 42: Ligplot of the StmPr1-*Agkistrodon bilineatus* venom peptide inhibitor complex.

It is stabilized by forming 8 hydrogen bonds between the six residues (His-105, Gly-143, Ser-176, Gly-178, Asn-207, and Ser-289) of StmPr1 and the main chain atoms of the inhibitor (Figure 42), in addition to three water molecules which form

Results and Discussion

three hydrogen bonds with the atoms N1, O4, and O7 of the ASPS peptide. In total 11 hydrogen bonds stabilize the inhibitor inside the active site of StmPr1. The hydrogen bonds to the peptide backbone are made primarily with side-chains on the enzyme. The C-terminal serine residue of the inhibitor is bound to the primary specificity pocket of the enzyme through four hydrogen bonds. The StmPr1-ASPS complex model was refined to a R-factor of 14.9%, with good geometry at 1.78 Å resolution. All four residues of the peptide inhibitor are well-localized in the putative S1 through S4 sites.

Table 14: The *Agkistrodon bilineatus* venom peptide inhibitor-StmPr1 complex interactions.

Nr.	Residue name	Atom name		Atom name	Distance (Å)	Bond type
1.	Ser-289	OG	↔	N4	2.88	H-bond
2.	Ser-289	N	↔	O6	2.89	H-bond
3.	Asn-207	ND2	↔	O6	2.64	H-bond
4.	His-105	NE2	↔	O7	3.06	H-bond
5.	Ser-176	O	↔	N4	3.35	H-bond
6.	Gly-178	O	↔	N2	2.94	H-bond
7.	Gly-178	N	↔	O3	2.87	H-bond
8.	Gly-143	N	↔	O1	2.80	H-bond

The oxygen atom occupying the “oxyanion hole” region of the enzyme accepts two hydrogen bonds; one from the polypeptide backbone and one from the positively charged amino group of Asn 207. It was noticed that there were no significant movements of the active site residues of StmPr1 upon complexation with inhibitors. The *Agkistrodon bilineatus* venom peptide forms a higher number of hydrogen bonds with StmPr1 active site residues than what was seen in the structures with leupeptin and chymostatin. Therefore, one may assume that it acts as a strong inhibitor in enzyme assays.

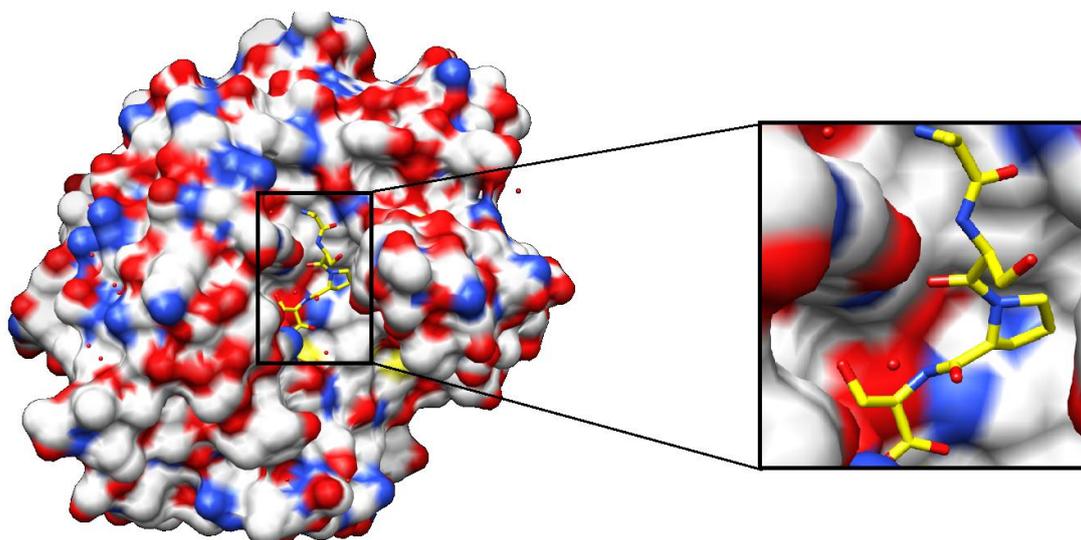


Figure 43: Surface view of StmPr1 showing the peptide inhibitor (in sticks) in the binding site of StmPr1.

This is in accordance with one preliminary experiment where the snake venom peptide fraction showed stronger inhibition on StmPr1 than on chymotrypsin (which is inhibited by chymostatin in the nano-molar range). Thus, the ASPS peptide seems to exhibit some specificity for StmPr1 (the structure of the complex supports this notion). Unfortunately, at the time of this work there was not sufficient material available to perform additional biochemical experiments necessary for further characterization (affinity, specificity, etc.). The crude snake venom fraction contains many compounds (anything < 10 kDa), however, crystals grew and only one peptide was found in the active site which is remarkable from a technical point of view. In this case the crystal served as a purification tool, and crystallography allowed to identify a new inhibitory compound.

Although the StmPr1 inhibitors described so far must be regarded as “non-druggable” due to lack of affinity and/or specificity. The crystal structure of the native StmPr1 can help to identify new molecules fitting into the active site. This can be done using a computational technique called virtual ligand screening (VLS). It involves the rapid *in silico* assessment of large libraries of chemical structures in order to identify compounds which are most likely to bind to StmPr1. This was performed in cooperation with European Screening Port at Hamburg, Germany. When this strategy was applied on StmPr1, some “hit” compounds were found. Unfortunately, when tested in enzyme activity assays, none of them displayed StmPr1 inhibition. One explanation could be that using the native protein structure is always tricky for doing

the docking experiments of the small molecules; there is no reference structure of ligands which interact with the protease to be used as an example of the binding. There is a running experiment which makes use of the elucidated complex structures. Once the coordinates of the complexes are known, a model can be built by exploiting the collective information contained in such complex structures. This is known as pharmacophore model, which can be simply defined as the portion of a ligand molecule which binds to the enzyme active site. Then a candidate ligand can be compared to the pharmacophore model to determine whether it is compatible with it and therefore likely to bind.

3.6 Identification of StmPr1 inhibitors by High-Throughput Screening:

High-throughput screening (HTS) is considered as one possible method to obtain lead compounds, in which thousands of compounds are assayed for StmPr1 inhibition in small scale experiments. Indeed, there are numerous compound libraries containing a collection of stored chemicals used in high-throughput screening.

The ChemBioNet libraries which contain more than thirty thousand compounds and in addition to the ENZO library which contains 640 drugs approved by Food and Drug Administration (FDA) were tested for StmPr1 inhibition. For that purpose, the protein was produced in bulk quantities in *E. coli* and an enzyme assay was established suitable for the robot-based screening procedure. In order to perform the high-throughput screening, the enzyme assay was optimized in order to scale down the assay volume. The next step is the assay transfer in which the assay is performed on the 384-well plates to optimize the incubation time, the substrate concentration, the speed of the reaction, etc. Next the primary screening was performed at which the enzyme was incubated with the compound libraries prior to substrate addition. This step resulted in 85 'hit' compounds. A secondary screening was applied to validate the hits, and at the same experiment the IC_{50} is calculated. Several potential inhibitors were identified with IC_{50} values $< 10 \mu\text{M}$. One of the positive hits was bortezomib which showed significant inhibition with IC_{50} of $0.3 \mu\text{M}$, as shown in Figure 44.

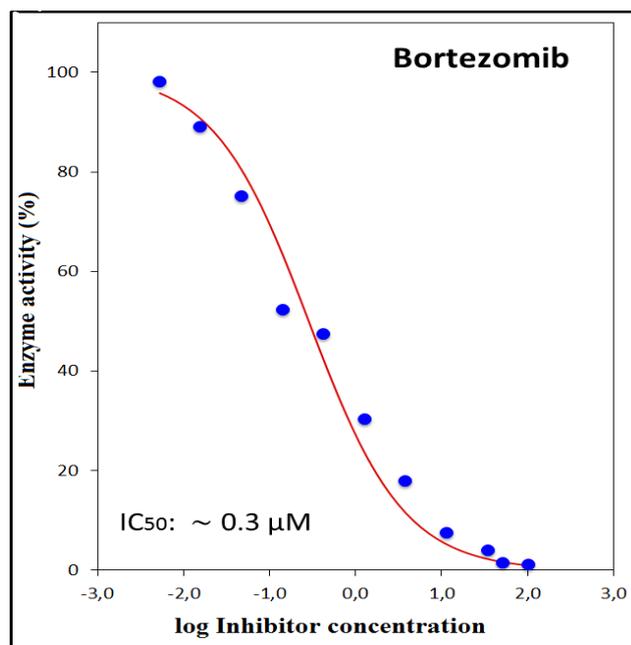


Figure 44: The Inhibitory effect of bortezomib on StmPr1.

The dipeptide boronic acid bortezomib (Velcade™) is a proteasome inhibitor which is currently used in the treatment of multiple myeloma. The chemical structure determined for bortezomib is shown in Figure 45. It is a tripeptide and can be written as Pyz-Phe-boroLeu, which stands for pyrazinoic acid, phenylalanine and Leucine with boronic acid instead of a carboxylic acid.

