

**Aus dem Institut für Biochemie  
Der Universität zu Lübeck  
Direktor: Prof. Dr. rer. nat. Dr. h.c. Rolf Hilgenfeld**

**Structural and biochemical analysis of the *Plasmodium falciparum*  
cysteine protease falcipain-2 and its interaction with host proteins**

Inauguraldissertation  
zur  
Erlangung der Doktorwürde  
der Universität zu Lübeck

- Aus der Technisch-Naturwissenschaftlichen Fakultät -

Vorgelegt von  
**Krishna Nagarajan**  
aus Mayiladuthurai, Indien  
Lübeck 2009

1. Berichterstatter/Berichterstatterin: **Prof. Dr. rer. nat. Dr. h.c. Rolf Hilgenfeld**

2. Berichterstatter/Berichterstatterin: **PD Dr. Volker Heussler**

Tag der mündlichen Prüfung: 01.09.2009

Zum Druck genehmigt. Lübeck, den 02.09.2009

## ABSTRACT

Malaria is one of the major infectious diseases in humans and results in 1 - 3 million deaths each year worldwide. *Plasmodium falciparum* is responsible for 80% of malaria-related deaths. Worldwide, malaria is considered to be one of the most pressing health priorities due to parasite resistance against various antimalarials, drug toxicity and high costs. Malarial protease, especially falcipain-2 (FP-2), is a potential target for new mode of chemotherapy. FP-2 belongs to the C1A family of cysteine proteases. It degrades the host-cell hemoglobin at acidic pH and cleaves erythrocyte cytoskeletal proteins at neutral pH. Biochemical and structural studies were carried out to identify the underlying mechanism of this multi-functional protease. Biochemical methods demonstrate that (i), mature FP-2 can proteolytically process its own precursor at basic pH; (ii), the binding of hemoglobin to FP-2 is strictly pH-dependent; and (iii), FP-2 preferentially binds methemoglobin over hemoglobin. The crystal structure of FP-2 was solved at 3.1 Å resolution and revealed several structural features (N-terminal extension, flexible  $\beta$ -hairpin motif etc) that are unique to plasmodial proteases. To understand the FP-2 : hemoglobin interaction and the degradation specificity, myoglobin, a structural and functional homologue of hemoglobin, was used as a model. FP-2 efficiently degraded myoglobin in a pH-dependent manner, producing distinct intermediates. Time-dependent analysis of cleavage products revealed primary and secondary cleavage sites with different sequence preferences. The results show that the removal of the heme moiety is systematic and degradation of substrate by FP-2 is a highly coordinated process. A more detailed model of hemoglobin degradation during the plasmodial life cycle is being proposed. Interference with proteolytic degradation of hemoglobin holds great promise as a new mode for chemotherapy of plasmodial infections.

## **ACKNOWLEDGEMENTS**

I express my sincere thanks with deep sense of gratitude to Prof. Dr. Dr. h.c. Rolf Hilgenfeld, Director of the Institute of Biochemistry, University of Lübeck, for offering me an opportunity to join his distinguished science group and to pursue my dream. His wide scientific knowledge, exceptional memory and involvement have always been truly inspirational and it means a lot to me. He showed how to approach a scientific problem and always emphasized on the need of hard work combined with persistence to accomplish any goal.

This work has been made possible with the financial support from The Deutsche Forschungsgemeinschaft (DFG Hi 611/5-1).

Special thanks to Dr. Christian Schmidt for his dedicated guidance throughout the project. Christian has been a friend and mentor. Also, special thanks to Dr. Tanis Hogg for his stimulating ideas and innovative discussions. They both played a great role in shaping my scientific life. I also extend my gratitude to Doris Mutschall for teaching me the basic techniques necessary for a scientific work and Dr. Guido Hansen. His timely help and support lately deserves a special mention.

Let me also say 'thank you' to the following people: Dr. Jeroen Mesters, Prof. Dr. Stefan Anemüller, Dr. Jörg Deiwick, Dr. Koen Verschueren, Dr. Jinzhi Tan, Dr. Saravanakumar Shanmuganathan, Dr. Ksenia Pumpor, Dr. Shuai Chen, Dr. Sebastián Klinke, Dr. Lili Zhu, Walter Verheyen, Hans-Joachim Kraus, Helga Lorenz, Susanne Zoske, Silke Schmidtke, Angelika von Keiser-Gerhus, Christina Leister and to my fellow PhD student folks: Rajesh, Raspudin, Mirko, Helgo, Yvonne, Sebastian,

Nele P, Robert, Santosh, Monarin, Jiajie, Yibei and many others for interesting discussions and being fun to be with. Thanks guys.