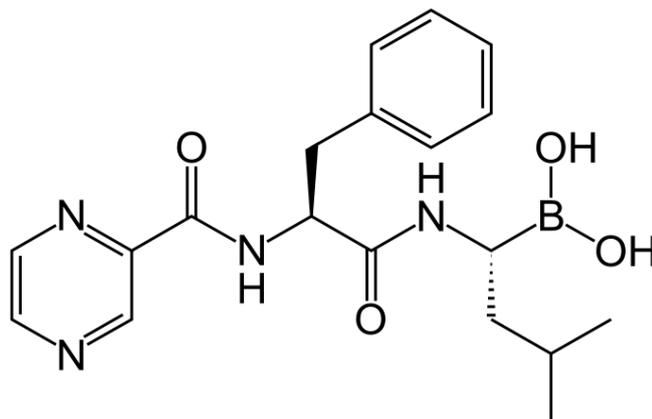


Figure 45: The chemical structure of bortezomib.

Bortezomib belongs to the peptide boronates which were developed from peptide aldehyde inhibitors [100]. Substitution of boronic acid for the aldehyde moiety produced highly potent inhibitors. Since bortezomib so far has not been used as inhibitor for bacterial proteases, it was important to gain more information about the mode of interaction of this drug with the StmPr1 protease. Therefore, enzyme assays

Results and Discussion

were performed using different concentration of the substrate in the presence or absence of the inhibitor, and the velocities of reactions were determined as well. From the saturation curves obtained, the K_m (Michaelis constant) and V_{max} (maximum velocity) were calculated by non-linear regression and/or by a double reciprocal plot (Lineweaver Burke). As shown in Figure 46, the K_m value was not significantly changed upon addition of bortezomib, while V_{max} was markedly decreased. Assuming Michaelis-Menten kinetics, one might suggest for bortezomib a non-competitive mechanism where the inhibitor appears to reduce the number of non-occupied enzyme molecules.

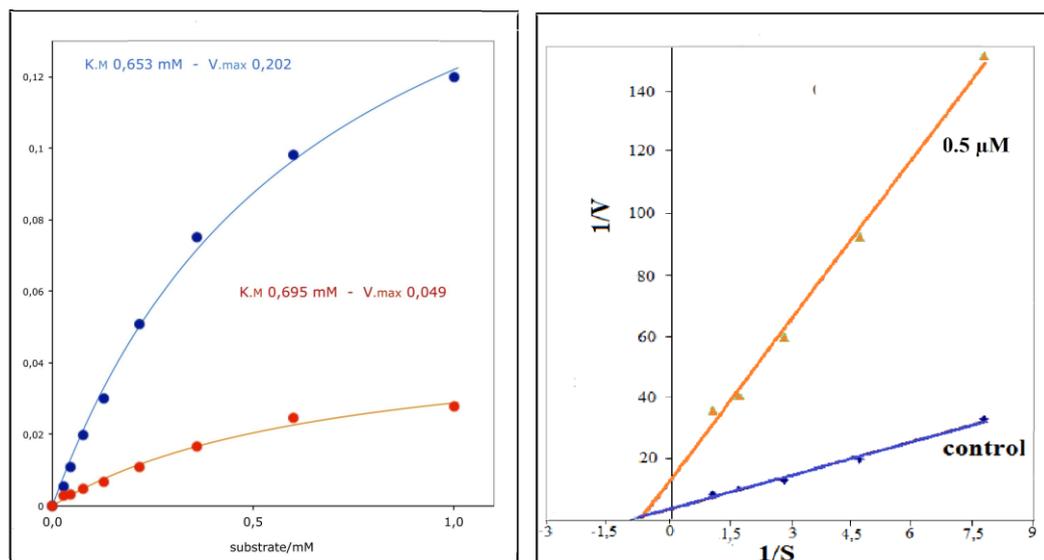


Figure 46: Inhibition of StmPr1 with bortezomib and chymostatin.

Bortezomib was preincubated with the StmPr1 for 20 minutes prior to addition of the chromogenic substrate Suc-A-A-P-F-PNA at different concentrations, 0.01-1 mM.

However, it seems to be that bortezomib does not follow Michaelis-Menten kinetics, since the inhibitory effect was seen only after preincubation of the enzyme in the absence of the substrate. When the inhibitor and the substrate were mixed and the reaction was started by addition of the enzyme, inhibition was hardly detectable. Essentially, similar observation was reported using chymostatin (data not shown). Possibly, longer time is required for the inhibitor to bind to the active site of the protease.

3.6.1 Effect of StmPr1 and bortezomib in cell culture

In order to observe the overall effects of bortezomib on living cells, a cell culture experiment was performed. Human lung epithelial tumor cells were selected because it is reported that *S. maltophilia* has originally been isolated from the lung of infected patients and has contributions in inflammations of the respiratory tract. Several questions will be addressed:

- Is StmPr1 responsible for cell damage effect?
- Can bortezomib block the damaging effect of recombinant StmPr1?
- And if so, does the drug block also the natural enzyme secreted by *S. maltophilia*?
- Is bortezomib itself toxic for the cells?

The first experiment was to demonstrate the effect of pure recombinant StmPr1 on the cultures of human lung epithelial tumor cells. The protease was added to confluent cells, and after one hour morphological changes were investigated (Figure 47B). The cell layer partially condensed forming cell-free areas, and finally detached from the culture plate. Importantly, this effect of StmPr1 could be prevented (Figure 47C) when the cells were incubated with StmPr1 together with bortezomib (40-fold molar excess).

The same bortezomib concentration was tested on control cells without protease to test whether bortezomib itself is toxic; figure 47D shows that the drug under these conditions did not cause morphological changes.

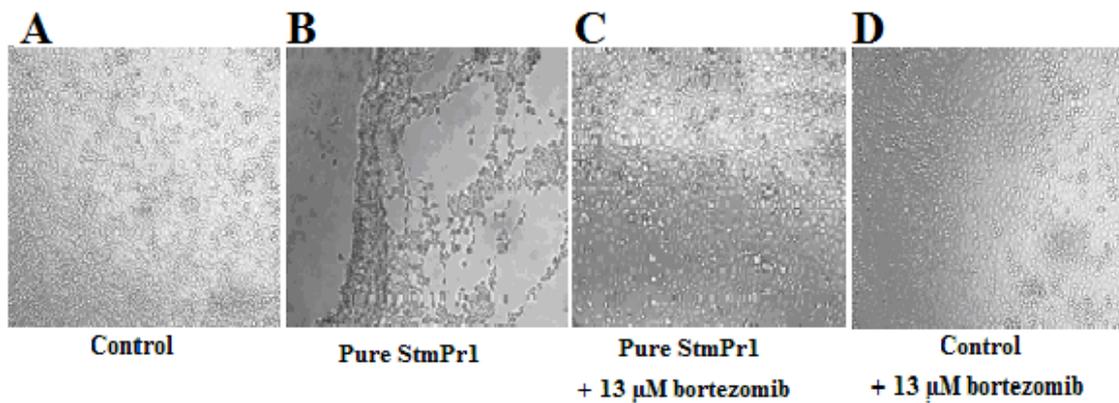


Figure 47: Effect of bortezomib on cultures of human lung cells.

A: Confluent cultures of human lung tumor cells were incubated **B:** with 12 µg/ml purified StmPr1 protease, dialyzed against Dulbecco's modified Eagle's medium DMEM were incubated for 1 hour at 37 °C, **C:** in presence of 13 µM bortezomib **D:** Confluent cultures of human lung cells were incubated for 1 hour at 37 °C 13 µM bortezomib in (DMEM).

So far, the experiments used the recombinant 37 kDa truncated product of the StmPr1 gene processed by *E coli*. More important issue is to demonstrate the effect of the “natural” protease of *S. maltophilia* in the same type of experiment, because this would be closer to the in-vivo situation. Therefore, a bacterial culture was raised from an infected with *S. maltophilia*. The cell-free supernatant of this culture was concentrated and applied to a buffer exchange column equilibrated with cell culture medium. The preparation was calibrated by enzyme assay, and the same amount of protease activity used in the case of recombinant StmPr1 was added to the cells. As demonstrated in Figure 48B, a similar destructive effect which was observed with the recombinant protease, was also reported in case of the natural protease of *S. maltophilia*. Most importantly, the cell damaging effect could be diminished with bortezomib (Fig. 48C).

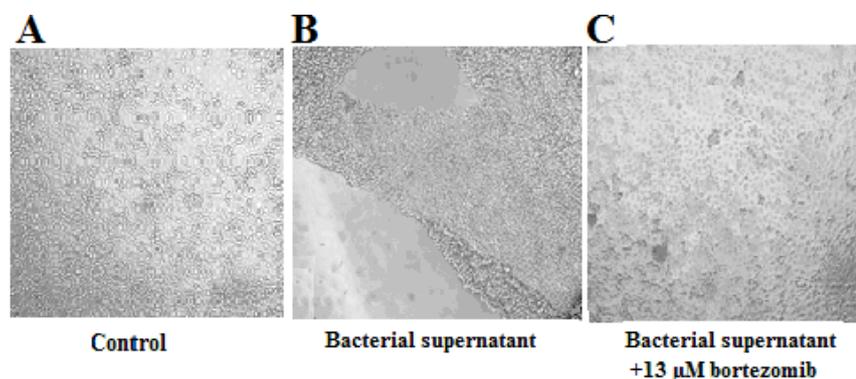


Figure 48: Effect of bortezomib on cultures of human lung cells pre-treated with the *Stenotrophomonas maltophilia* culture medium.

Confluent cultures of human lung cells were incubated for 1 hour at 37 °C with bacterial supernatant, dialyzed against DMEM, in the absence and presence of 15 µM chymostatin and 13 µM bortezomib.

Altogether, these experiments demonstrated that the secretions of *S. maltophilia* are able to destroy living cells, and that the StmPr1 protease is the major factor responsible for this effect. Therefore, it is likely that the tissue lesions seen in infected patients are consequences of StmPr1 action. StmPr1 inhibitors like bortezomib, therefore, should have the potential to provide therapeutic benefit in *S. maltophilia* infections. Since the damaging effects of the recombinant gene product as well as the natural protease were neutralized by bortezomib, it seems allowable that insights gained from the crystallographic studies of recombinant StmPr1 and of its complex with inhibitors could be utilized for future drug discovery projects. The next step, therefore, is to analyze the three-dimensional structure of a complex of StmPr1 with the drug bortezomib.

3.6.2 Crystal structure the bortezomib-StmPr1 complex

In order to form complexes of StmPr1 with bortezomib, crystals of the native protease were soaked in solutions of the inhibitor. Although this was possible only under conditions with a low content of compound and/or organic solvent without crystal cracking, bortezomib was specifically bound as will be demonstrated in the following section. The diffraction data were collected as outlined above. The data collection parameters and refinement statistics are summarized in Table 14.

Results and Discussion

Table 15: Data collection parameters and refinement statistics of the bortezomib StmPr1 complex

Parameters	
Space group	C222 ₁
a (Å)	60.79
b (Å)	86.62
c (Å)	132.09
V _M (Å ³ /Da)	2.12
Solvent content (%)	42%
Completeness of data (%)	97.8 (85.4)
No. of total reflections	53279(6261)
No. of unique reflections	11224(1397)
Average I/sigma intensity	13.6(6.3)
Resolution (Å)	39.0-2.0(2.19-2.0)
Redundancy	4.8(4.5)
*Rmerge (%)	7.7(19.3)
No. of reflections used in refinement	10641
Rcrystal (%)	15.7
No. of reflections used in Rfree	532
Rfree (%)	21.0
Protein atoms	2734
Amino Acids:	356
Calcium Ions	1
Sulphate Ions	5
Glycerol molecules	2
Water molecules	252
Root mean square deviation	
Bonds (Å)	0.015
Bond angles (°)	1.60
Residues in regions of the Ramachandran plot (%)	
Most favored	90.5
Allowed	9.5
Generally allowed	0
Disallowed	0

3.6.2.1 Structure analysis

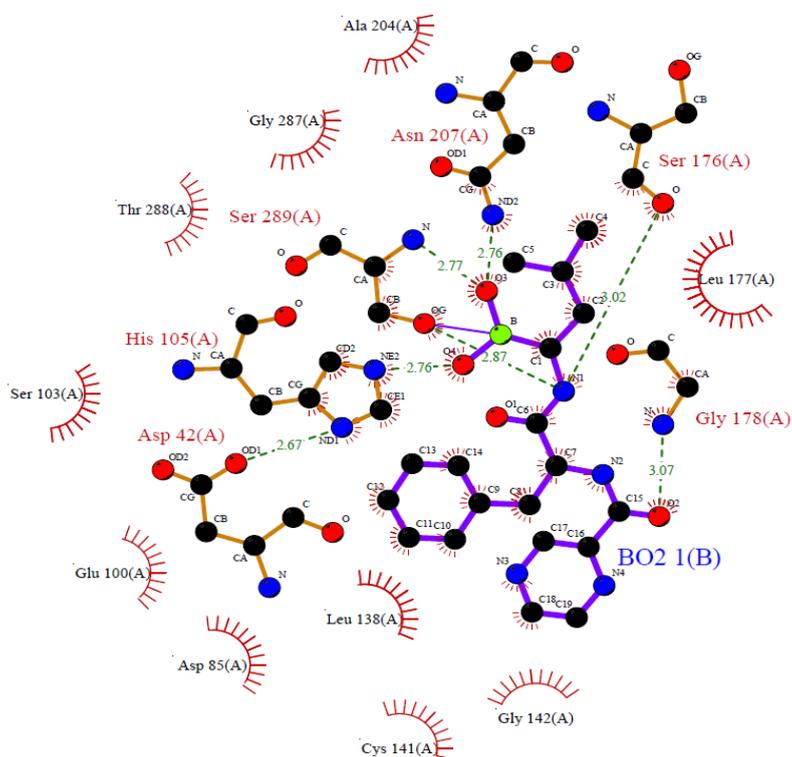


Figure 49: Ligplot of StmPr1 showing bortezomib in the binding site.

Table 16: The interactions of StmPr1-bortezomib complex

Nr.	Residue name	Atom name		Atom name	Distance (Å)	Bond type
1.	Ser-289	OG	↔	B	1.42	Covalent bond
2.	Ser-289	N	↔	O3	2.77	H-bond
3.	Ser-289	OG	↔	N1	2.87	H-bond
4.	Asn-207	ND2	↔	O3	2.76	H-bond
5.	His-105	NE2	↔	O4	2.76	H-bond
6.	Ser-176	O	↔	N1	3.02	H-bond
7.	Gly-178	N	↔	O2	3.07	H-bond

Analysis of the StmPr1-bortezomib complex showed that the active site is occupied with the inhibitor. In its bound conformation; bortezomib adopts an anti-parallel β -sheet conformation. This sheet is stabilized by forming 6 hydrogen bonds between the residues (His-105, Ser-176, Gly-178, Asn-207, and Ser-289) of StmPr1 and the main chain atoms of bortezomib (Figure 49 and 50).

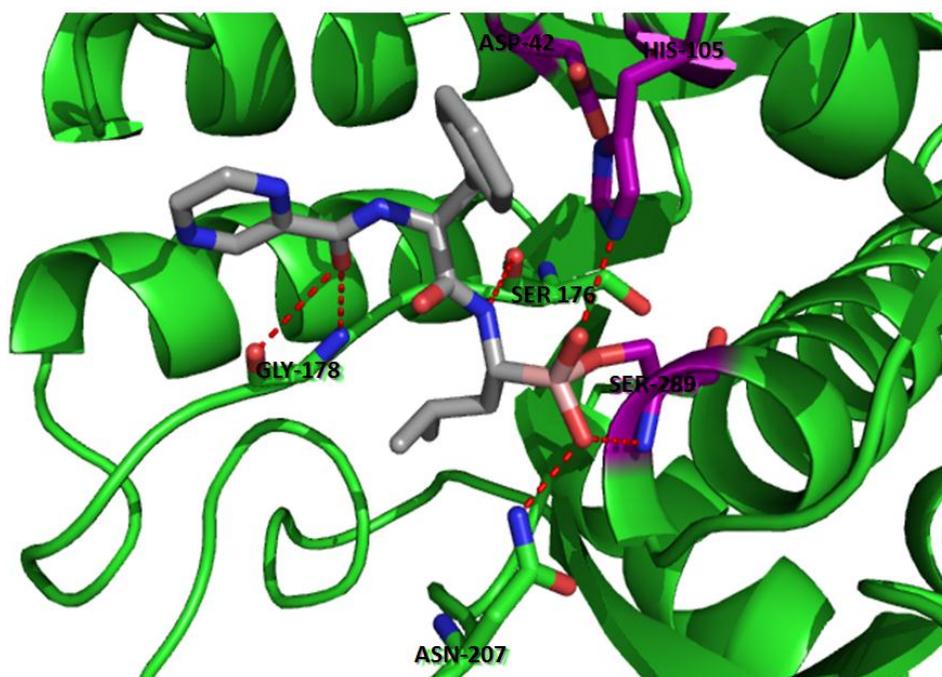


Figure 50: Active site of StmPr1 (top view). Catalytic triad residues (Asp-42, His-105, and Ser-289) are colored and labeled in purple. Main chains and side chains of StmPr1 residues forming polar contacts with Bortezomib (red dashed lines) are shown. Inhibitor is in gray.