I thank my family: my parents, Nagarajan Krishnamurthy and Devaki Nagarajan for giving me life in the first place, for providing me with the finest education with aspects from both arts and sciences, for unconditional love, support and encouragement to pursue my interests, whatever the interests may be. I thank my brothers, Raghavan and Murari for their innocent intriguing questions about my project and moral support. I thank all my relatives, friends and well wishers for their undiluted love and constant presence. Last, but not least, I thank Karin for her constant encouragement and guiding me through the difficult times and for believing in me. I dedicate this work to them.



## TABLE OF CONTENTS

<b>1.</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	Infections/ Diseases	1
1.2	<i>Plasmodium falciparum</i> malaria	2
1.2.1	Erythrocyte life cycle of <i>Plasmodium falciparum</i>	3
1.2.2	Overview of <i>P. falciparum</i> proteins	4
1.3	Established antimalarial drugs and their limitations	5
1.4	Malarial proteases: New target for chemotherapy	7
1.4.1	Aspartic proteases of the food vacuole	8
1.4.2	Cysteine proteases of the food vacuole	8
1.4.3	Metalloprotease present in the food vacuole	9
1.5	FP-2 as promising drug target	9
1.5.1	FP-2 biochemical properties	10
1.5.2	Uncertainty in native hemoglobin cleavage	11
1.5.3	FP-2: hemoglobin binding models	12
1.6	Scope and objectives of this work	13
<b>2.</b>	<b>MATERIALS AND METHODS</b>	<b>14</b>
2.1	Materials	14
2.1.1	Equipments	14
2.1.2	Chemicals	15
2.1.3	Crystallization materials and cryo-tools	15
2.1.4	Buffers and solutions	16
2.2	Methods	16
2.2.1	Large-scale expression, purification and refolding	16

2.2.2	General activity assay	21
2.2.3	Self-processing of the FP-2 prodomain	21
2.2.4	Self-degradation of wild-type active mature FP-2	21
2.2.5	Crystallization experiments	22
2.2.5.1	Crystallization of FP-2	23
2.2.5.2	Crystallization of <i>Plasmodium berghei</i> inhibitor cysteine protease C-terminal (PbICP)	24
2.2.6	Characterization of protein crystals and optimization of Cryo-conditions	24
2.2.7	Diffraction-data collection	25
2.2.8	Structure determination and refinement	25
2.2.9	Electrostatic potential map calculations	27
2.2.10	Homology modeling of different FP-2 constructs	27
2.2.11	Hemoglobinase and myoglobinase assay	29
2.2.12	Inhibition assay studies	29
2.2.13	Surface plasmon resonance (SPR)	30
2.2.13.1	Surface plasmon resonance studies	31
2.2.14	MALDI-TOF mass spectrometry	32
2.2.14.1	MALDI-TOF analysis of substrate cleavage products	34
2.2.15	NH <sub>2</sub> -terminal sequence analysis	35
2.2.16	Protein-protein docking studies	35
2.2.17	PbICP C-terminal sequence alignment	36
<b>3.</b>	<b>RESULT AND DISCUSSION</b>	<b>37</b>
3.1	FP-2 general activity assays	37
3.1.1	FP-2 enzyme activity determination assays	37

3.1.2	pH-dependence of the activity with artificial substrate	38
3.2	FP-2 self processing assays	39
3.2.1	Auto-processing of FP-2 prodomain	39
3.2.2	Self-degradation assay for FP-2 at pH 5.0 to 9.0	41
3.3	Crystallization and Structure determination of FP-2	44
3.3.1	Crystallization of recombinant wild-type FP-2	44
3.3.2	X-ray diffraction data	46
3.3.3	Overall crystal structure	48
3.4	Sequential and structural comparison with C1A proteases	51
3.5	Three-dimensional structure insights of FP-2	55
3.5.1	The active-site cleft	55
3.5.2	Subsite-specificity of FP-2	57
3.5.3	An additional Cys99-Cys119 disulphide bond in FP-2	58
3.5.4	Structural features unique to the FP-2 subfamily	58
3.5.4.1	The N-terminal extension	60
3.5.4.2	N-terminus as a structural determinant of protein folding	61
3.5.4.3	The unique motif / loop	62
3.5.4.4	Comparison with the FP-2 : cystatin complex	63
3.6	Mechanism of hemoglobin degradation	65
3.6.1	Electrostatic potential map calculation of FP-2	65
3.6.2	Homology modeling of different FP-2 constructs	66
3.7	Hemoglobin degradation assays	69
3.7.1	The pH-dependent activity of FP-2 with hemoglobin	69
3.7.2	Preference for methemoglobin over hemoglobin	71
3.7.3	FP-2 $\Delta$ 185-194 mutant cleavage assays	72