The three residues of the inhibitor (Pyz-Phe-boroLeu) are well-localized in the putative S1 through S3 sites. Several structural features of bortezomib seem to be important for its mechanism of inhibition. While in peptide aldehyde inhibitors it is the carbonyl group forming a hemiacetal linkage with Ser-289 (as also shown in the chymostatin and leupeptin structures above), in bortezomib this function is realized by the boronic acid moiety, which ensures high affinity for strong oxygen nucleophiles. The boron atom covalently interacts with the nucleophilic oxygen lone pair of Ser-289 (Figure 51), while Asn-207 stabilizes the oxyanion hole formed hydrogen bond with one of the acidic boronate hydroxyl groups. The tetrahedral boronate adduct is further stabilized by a second acidic boronate hydroxyl moiety which forms hydrogen bond with the nitrogen atom of Ser-289, functioning as a catalytic proton acceptor. The oxygen atom occupying the ‘‘oxyanion hole’’ region of the enzyme accepts two

hydrogen bonds, one from the polypeptide backbone and one from the positively charged amino group of Asn-207.

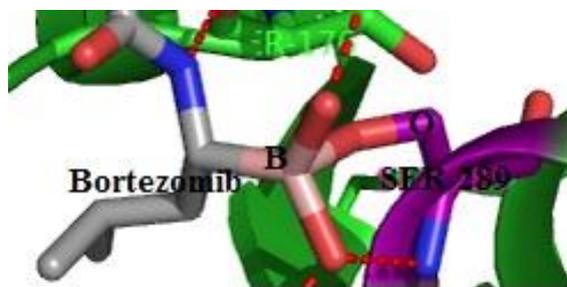


Figure 51: Covalent bond between Ser-289 and boronic acid moiety of bortezomib.

Water molecules play an important role in stabilizing the inhibitor in the active site cleft. It was found that four water molecules (W166, W316, W317, and W321) are bound to the inhibitor back bone (Figure 52). Three of them coordinate a tight hydrogen bonding network which interacts with one of the acidic boronate hydroxyl groups, which stabilizes it because of strong interaction as evidenced by the short bond length. The fourth water molecule interacts with the nitrogen atom of the pyrazine ring of the P3 site.

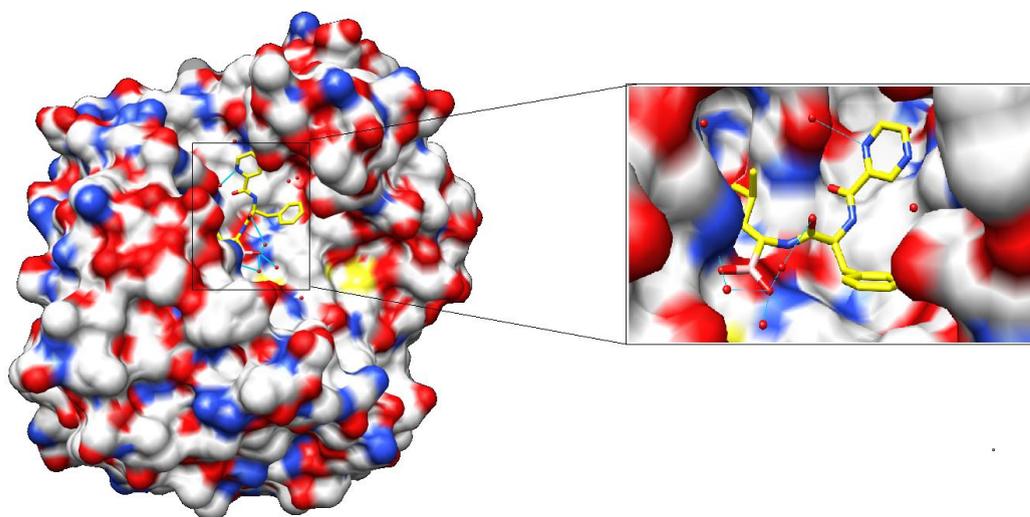


Figure 52: Surface representation of StmPr1 showing bortezomib (in yellow sticks) in the binding site of StmPr1.

The StmPr1 bortezomib complex structure has been superimposed to the StmPr1 native structure, and no significant movements of the active site residues of StmPr1 upon complexation with the inhibitor were observed. The phenylalanine moiety of bortezomib points into the empty space and is not in contact with the protein,

and, because of the unique topology of the S2 pocket, it can accept space demanding side chain residues. It means that the P2 site of bortezomib does not contribute to the kinetics of the inhibition; specific gain may be obtained by introducing larger hydrophobic residues at the P2 site. The electron density reveals a well-defined water molecule which forms a hydrogen bond with a nitrogen atom of the pyrazine ring of bortezomib (the P2 site) which leads to stabilization of this nitrogen atom. Thus, the structure reveals that although bortezomib accommodates the active site cleft of StmPr1, it may not represent the optimal fit. Future alterations of the P1 and P3 sites may improve the binding properties.

All inhibitors tested so far had peptide-like structures. A potential therapeutic use of those compounds would, therefore, be limited to parenteral application. Interestingly, high-throughput screening with StmPr1 as a target had generated further but non-peptidic “hits”. Five low-molecular weight compounds inhibited StmPr1 with IC_{50} less than 10 μ M. Figure 53 shows the chemical structures of these compounds.

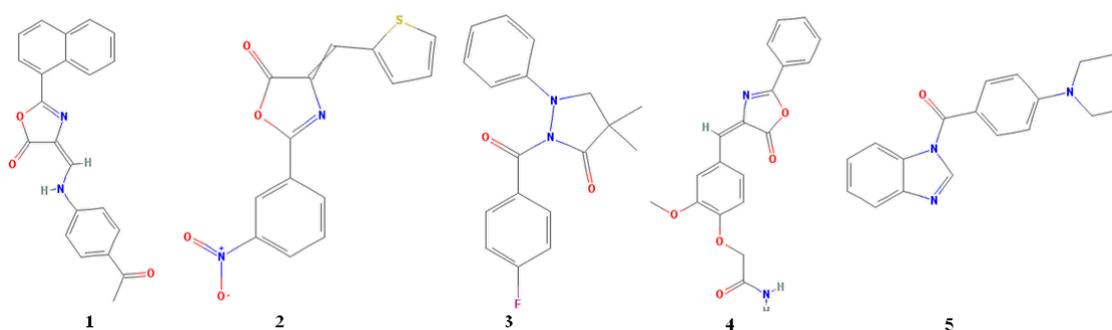


Figure 53: Five positive HTS hits represent non-peptidic compounds.

Each of these compounds was incubated in different concentrations with the StmPr1 for 20 minutes prior to addition of Suc-Ala-Ala-Pro-Phe-PNA substrate and which resulted in significant inhibition of StmPr1 and PNA release.

Out of them, compound #5 (benzimidazol-1-yl-[4-(diethylamino) phenyl] methanone) revealed to be the most effective with IC_{50} of about 0.5 μ M. Soaking of StmPr1 crystals with this compound was successful, and also, diffraction data with maximum resolution of 2.0 Å could be collected at the Consortium-Beam Line X13, DESY Hamburg. The evaluation of these data is ongoing and could not be accomplished within the time frame of this work. But preliminary refinement of the complex structure displayed that the compound is covalently bound to the catalytic Ser-289 via its carbonyl group. It appears that inhibition of StmPr1 requires a reactive carbonyl group which is present in all five hits selected by high-throughput screening.

4. Summary and Outlook

Stenotrophomonas maltophilia is known to cause human diseases as a result of its ability to colonize immune-compromised patients. The pathogen attacks a group of high risk patients, such as those with chronic obstructive pulmonary disease, cancer patients, patients with immune suppression (AIDS, treatment with immune-suppressors, and/or steroids), or patients suffering from acute illnesses, such as myocardial infarction, acute pancreatitis, and patients with a central venous catheter in place. There are further reports demonstrating the involvement of this bacterium in massive hemorrhagic processes of the small intestine and of the subclavian artery accompanied by severe lesions of the tissue. These observations strongly suggest participation of proteolytic activity, produced by the bacteria, which can damage the infected tissue.

Stenotrophomonas maltophilia had been shown to produce two extracellular proteases. The genes have been cloned, termed StmPr1 and StmPr2. This work is limited to the StmPr1 protease as the major secretory protease of *S. maltophilia*. The gene codes for a 63-kDa precursor that is processed to the mature protein of 47 kDa. The amino acid sequence predicted differences in the architecture of the active site as indicated by inserts adjacent to the catalytic His and Ser residues. The StmPr1 protease is able to degrade several human proteins from serum and connective tissue.

The objectives of this work were to analyze the crystal structure of the StmPr1 protease as a prerequisite in searching for inhibiting molecules thereby paving the way for future drug discovery projects. The protein was expressed in *E. coli* in sufficient amounts. Remarkably, the protease was processed by the host bacteria to a core protein of 36 kDa which was fully active and could be crystallized.

The key step in the structure determination process is the collection of X-ray diffraction data. Native diffraction data were collected at the Consortium-Beam line X13, DESY Hamburg. Single crystals were obtained and diffracted synchrotron X-radiation up to a resolution of 1.4 Å. The crystal belongs to the space group C222₁. Matthews coefficient calculations indicated one molecule per asymmetric unit, which corresponds to a packing parameter V_M of 2.17 Å³ Da⁻¹, and a solvent content of approximately 43 %. The phase problem was solved by molecular replacement using

Summary and Outlook

the structure of a homologous subtilisin. The model was refined at a resolution of 1.4 Å with R_{crystal} of 15.36 % and R_{free} of 16.67%.