3.7.4	Hemoglobin cleavage assay by FP-2 $\Delta$ KK/G	74
3.8	FP-2 studies with myoglobin as a substrate	76
3.8.1	Hemoglobin binding models	76
3.8.2	Proteolytic cleavage of myoglobin by FP-2	77
3.8.3	Proteolytic cleavage of myoglobin by FP-2 $\Delta$ 185-194	79
3.8.4	Analysis of myoglobin primary cleavage products	80
3.8.5	Analysis of secondary cleavage products	83
3.8.6	NH <sub>2</sub> -terminal amino-acid analysis	85
3.8.7	Analysis of hemoglobin cleavage products	86
3.8.8	Analysis of FP-2 substrate specificity	89
3.9	Structural significance of cleavage products	91
3.9.1	Significance of FP-2 subsite preferences	91
3.9.2	Prediction of additional hemoglobin cleavage sites	93
3.10	Role of FP-2 in heme detoxification	96
3.11	Residual activity of the FPc285aM mutant	102
3.11.1	Active-site mutant protease activity	102
3.11.2	FPc285aM mutant activity alongside wild-type FP-2	103
3.11.3	MALDI-TOF analysis of cleavage products	104
3.11.4	Degradation assay in the presence of reducing agent	105
3.11.5	Protein-protein docking studies	106
3.12	FP-2 Inhibitor studies	110
3.12.1	Falstatin, the natural inhibitor of FP-2	110
3.12.2	Sequence alignment and homology modeling of PbICP	110
3.12.3	Activity assay in the presence of PbICP C-terminal	112
3.12.4	Crystallization of PbICP C-terminal	113

3.13	Search for other potential <i>in-vivo</i> substrates	114
3.13.1	FP-2 cleavage assays in presence of Carbonic anhydrase 1	114
3.14	Search for novel cysteine protease Inhibitors	115
3.14.1	Virtual screen and binding studies of FP-2 inhibitors	115
3.14.2	Myoglobin degradation assay in the presence of inhibitors	115
3.14.3	Degradation assay in the presence of compound X	117
<b>4.</b>	<b>SUMMARY AND OUTLOOK</b>	<b>118</b>
<b>5.</b>	<b>REFERENCES</b>	<b>121</b>

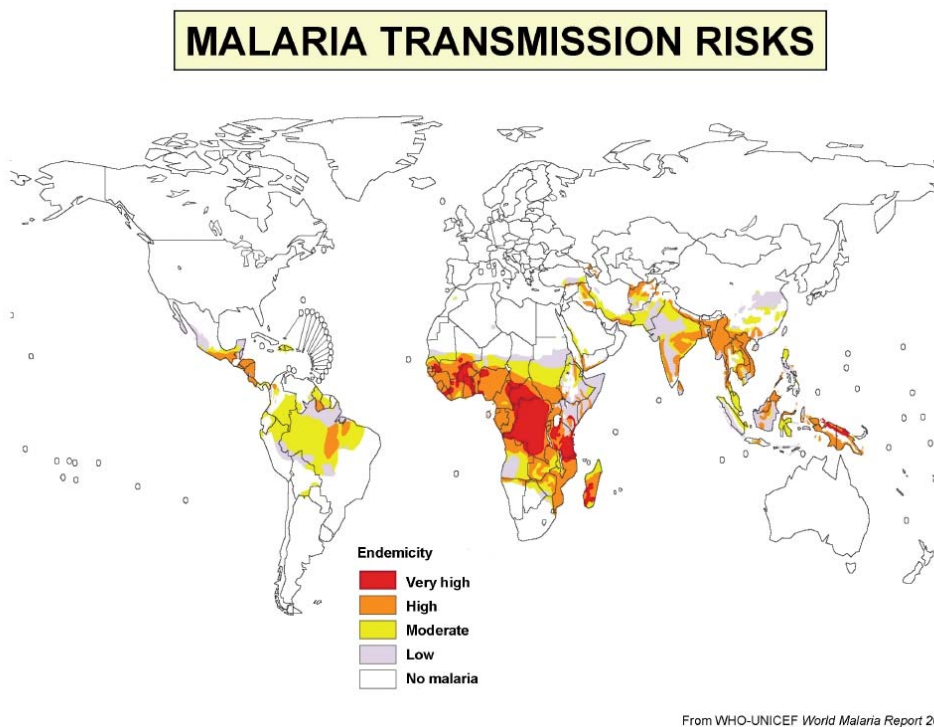
## ABBREVIATIONS, UNITS, and SYMBOLS

AA	Amino acid
ACN	Acetonitrile
AMBER	Assisted Model Building with Energy Refinement
B-factor	Temperature factor
CA1	Carbonic anhydrase Isoform 1
CCP4	Collaborative Computational Project Number 4
CNS	Crystallography and NMR System
Da/kDa	Dalton/kiloDalton
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
E64	L-trans-epoxysuccinyl-leucyl-amido(4-guanidino)butane
EDTA	Ethylene Diamine Tetraacetic Acid
<i>E.coli</i>	<i>Escherichia coli</i>
FP-2	falcipain-2
G6PDH	Glucose-6-phosphate dehydrogenase
GSH	Glutathione
GSSG	Glutathione disulphide
Hb	hemoglobin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HHb/RHb	human hemoglobin/rat hemoglobin
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
<i>K<sub>d</sub></i>	Dissociation Constant
MALDI	Matrix-Assisted Laser Desorption/Ionization
Mb	myoglobin

MS	Mass Spectrometry
MultAlin	Multiple sequence Alignment
NO	Nitrous Oxide
PBE	Poisson-Boltzmann Equation
PBICP C	<i>Plasmodium berghei</i> inhibitor of cysteine proteases (C-terminal)
PDB	Protein Data Bank
PEG	Polyethylene glycol
PHYRE	Protein Homology/analogY Recognition Engine
PM	plasmepsins
r.m.s.d.	Root Mean Square Deviation
RU	Response unit
SDA	Simulation of Diffusional Association
SDS-PAGE	Sodium Dodecyl Sulfate polyacrylamide gel electrophoresis
SPDBV	Swiss PDB Viewer
SPR	Surface plasmon resonance
SSE	Secondary Structural Elements
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
TOF	Time of Flight
UHBD	University of Houston Brownian Dynamics
wt	wild-type
Z-Phe-Arg-pNA	Benzoyl Phe-Arg- <i>p</i> -nitroanilide

## 1. INTRODUCTION

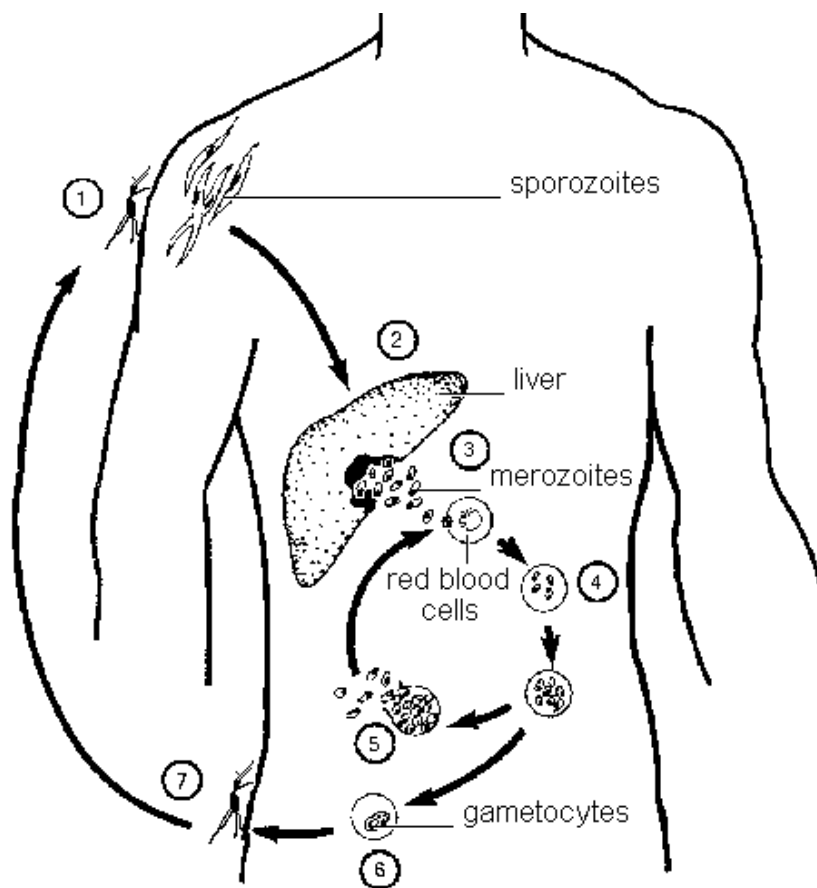
**1.1 Infections / Diseases** - Malaria is one of the major vector-borne infectious diseases known to man for several thousands of years. The disease is wide-spread and infects approximately 40% of the world's population living in the tropical and subtropical regions of Asia, America and Africa (Fig. 1.1, Snow et al., 2005). Malaria infects an estimated 300 - 500 million people and is responsible for 1 - 3 million deaths each year or 3000 deaths per day (Greenwood et al., 2005; Greenbaum et al., 2002; WHO, 2005). In many of the poorest countries of the world, malaria is considered to be a major hindrance to economic development. The costs of malaria control and medical care are estimated to 12 billion US\$ every year for African countries alone (Sachs et al., 2002). The disease is caused by protozoan parasites of the genus *Plasmodium*, four species of which are capable of infecting humans (*P. falciparum*, *P. vivax*, *P. malariae*, and *P. ovale*).