A total of 356 amino acids has been detected in the three-dimensional structure which corresponds to a molecular weight 36 kDa, confirming the C-terminal truncation of the precursor protein during processing. The structure shows the scaffold characteristic of subtilisin-like serine proteinases. Two disulfide bonds (Cys93–Cys141 and Cys183–Cys220) and one calcium-binding site were detected in the structure. The calcium ion is coordinated in a pentagonal bipyramidal manner. Noteworthy, calcium ion is well conserved among the most related enzymes of subtilisin family.

Comparison with other subtilisin-like proteases reveals several major insertions (termed L1-L5) in the loops that surround the active site cleft. The L2 loop is well defined in the electron density with a moderately high B-factor suggesting that it has some mobility and flexibility characters. However, the apparent stability of the L2 loop is likely to arise from crystal packing. These loops - together with the larger active site cleft - make the StmPr1 protease different from the classical subtilisins and may allow defining a new sub-family of non-classical subtilisins.

The second part of this work focuses on StmPr1 interactions with inhibitory ligands. Two crystal structures of complexes with commercially available protease inhibitors (leupeptin and chymostatin) are presented. Both structures have in common the following features: (i) when superposed to the native structure of StmPr1 no significant movements of the active site occur upon binding of the ligand. (ii) The catalytic Ser-289 of the protease is covalently modified by aldehyde groups of both inhibitors. (iii) Their P3 sites are located at the surface of the protein. The structure indicates that the S1 subsite of StmPr1 can accept a positively charged amino acid (leupeptin) or an aromatic amino acid (chymostatin). The number of hydrogen bonds seen in the structures is higher in chymostatin which is in accordance with the stronger inhibitory effect of this molecule compared to leupeptin.

One further inhibitory molecule for StmPr1 was detected in connection with a project dealing with snake venom peptides. Co-crystallization with such crude preparation from *Agkistrodon bilineatus* venom served to identify in the crystal structure one particular peptide (Ala-Ser-Pro-Ser). The crystal structure analysis revealed a high number of hydrogen bonds between the peptide and active site residues

of StmPr1. This indicates that this peptide – though without a carbonyl group – could be a strong StmPr1 inhibitor. Further analysis is required for conformation.

Crystallography of StmPr1 complexes with the inhibitors described so far provided a lot of information on active site interactions. However, these compounds can be considered as “non-druggable” due to lack of affinity and/or specificity. The search for novel StmPr1 inhibitors was addressed in the current work by two strategies; (i) virtual (“in-silico”) ligand screening (VLS) using the reported crystal structures and (ii) high-throughput screening (HTS) on the basis of enzyme activity assays. While the VLS strategy is an ongoing cooperation project, the HTS strategy has yielded the first positive results.

The most promising “hit” was the drug bortezomib which had been developed as a proteasome inhibitor. Bortezomib is in clinical use since 2004 for the treatment of multiple myeloma. The compound can be regarded as a modified dipeptide with a C-terminal boronic acid moiety. Interestingly, bortezomib showed significant inhibition of StmPr1 in the sub-micromolar range. Most importantly, bortezomib was effective in neutralizing the damaging effect of *S. maltophilia* secretions on lung epithelial tumor cells. The crystal structure of the complex with bortezomib revealed that a covalent bond of the boronic acid moiety with the catalytic Ser-289 represents the basis for its inhibitory action. This close binding by a covalent bond seems to be unique, as it was not reported in case of the proteasome. By forming a covalent bond in the active site of StmPr1, bortezomib may act as a suicide inhibitor reducing the number of functional enzyme molecules. However, the structure reveals that although bortezomib accommodates the active site cleft of StmPr1, it may not represent the optimal fit. Future alterations of the P1 and P3 sites may still improve the binding properties. Since bortezomib is an approved drug, it will be tempting to study whether its therapeutic indication may be extended from an anti-cancer drug to infections with *S. maltophilia*. Here, we present the first pre-clinical data for the use of bortezomib to treat *S. maltophilia* infections. Animal models would be the next step to validate the current data.

Further new compounds were discovered here as StmPr1 inhibitors using high-throughput screening. One of them (benzimidazol-1-yl-[4-(diethylamino) phenyl] methanone) is likely a good candidate for future studies based on: (i) It inhibited StmPr1 in a concentration range comparable to bortezomib. (ii) It did not inhibit

Summary and Outlook

thrombin and kallikrein as human proteases in serum. (iii) It has a non-peptidic structure. (iv) Preliminary crystallographic data of the complex with StmPr1 suggest that the compound forms a covalent bond with the catalytic serine. Therefore, it has the potential to be used as a lead compound for future drug discovery.

Exploiting the collective information derived from the numerous crystal structures presented in the current study can serve to build a “pharmacophore” model which will define how optimal drug looks like.

5. References

1. Swings, J., Devos, P., Vandenmooter, M. and Deley, J., *Transfer of Pseudomonas maltophilia Hugh 1981 to the genus Xanthomonas maltophilia (Hugh 1981)*. Int. J. Syst. Bacteriol., 1983. **33**: p. 409–413.
2. Palleroni, N.J. and Bradbury, J. F., *Stenotrophomonas, a new bacterial genus for Xanthomonas maltophilia*. Int J Syst Bacteriol, 1993. **43**(3): p. 606-9.
3. Coenye, T., Vanlaere, E., Falsen, E. and Vandamme, P., *Stenotrophomonas africana is a later synonym of Stenotrophomonas maltophilia*. Int J Syst Evol Microbiol, 2004. **54**(4): p. 1235-7.
4. Finkmann, W., Altendorf, K., Stackebrandt, E., and Lipski, A., *Characterization of N₂O-producing Xanthomonas-like isolates from biofilters as Stenotrophomonas nitritireducens sp. nov., Luteimonas mephitis gen. nov., sp. nov. and Pseudoxanthomonas broegbernensis gen. nov., sp. nov.* Int J Syst Evol Microbiol, 2000. **50** (1): p. 273-82.
5. Ryan, R.P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M. B., Berg, G., van der Lelie, D., and Dow, J. M., *The versatility and adaptation of bacteria from the genus Stenotrophomonas*. Nat Rev Microbiol, 2009. **7**(7): p. 514-25.
6. Berg, G., Marten, P. and Ballin, G., *Stenotrophomonas maltophilia in the rhizosphere of oil seed rape — occurrence, characterization and interaction with phytopathogenic fungi*. Microbiol. Res., 1996. **151**: p. 19-27.
7. Mahaffee, W. F. and Kloepper, J. W., *Temporal Changes in the Bacterial Communities of Soil, Rhizosphere, and Endorhiza Associated with Field-Grown Cucumber (Cucumis sativus L.)*. Microb Ecol, 1997. **34**(3): p. 210-23.
8. Berg, G., Roskot, N., Steidle, A., Eberl, L., Zock, A. and Smalla, K., *Plant-dependent genotypic and phenotypic diversity of antagonistic rhizobacteria isolated from different Verticillium host plants*. Appl Environ Microbiol, 2002. **68**(7): p. 3328-38.
9. Schwieger, F. and Tebbe, C. C., *Effect of field inoculation with Sinorhizobium meliloti L33 on the composition of bacterial communities in rhizospheres of a target plant (Medicago sativa) and a non-target plant (Chenopodium album)-linking of 16S rRNA gene-based single-strand conformation polymorphism*

References

- community profiles to the diversity of cultivated bacteria. *Appl Environ Microbiol*, 2000. **66**(8): p. 3556-65.
10. Chelius, M. K. and Triplett, E. W., *Immunolocalization of dinitrogenase reductase produced by Klebsiella pneumoniae in association with Zea mays L.* *Appl Environ Microbiol*, 2000. **66**(2): p. 783-7.
 11. Mehnaz, S., Mirza, M. S., Haurat, J., Bally, R., Normand, P., Bano, A. and Malik, K. A., *Isolation and 16S rRNA sequence analysis of the beneficial bacteria from the rhizosphere of rice.* *Can J Microbiol*, 2001. **47**(2): p. 110-7.
 12. Germida, J. and Siciliano, S., *Taxonomic diversity of bacteria associated with the roots of modern, recent and ancient wheat cultivars.* *Biology and Fertility of Soils*, 2001. **33**(5): p. 410-415.
 13. Galai, S., Limam, F. and Marzouki, M. N., *A new Stenotrophomonas maltophilia strain producing laccase. Use in decolorization of synthetic dyes.* *Appl Biochem Biotechnol*, 2009. **158**(2): p. 416-31.
 14. Zhang, Z., Yuen, G. Y., Sarath, G. and Penheiter, A. R., *Chitinases from the Plant Disease Biocontrol Agent, Stenotrophomonas maltophilia C3.* *Phytopathology*, 2001. **91**(2): p. 204-11.
 15. Mastretta, C., Barac, T., Vangronsveld, J., Newman, L., Taghavi, S. and van der Lelie, D., *Endophytic Bacteria and their Potential Application to Improve the Phytoremediation of Contaminated Environments.* *Biotechnology and Genetic Engineering Reviews*, 2006. **23**: p. 175-207.
 16. Park, Y. S., Kim, S. Y., Park, S. Y., Kang, J. H., Lee, H. S., Seo, Y. H. and Cho, Y. K., *Pseudooutbreak of Stenotrophomonas maltophilia bacteremia in a general ward.* *Am J Infect Control*, 2008. **36**(1): p. 29-32.
 17. Paez, J. I. and Costa, S. F., *Risk factors associated with mortality of infections caused by Stenotrophomonas maltophilia: a systematic review.* *J Hosp Infect*, 2008. **70**(2): p. 101-8.
 18. Cheong, H.S., Lee, J. A., Kang, C. I., Chung, D. R., Peck, K. R., Kim, E. S., Lee, J. S., Son, J. S., Lee, N. Y. and Song, J. H., *Risk factors for mortality and clinical implications of catheter-related infections in patients with bacteraemia caused by Stenotrophomonas maltophilia.* *Int J Antimicrob Agents*, 2008. **32**(6): p. 538-40.