**Fig 1.1:** Distribution of malaria worldwide. Note that the tropical and subtropical regions are at high risk. This map was taken from WHO-UNICEF World Malaria Report 2005.



**1.2 *Plasmodium falciparum* malaria** - The most serious form of the disease is caused by *Plasmodium falciparum*, which account for most of the deaths, especially in children and pregnant women (Greenwood et al., 2005). The parasite is transmitted by the female *Anopheles* mosquito. The life-cycle of the parasites within the human host can be divided into a liver and a blood stage (Fig. 1.2).

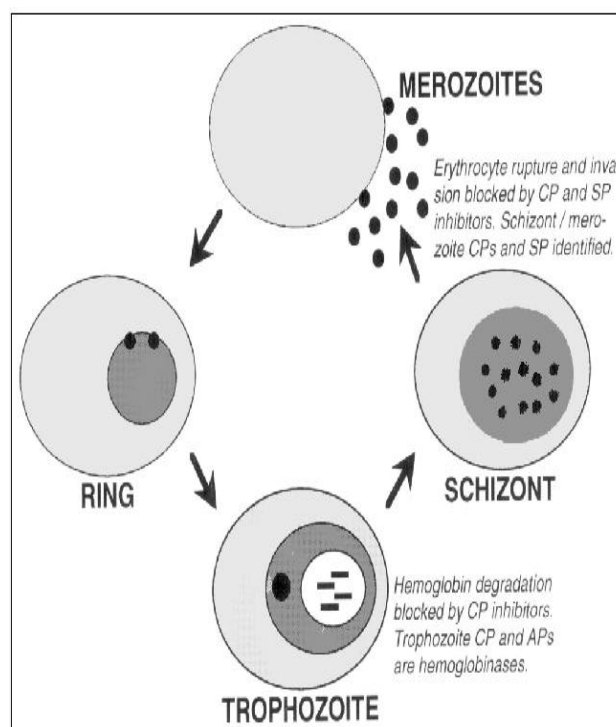


**Fig 1.2:** Diagram showing the life cycle of *Plasmodium* infection (Jones and Hoffman, 1994).

Sporozoites injected into the human blood stream are transported into the liver and penetrate hepatocytes. After multiplying asexually for a period of 6 - 15 days, sporozoites mature into schizonts that differentiate into thousands of merozoites and rupture their host cell's membrane to escape into the blood. In the case of *P. vivax* and *P. ovale*, some of the sporozoites do not start to differentiate in the hepatocytes immediately but enter a resting or dormant phase (hypnozoites) which lasts for days

or even months. Once reactivated, hypnozoites can cause a clinical relapse several months after the initial infection (Greenwood et al., 2005). During the pre-erythrocytic stage, the parasitic infection proceeds without noticeable symptoms.