References

19. Figueiredo, P.M., Furumura, M. T., Santos, A. M., Sousa, A. C., Kota, D. J., Levy, C. E. and Yano, T., *Cytotoxic activity of clinical Stenotrophomonas maltophilia*. Lett Appl Microbiol, 2006. **43**(4): p. 443-9.
20. Fouhy, Y., Scanlon, K., Schouest, K., Spillane, C., Crossman, L., Avison, M. B., Ryan, R. P. and Dow, J. M., *Diffusible signal factor-dependent cell-cell signaling and virulence in the nosocomial pathogen Stenotrophomonas maltophilia*. J Bacteriol, 2007. **189**(13): p. 4964-8.
21. Waters, V. J., Gomez, M. I., Soong, G., Amin, S., Ernst, R. K. and Prince, A., *Immunostimulatory properties of the emerging pathogen Stenotrophomonas maltophilia*. Infect Immun, 2007. **75**(4): p. 1698-703.
22. Weber, D. J., Rutala, W. A., Sickbert-Bennett, E. E., Samsa, G. P., Brown, V. and Niederman, M. S., *Microbiology of ventilator-associated pneumonia compared with that of hospital-acquired pneumonia*. Infect Control Hosp Epidemiol, 2007. **28**(7): p. 825-31.
23. Avison, M. B., Higgins, C. S., Ford, P. J., von Heldreich, C. J., Walsh, T. R., and Bennett, P. M., *Differential regulation of L1 and L2 beta-lactamase expression in Stenotrophomonas maltophilia*. J Antimicrob Chemother, 2002. **49**(2): p. 387-9.
24. Okazaki, A. and Avison, M. B., *Aph(3')-IIc, an aminoglycoside resistance determinant from Stenotrophomonas maltophilia*. Antimicrob Agents Chemother, 2007. **51**(1): p. 359-60.
25. Zhang, L., Li, X. Z. and Poole, K., *SmeDEF multidrug efflux pump contributes to intrinsic multidrug resistance in Stenotrophomonas maltophilia*. Antimicrob Agents Chemother, 2001. **45**(12): p. 3497-503.
26. Gould, V. C. and Avison, M. B., *SmeDEF-mediated antimicrobial drug resistance in Stenotrophomonas maltophilia clinical isolates having defined phylogenetic relationships*. J Antimicrob Chemother, 2006. **57**(6): p. 1070-6.
27. Falagas, M. E., Valkimadi, P. E., Huang, Y. T., Matthaiou, D. K. and Hsueh, P. R., *Therapeutic options for Stenotrophomonas maltophilia infections beyond co-trimoxazole: a systematic review*. J Antimicrob Chemother, 2008. **62**(5): p. 889-94.
28. Toleman, M. A., Bennett, P. M., Bennett, D. M., Jones, R. N. and Walsh, T. R., *Global emergence of trimethoprim/sulfamethoxazole resistance in*

References

- Stenotrophomonas maltophilia* mediated by acquisition of *sul* genes. Emerg Infect Dis, 2007. **13**(4): p. 559-65.
29. Lawson, D. H. and Paice, B. J., *Adverse reactions to trimethoprim-sulfamethoxazole*. Rev. Infect. Dis., 1982. **4**: p. 429-433.
30. Elting, L. S. and Bodey, G. P., *Septicemia due to Xanthomonas species and non-aeruginosa Pseudomonas species: increasing incidence of catheter-related infections*. Medicine (Baltimore), 1990. **69**(5): p. 296-306.
31. Zuravleff, J. J. and Yu, V. L., *Infections caused by Pseudomonas maltophilia with emphasis on bacteremia: case reports and a review of the literature*. Rev Infect Dis, 1982. **4**(6): p. 1236-46.
32. Muder, R. R., Harris, A. P., Muller, S., Edmond, M., Chow, J. W., Papadakis, K., Wagener, M. W., Bodey, G. P. and Steckelberg, J. M., *Bacteremia due to Stenotrophomonas (Xanthomonas) maltophilia: a prospective, multicenter study of 91 episodes*. Clin Infect Dis, 1996. **22**(3): p. 508-12.
33. Shimoni, S., Abend, Y., Shimon, A., Landau, Z. and Caspi, A., *Stenotrophomonas maltophilia endocarditis following dental treatment in a previously healthy patient*. J Infect, 1998. **37**(3): p. 305-6.
34. Elsner, H. A., Duhrsen, U., Hollwitz, B., Kaulfers, P. M. and Hossfeld, D. K., *Fatal pulmonary hemorrhage in patients with acute leukemia and fulminant pneumonia caused by Stenotrophomonas maltophilia*. Ann Hematol, 1997. **74**(4): p. 155-61.
35. Harlowe, H., *Acute mastoiditis following Pseudomonas maltophilia infection: case report*. Laryngoscope, 1972. **82**: p. 882-883.
36. Caylan, R., Aydin, K. and Koksall, I., *Meningitis caused by Stenotrophomonas maltophilia: case report and review of the literature*. Ann Saudi Med, 2002. **22**(3-4): p. 216-8.
37. Chen, Y.F., Chung, P. C. and Hsiao, C. H., *Stenotrophomonas maltophilia keratitis and scleritis*. Chang Gung Med J, 2005. **28**(3): p. 142-50.
38. Papadakis, K.A., Vartivarian, S. E., Vassilaki, M. E. and Anaissie, E. J., *Septic prepatellar bursitis caused by Stenotrophomonas (Xanthomonas) maltophilia*. Clin Infect Dis, 1996. **22**(2): p. 388-9.
39. Barchitta, M., Cipresso, R., Giaquinta, L., Romeo, M. A., Denaro, C., Pennisi, C. and Agodi, A., *Acquisition and spread of Acinetobacter baumannii and*

References

- Stenotrophomonas maltophilia* in intensive care patients. Int J Hyg Environ Health, 2009. **212**(3): p. 330-7.
40. Marshall W. F., Keating, M. R., Anhalt J. P. and Steckelberg J.M. , *Xanthomonas maltophilia: an emerging nosocomial pathogen*. Mayo Clin Proc., 1989. **64**: p. 1097–1104.
41. Elting L. S., Khardori N., Bodey G. P. and Fainstein V., *Nosocomial infection caused by Xanthomonas maltophilia: a case-control study of predisposing factors*. Infect Control Hosp Epidemiol, 1990. **11**: p. 134-138.
42. Valenza, G., Tappe, D., Turnwald, D., Frosch, M., Konig, C., Hebestreit, H. and Abele-Horn, M., *Prevalence and antimicrobial susceptibility of microorganisms isolated from sputa of patients with cystic fibrosis*. J Cyst Fibros, 2008. **7**(2): p. 123-7.
43. McGowan, J. E., *Resistance in nonfermenting gram-negative bacteria: multidrug resistance to the maximum*. Am J Med, 2006. **119**(6 Suppl 1): p. S29-36; discussion S62-70.
44. Ferrara, A. M., *Potentially multidrug-resistant non-fermentative Gram-negative pathogens causing nosocomial pneumonia*. Int J Antimicrob Agents, 2006. **27**(3): p. 183-95.
45. Spencer, R. C., *The emergence of epidemic, multiple-antibiotic-resistant Stenotrophomonas (Xanthomonas) maltophilia and Burkholderia (Pseudomonas) cepacia*. J Hosp Infect, 1995. **30 Suppl**: p. 453-64.
46. Tan, C. K., Liaw, S. J., Yu, C. J., Teng, L. J. and Hsueh, P. R., *Extensively drug-resistant Stenotrophomonas maltophilia in a tertiary care hospital in Taiwan: microbiologic characteristics, clinical features, and outcomes*. Diagn Microbiol Infect Dis, 2008. **60**(2): p. 205-10.
47. Aisenberg G., Rolston K. V., Dickey B. F., Kontoyiannis D. P., Raad I. I. and Safdar A., *Stenotrophomonas maltophilia pneumonia in cancer patients without traditional risk factors for infection, 1997-2004*. Eur J Clin Microbiol Infect Dis 2007. **26**: p. 13-20.
48. Hanes, S. D., Demirkan, K., Tolley, E., Boucher, B. A., Croce, M. A., Wood, G. C. and Fabian, T. C., *Risk factors for late-onset nosocomial pneumonia caused by Stenotrophomonas maltophilia in critically ill trauma patients*. Clin Infect Dis, 2002. **35**(3): p. 228-35.