**1.2.1 Erythrocyte life cycle of *Plasmodium falciparum*** - In the blood stream, merozoites penetrate and infect red blood cells, thus beginning the erythrocytic stage of the life cycle (Fig 1.3).

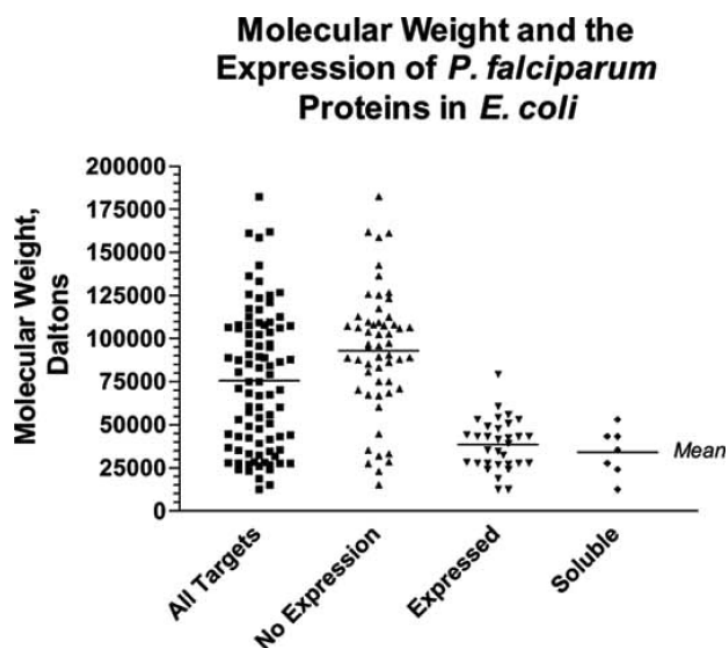


**Fig 1.3:** *Plasmodium falciparum* erythrocyte life cycle (Rosenthal, 1998). The cysteine (CP), serine (SP), and aspartic (AP) proteases involved in the process are given in *italics*.

Within erythrocytes, the merozoites grow through feeding on the cell contents and develop into ring stage and trophozoitic stage parasites, which in turn develop into erythrocytic schizonts. Mature erythrocytic schizonts divide into 6 - 32 merozoites which are released into the blood stream by rupturing the host-cell membrane. The infection proceeds in a repeated cycle of merozoite penetration of erythrocytes, intracellular development, multiplication, and escape from the host cell. Normally, a variable number of cycles occur, with each cycle lasting approximately 48 hours. A

small number of merozoites develop into male and female gametocytes. Gametocytes taken up by an Anopheles mosquito can enter the so-called sporogonic cycle leading to mature sporocytes, which are then able to infect another human host during the next blood meal of the mosquito (Cowman and Crabb, 2006).

**1.2.2 Overview of *P. falciparum* proteins** - The sequencing of *P. falciparum* genome revealed over 5,400 genes, of which 60% encode proteins of unknown function (Bozdech et al., 2003). Insights into the biochemical function and regulation of these genes will provide the foundation for future drug and vaccine development efforts toward eradication of this disease. The soluble expression rate of *P. falciparum* genes encoding the proteins has reported to be just about 5% out of 368 targets (Fig. 1.4). This is due to the fact that most of these plasmids encoding these proteins are either being insoluble or unexpressed (Mehlin, 2005).



**Fig 1.4:** Molecular weight of all targets at each stage of the process. It can be seen that the genes encoding larger proteins tend not to express at all while the genes encoding smaller proteins may express but tend to do so insolubly (Mehlin, 2005).

Furthermore, 90 *P. falciparum* proteins were targeted in the Structural Genomics of Pathogenic Protozoa pipeline (Mehlin, 2005), out of which structural information is available only for 21 different protein targets as shown in Table. 1.1.