References

49. Nseir, S., Di Pompeo, C., Brisson, H., Dewavrin, F., Tissier, S., Diarra, M., Boulo, M. and Durocher, A., *Intensive care unit-acquired Stenotrophomonas maltophilia: incidence, risk factors, and outcome*. Crit Care, 2006. **10**(5): p. R143.
50. Tamura, H., Yamashita, S., Kusano, N., Suzuki, C., Yamaguchi, Y., Tanigawa, K., Masuhara, M., Okita, K. and Murakami, F., *Fulminant hepatitis complicated by small intestine infection and massive hemorrhage*. J Gastroenterol, 1998. **33**(3): p. 412-8.
51. Ohkoshi, Y., Ninomiya, H., Mukai, H. Y., Mochizuki, N., Hori, M., Nagasawa, T., Jikuya, T. and Saida, Y., *Pseudoaneurysm of the subclavian artery due to Xanthomonas pneumonia in a patient with acute myeloid leukemia: its rupture treated by transcatheter coil embolization*. Intern Med, 1999. **38**(8): p. 671-4.
52. Siezen, R. J. and Leunissen, J. A., *Subtilases: the superfamily of subtilisin-like serine proteases*. Protein Sci, 1997. **6**(3): p. 501-23.
53. Billington, S. J., Johnston, J. L. and Rood, J. I., *Virulence regions and virulence factors of the ovine footrot pathogen, Dichelobacter nodosus*. FEMS Microbiol Lett, 1996. **145**(2): p. 147-56.
54. Windhorst, S., Frank, E., Georgieva, D. N., Genov, N., Buck, F., Borowski, P. and Weber, W., *The major extracellular protease of the nosocomial pathogen Stenotrophomonas maltophilia: characterization of the protein and molecular cloning of the gene*. J Biol Chem, 2002. **277**(13): p. 11042-9.
55. Windhorst, S., *Neue extrazelluläre Proteinasen des Krankenhauskeims Stenotrophomonas maltophilia als potentiell Pathogenitätsfaktoren: Sequenzierung, Klonierung, rekombinante Produktion und vergleichende Charakterisierung*, in *Biology Department*. 2002, Hamburg: Hamburg. p. 101.
56. Sims, G. K., *Nitrogen Starvation Promotes Biodegradation of N-Heterocyclic Compounds in Soil*. Soil Biology & Biochemistry 2006. **38**: p. 2478-2480.
57. Rawlings, N. D., Barrett, A. J. and Bateman, A., *MEROPS: the peptidase database*. Nucleic Acids Res, 2010. **38**(Database issue): p. D227-33.
58. Rawlings, N. D. and Barrett, A. J., *Families of serine peptidases*. Methods Enzymol, 1994. **244**: p. 19-61.
59. Dias, D. R., Vilela, D. M., Silvestre, M. P. C. and Schwan, R. F., *Alkaline protease from Bacillus sp. isolated from coffee bean grown on cheese whey*. World J. Microbiol. Biotechnol., 2008. **24**: p. 2027-2034.

References

60. Rawlings, N. D. and Barrett, A. J., *Evolutionary families of peptidases*. *Biochem J*, 1993. **290** (1): p. 205-18.
61. Saeki, K., Ozaki, K., Kobayashi, T. and Ito, S., *Detergent alkaline proteases: enzymatic properties, genes, and crystal structures*. *J Biosci Bioeng*, 2007. **103**(6): p. 501-8.
62. Betzel, C., Gourinath, S., Kumar, P., Kaur, P., Perbandt, M., Eschenburg, S. and Singh, T. P., *Structure of a serine protease proteinase K from *Tritirachium album limber* at 0.98 Å resolution*. *Biochemistry*, 2001. **40**(10): p. 3080-8.
63. Sahl, H. G., Jack, R. W. and Bierbaum, G., *Biosynthesis and biological activities of lantibiotics with unique post-translational modifications*. *Eur J Biochem*, 1995. **230**(3): p. 827-53.
64. Siezen, R. J., Rollema, H. S., Kuipers, O. P. and de Vos, W. M., *Homology modelling of the *Lactococcus lactis* leader peptidase NisP and its interaction with the precursor of the lantibiotic nisin*. *Protein Eng*, 1995. **8**(2): p. 117-25.
65. Booth, M. C., Bogie, C. P., Sahl, H. G., Siezen, R. J., Hatter, K. L. and Gilmore, M. S., *Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic*. *Mol Microbiol*, 1996. **21**(6): p. 1175-84.
66. Van de Ven, W. J., Roebroek, A. J. and Van Duijnhoven, H. L., *Structure and function of eukaryotic proprotein processing enzymes of the subtilisin family of serine proteases*. *Crit Rev Oncog*, 1993. **4**(2): p. 115-36.
67. Barr, P. J., *Mammalian subtilisins: the long-sought dibasic processing endoproteases*. *Cell*, 1991. **66**(1): p. 1-3.
68. Siezen, R. J., de Vos, W. M., Leunissen, J. A. and Dijkstra, B. W., *Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases*. *Protein Eng*, 1991. **4**(7): p. 719-37.
69. Blow, D., *Structure and Mechanism of Chymotrypsin*. *Accounts Chem. Res.*, 1976. **9**: p. 145-152.
70. Kennan, R. M., Wong, W., Dhungyel, O. P., Han, X., Wong, D., Parker, D., Rosado, C. J., Law, R. H., McGowan, S., Reeve, S. B., Levina, V., Powers, G. A., Pike, R. N., Bottomley, S. P., Smith, A. I., Marsh, I., Whittington, R. J., Whisstock, J. C., Porter, C. J. and Rood, J. I., *The subtilisin-like protease AprV2 is required for virulence and uses a novel disulphide-tethered exosite to bind substrates*. *PLoS Pathog*, 2010. **6**(11): p. e1001210.

References

71. Georgieva, D. N., Stoeva, S., Voelter, W., Genov, N. and Betzel, C., *Substrate specificity of the highly alkalophilic bacterial proteinase esperase: relation to the x-ray structure*. *Curr Microbiol*, 2001. **42**(5): p. 368-71.
72. Laemmli, U. K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. *Nature*, 1970. **227**(5259): p. 680-5.
73. Ollis, D. and S. White, *Protein crystallization*. *Methods Enzymol*, 1990. **182**: p. 646-59.
74. McPherson, A., Malkin, A. J. and Y.G. Kuznetsov, *The science of macromolecular crystallization*. *Structure*, 1995. **3**(8): p. 759-68.
75. Provencher, S. W., *CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations*. *Comput. Phys. Commun*, 1982. **27**: p. 213-229.
76. Walter, T. S., Mancini, E. J., Kadlec, J., Graham, S. C., Assenberg, R., Ren, J., Sainsbury, S., Owens, R. J., Stuart, D. I., Grimes, J. M. and Harlos, K., *Semi-automated microseeding of nanolitre crystallization experiments*. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 2008. **64**(1): p. 14-8.
77. Otwinowski, Z. and Minor, W., *Processing of X-ray Diffraction Data Collected in Oscillation Mode*. *Methods Enzymol*, 1997. **276**: p. 307–326.
78. Vagin, A. and Teplyakov, A., *Molecular replacement with MOLREP*. *Acta Crystallogr D Biol Crystallogr*, 2010. **66**(1): p. 22-5.
79. Collaborative Computational Project, Number 4, *The CCP4 suite: programs for protein crystallography*. *Acta Crystallogr D Biol Crystallogr*, 1994. **50**(5): p. 760-3.
80. Matthews, B.W., *Solvent content of protein crystals*. *J Mol Biol*, 1968. **33**(2): p. 491-7.
81. Emsley, P. and Cowtan, K., *Coot: model-building tools for molecular graphics*. *Acta Crystallogr D Biol Crystallogr*, 2004. **60**(12): p. 2126-32.
82. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. and Murshudov, G. N., *REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use*. *Acta Crystallogr D Biol Crystallogr*, 2004. **60**(12): p. 2184-95.
83. Brunger, A. T., *Free R value: a novel statistical quantity for assessing the accuracy of crystal structures*. *Nature*, 1992. **355**(6359): p. 472-5.