PlasmoDB Name [127]	Function	Number of Structures		Pathway
		Total	w/inhibitor	
PF13_0141	L-lactate dehydrogenase	8	7	glycolysis
PF14_0077	Plasmeprin 2	7	5	heme digestion
PF14_0378	triose-phosphate isomerase	5	4	glycolysis
MAL6P1.275	enoyl-acyl carrier reductase	5	3	fatty acid synthesis
PFD0830w	dihydrofolate reductase-thymidylate synthase	3	3	folate synthesis
MAL13P1.279	PK5	4	2	cell cycle
PFE0660c	Purine nucleotide phosphorylase	3	2	purine salvage
PFC0975c	cyclophilin (PFCYP19)	2	2	unknown
PF14_0187	Glutathione S-transferase	3	1	detoxification
PF13_0287	adenylosuccinate synthetase	1	1	purine salvage
PFI0380c	formylmethionine deformylase	1	1	protein processing
PF10_0121	hypoxanthine phosphoribosyltransferase	1	1	purine salvage
PF14-0075	Plasmeprin IV	1	1	heme digestion
PF14_0749	acyl CoA binding protein	1	0	fatty acid metabolism
MAL13P1.95	ferredoxin	1	0	unknown
PF14_0425	fructose-bisphosphate aldolase	1	0	glycolysis
PF13_011	Gamete antigen 27/25	1	0	unknown
PF14_0192	Glutathione reductase	1	0	detoxification, redox
PF1105w	phosphoglycerate kinase	1	0	glycolysis
PF11_0461	Rab6	1	0	vesicular transport
PF14_0545	thioredoxin	1	0	redox maintenance

**Table 1.1:** *P. falciparum* soluble (non-surface) proteins with three-dimensional structure (Mehlin, 2005).

**1.3 Established antimalarial drugs and their limitations** - Drugs against malaria infection have been in use for centuries. Quinine was widely used in the 17<sup>th</sup> and 18<sup>th</sup> centuries and is still the main stay for treating severe malaria in many countries due to the low incidence of adverse effects (Alkadi, 2007). However, the spreading of resistance and the occurrence of some side effects in patients with G6PDH deficiency unfortunately limits the application of quinine (Panisko et al., 1990; Alkadi, 2007).

Extensive research and screening of several compounds led to the discovery of chloroquine. Due to its low cost, wide availability, minimal toxicity and rapid eradication of parasites, chloroquine was the most efficacious and safest drug

of all antimalarials for the past 50 years (Rosenthal, 1998). In the last two decades, resistance to chloroquine has become wide-spread in most *P. falciparum* endemic areas and thus the use of chloroquine alone for treating malaria is no longer effective (Olliaro et al., 1996). Resistance of *P. vivax* to chloroquine has also been reported in parts of Southeast Asia (Murphy et al., 1993). Recently, Laufer and co-workers (Laufer et al., 2006) reported that chloroquine is once again effective against uncomplicated *falciparum* malaria in Malawi, but advised against chloroquine monotherapy as the study is confined to a small area and there is high chance for reemergence of chloroquine resistance.

Drug	Role	Best Feature(s)	Limitations
Chloroquine	TX of and CP against non-Pf and sensitive Pf parasites	Very safe; low cost; long half-life	Widespread R
Quinine/quinidine	Best TX for Pf malaria; low cost	Limited R; rapidly acting	Fairly toxic (cinchonism, cardiac)
Amodiaquine <sup>b</sup>	TX of R Pf malaria	Low cost	Toxicity (bone marrow, liver); R common
Mefloquine	CP against R Pf malaria; not approved for TX in United States	Relatively little R, though increasing; long half-life	Moderately toxic (mostly CNS); high cost; R in SE Asia
Fansidar	TX of Pf malaria; no longer recommended for CP	Relatively low cost; long half-life	Skin toxicity (can be fatal); increasing R
Primaquine	Eradication of chronic liver stage Pv, Po malaria	Only drug for this indication	Hemolysis with G6PD deficiency; increasing R
Proguanil <sup>b</sup>	CP only (often with chloroquine)	Low cost; nontoxic	R common
Maloprim <sup>b</sup>	CP only (often with chloroquine)	Low cost	R common; skin rashes
Tetracyclines	CP; TX of Pf malaria in combination with quinine	Low cost	Skin and gastrointestinal toxicity

<sup>a</sup>TX, therapy; CP, chemoprophylaxis; R, resistance/resistant; Pf, *Plasmodium falciparum*; Pv, *P. vivax*; Po, *P. ovale*; CNS, central nervous system; G6PD, glucose 6-phosphate dehydrogenase.

<sup>b</sup>Not available in the United States.