References

84. Hofmann, K. and Baron, M. D., *BOXSHADE: printouts from Aligned Protein or DNA Sequences*. 2001 [cited; Available from: <http://bioweb.pasteur.fr/seqanal/interfaces/boxshade.html>].
85. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*. *Nucleic Acids Res*, 1997. **25**(17): p. 3389-3402.
86. Rawlings, N. D., Morton, F. R. and Barrett, A. J., *MEROPS: the peptidase database*. *Nucleic Acids Res*, 2006. **34**(Database issue): p. D270-2.
87. Schechter, I. and Berger, A., *On the size of the active site in proteases. I. Papain*. *Biochem Biophys Res Commun*, 1967. **27**(2): p. 157-62.
88. Takeuchi, Y., Noguchi, S., Satow, Y., Kojima, S., Kumagai, I., Miura, K., Nakamura, K. T. and Mitsui, Y., *Molecular recognition at the active site of subtilisin BPN': crystallographic studies using genetically engineered proteinaceous inhibitor SSI (Streptomyces subtilisin inhibitor)*. *Protein Eng*, 1991. **4**(5): p. 501-8.
89. Betzel, C., Teplyakov, A. V., Harutyunyan, E. H., Saenger, W. and Wilson, K. S., *Thermitase and proteinase K: a comparison of the refined three-dimensional structures of the native enzymes*. *Protein Eng*, 1990. **3**(3): p. 161-72.
90. Almog, O., Gonzalez, A., Klein, D., Greenblatt, H. M., Braun, S. and Shoham, G., *The 0.93Å crystal structure of sphericase: a calcium-loaded serine protease from Bacillus sphaericus*. *J Mol Biol*, 2003. **332**(5): p. 1071-82.
91. Teplyakov, A. V., Kuranova, I. P., Harutyunyan, E. H., Vainshtein, B. K., Frommel, C., Hohne, W. E. and Wilson, K. S., *Crystal structure of thermitase at 1.4 Å resolution*. *J Mol Biol*, 1990. **214**(1): p. 261-79.
92. Betzel, C., Pal, G. P. and Saenger, W., *Three-dimensional structure of proteinase K at 0.15-nm resolution*. *Eur J Biochem*, 1988. **178**(1): p. 155-71.
93. Alexander, P. A., Ruan, B. and Bryan, P. N., *Cation-dependent stability of subtilisin*. *Biochemistry*, 2001. **40**(35): p. 10634-9.
94. Alexander, P. A., Ruan, B., Strausberg, S. L. and Bryan, P. N., *Stabilizing mutations and calcium-dependent stability of subtilisin*. *Biochemistry*, 2001. **40**(35): p. 10640-4.

References

95. Morihara, K. and Oka, T., *The complex active sites of bacterial neutral proteases in relation to their specificities*. Biochem Biophys Res Commun, 1968. **30**(6): p. 625-30.
96. Kulakova, L., Galkin, A., Kurihara, T., Yoshimura, T. and Esaki, N., *Cold-active serine alkaline protease from the psychrotrophic bacterium Shewanella strain ac10: gene cloning and enzyme purification and characterization*. Appl Environ Microbiol, 1999. **65**(2): p. 611-7.
97. Kuramochi, H., Nakata, H. and Ishii, S., *Mechanism of association of a specific aldehyde inhibitor, leupeptin, with bovine trypsin*. J Biochem, 1979. **86**(5): p. 1403-10.
98. Tatsuta, K., Mikami, N., Fujimoto, K., Umezawa, S. and Umezawa, H., *The structure of chymostatin, a chymotrypsin inhibitor*. J Antibiot (Tokyo), 1973. **26**(11): p. 625-46.
99. Gorenstein, D. G. and Shah, D. O., *Proton and fluorine nuclear magnetic resonance spectroscopic observation of hemiacetal formation between N-acyl-p-fluorophenylalaninals and alpha-chymotrypsin*. Biochemistry, 1982. **21**(19): p. 4679-86.
100. Adams, J., Behnke, M., Chen, S., Cruickshank, A. A., Dick, L. R., Grenier, L., Klunder, J. M., Ma, Y. T., Plamondon, L. and Stein, R. L., *Potent and selective inhibitors of the proteasome: dipeptidyl boronic acids*. Bioorg Med Chem Lett, 1998. **8**(4): p. 333-8.

Risk and safety statements for the compounds used in the study

Name of the reagent	Risk phrases	Safety Phrases
Acetic acid	10-35	23-26-45
Ampicillin	42/43	23-36/37-45
Bromophenol blue		22-24/25
Calcium chloride	36	22-24
Coomassie brilliant blue		22-24/25
Disodium hydrogen phosphate		22-24/25
Dithiothreitol (DTT)	22-36/37/38	26-36
DMSO	36/37/38	23-26-36
EDTA	36-52/53	61
Ethanol	11	7-16
Hydrochloric acid	34-37	26-36/37/39-45
Leupeptin		22-24/25
Magnesium chloride		22-24/25
Methanol	11-23/24/25-39/23-24/25	7-16-36/37-45
Penicillin	42/43	22-36/37
Potassium chloride		22-24/25
Potassium dihydrogen phosphate	34	26-45
Qiagen plasmid mini kit	10-35/36/38/11-36-67/42/43	13/26/26/46

Risk and Safety Statements

Quiagen plasmid Maxi kit	10-35/36/38/11-36-67/42/43	13/26/26/46
SDS	11-21/22-36/37/38	26-36/37
Sodium hydroxide	35	26-37/39-45
β-mercaptoethanol	20/22-24-34-51/53	26-36/37/39-45-61

Acknowledgement

- First of all I kneel to Allah truly thankful for supporting me to continue my research and compile this thesis.
- I send all my respect and gratitude to the Egyptian Ministry of High Education for financially supporting my Ph.D. program.
- I'd like to dedicate this work to my family and all my friends. They are pleased, I hope that they are a bit proud of me too, and I guess I owe them a lot. My Mom and Dad have to be thanked for always being enthusiastic about everything. They're brilliant to talk a problem through with, and I've appreciated their help and advice on loads of things. Thanks to my brothers Khalid, Youssef and Mohammed for being my closest friends. Thanks to my lovely wife Shaimaa and my children, Abdelrahman and Judi, for cheering me up when necessary and making me feel the warm during what I have been doing. Also, thanks for my wife's family specially her Mom and Dad for supporting me a lot during my work here and in Egypt.
- I would like to express my deepest appreciation to Prof. Christian Betzel. It would have been very difficult to continue my life in Germany without his support. More importantly, his kindness and scientific guidance were major factors that helped me to continue in this thesis. I thank him for introducing me into the field of "crystallography", for his non-pressing guidance and pushing me to think more and more.
- I am deeply indebted to PD. Dr. Wolfgang Weber, my research advisor; for a lot of helpful conversations, sharing ideas and personal experience, pushing me to think more about biological meaning of phenomena than about the technical questions. His brilliant discussions and comments have made me enjoy working with him. Without his help, it would have been impossible to continue my stay in Germany. Also thanks for putting up with my stresses during my writing over the past few months, and has been lovely all the time, even when his attempts to get me to relax have been thwarted.

Acknowledgement

- I'd like to thank all of the people in the Laboratory for Structural Biology of Infection and Inflammation at DESY, and also in the Laboratory of Biochemistry and Molecular Biology Institute at UKE and also for the open and collaborative atmosphere.
- I deeply thank my colleague Ahmed Akrem for supporting me in some experiments performed in this thesis. I wish him a great success in his career.
- I send my greetings to the whole staff in Biochemistry Department in Faculty of Science, Mansoura University in Egypt for facilitating my stay here in Germany.
- Thanks to everyone else who's been in the lab at any stage in the past four years for providing me with help, support and other fun.

AMR NEGM

Curriculum Vitae

Personal Data

Family Name: Negm

First Name: Amr

Address: Kaemmererufer 16, 22303-Hamburg

Date of Birth: 1st of January 1981

Birthplace: El-Dakahlia, Egypt

Nationality: Egyptian

Graduation

1997: High Secondary School, ElDakahlia, Egypt

1997-2001: B. Sc. In Biochemistry, Faculty of Science, Mansoura University, Egypt

2002-2006: M. Sc. in Biochemistry, Faculty of Science, Mansoura University, Egypt

Since 2007: Studying Ph. D. at the University of Hamburg, Germany

Work experience

2002 – 2006: Demonstrator of Biochemistry in the Faculty of Science, Mansoura University, Egypt. During this period I pursued my Master degree.

Topic: Hypoglycemic effects of some plant extracts in diabetic rats.

2006 – till now: Assistant Lecturer of Biochemistry in the Faculty of Science, Mansoura University, Egypt.

2007 – 2011: Scholarship to obtain a Ph.D. degree at the University of Hamburg in cooperation with University Hospital Hamburg-Eppendorf (UKE), funded by The Ministry of Higher Education.

Selbständigkeitserklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne unerlaubte fremde Hilfe angefertigt habe. Ich habe keine anderen als die im Literaturverzeichnis angeführten Quellen benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Zusammenarbeiten sind im Text genannt. Die Arbeit wurde zuvor keiner Prüfungsbehörde in gleicher oder ähnlicher Form vorgelegt.

Amr Negm

Hamburg, 28.November 2011