**Table 1.2:** Role and limitations of some established antimalarials (Rosenthal, 1998).

The other major anti-malarial drugs also have their own limitations (Table. 1.2). Amodiaquine and mefloquine, widely used to treat chloroquine-resistant malaria, are fairly toxic (Rosenthal, 1998). In some parts of Southeast Asia, especially in Myanmar, Thailand, and Cambodia, 10 - 15% of the cases do not respond to mefloquine (Baird et al., 2002). Fansidar, a drug combination of sulfadoxine and

pyrimethamine is also known to cause severe side effects (Rosenthal, 1998). Halofantrine, a relatively new drug, is unsuitable for widespread use because of its potential for cardiotoxicity (Rosenthal, 1998). Of the currently available antimalarials, a combination of atovaquone and proguanil functions with good efficacy and tolerability in prophylaxis and treatment (Taylor & White, 2004; Alkadi, 2007). Artemisinin and its derivatives are the most potent and rapidly acting antimalarial drugs that could have a major impact on malaria control. It has been shown that artemisinin is effective against chloroquine-resistant *P. falciparum* and has remarkable efficacy and an excellent safety record (Baird, 2005). However, recently mutations that could lead to artemisinin resistance have been reported (Jambou et al., 2005). At present, combining an artemisinin derivative with another efficacious antimalarial drug is increasingly being viewed as the state of art in malaria therapy (Taylor and White, 2004). Since the use of currently available antimalarial agents is restricted by toxicity, drug resistance, and high costs, there is an urgent need for new drugs, ideally against new targets. Potential and so far relatively unexplored targets for chemotherapy include malarial proteases.

**1.4 Malarial proteases: A new target for chemotherapy** - Malarial proteases play a major role in both the invasion of erythrocytes and in the release of merozoites. Recent studies illustrate that during invasion, malarial proteases cleave proteins of the erythrocyte cytoskeleton (Dua et al., 2001). Intraerythrocytic plasmodia complete their asexual life within a lysosome-like acidic organelle (pH 5.0-5.4) commonly known as the digestive food vacuole (Goldberg, 1993; Rosenthal, 2004). Intraerythrocyte parasites consume hemoglobin as a major nutrient source essential to gain free amino acids for parasite protein synthesis (Francis et al., 1997; Goldberg, 1993; Rosenthal et al., 1988; Bozdech et al., 2003; Kolakovich et al., 1997).

Furthermore, it has been suggested that excessive hemoglobin degradation is essential to maintain osmotic stability of the host cell, preventing a premature rupture of the cell membrane (Lew et al., 2003). Hemoglobin degradation is carried out predominantly in trophozoites and early schizont stages (Fig. 1.3). This massive, distinct catabolic activity involves several parasitic proteases. Plasmodial proteases responsible for Hb degradation fall into three major classes: aspartic-, cysteine-, and metallo-proteases (Goldberg, 1993; Rosenthal et al., 1988 & 2001; Vander Jagt et al., 1984; Banerjee et al., 2002; Francis et al., 2003; Sijwali et al., 2001b; Eggleston et al., 1999; Murata et al., 2003a, Murata et al., 2003b).

**1.4.1 Aspartic proteases of the food vacuole** - The genome of *P. falciparum* encodes four aspartic proteases called plasmepsins that function within the food vacuole (Banerjee et al., 2002). Plasmepsins (PMs) I and II are well characterized and believed to be required for the initial cleavage in the  $\alpha$ -chains of hemoglobin (Gluzman et al., 1994). They further reported that the cleavage site is located in a region of the  $\alpha$ -chain of hemoglobin that is crucial for the stability of the overall tetrameric structure. Cleavage at this site presumably causes the globin subunits to dissociate and partially unfold (Goldberg et al., 1991; Gluzman et al., 1994). Subsequently, additional cleavage sites within the globin polypeptide chains are exposed as a prerequisite for further efficient proteolysis by cysteine proteases (Shenai et al., 2000; Sijwali et al., 2004; Singh et al., 2006).

**1.4.2 Cysteine proteases of the food vacuole** - Four papain-like cysteine proteases have been identified in *P. falciparum*, namely falcipain-1 (FP-1), falcipain-2 (FP-2), falcipain-2' (FP-2' or FP-2B) and falcipain-3 (FP-3). Although have been the first cysteine protease to be described, the role of FP-1 in the plasmodial life cycle is still

unclear. It has been implicated in erythrocyte invasion by merozoites (Greenbaum et al., 2002). More recently, two independent gene disruption studies showed that FP-1 is not essential for hemoglobin degradation but plays a role in oocyst production (Sijwali & Rosenthal, 2004; Eski et al., 2004). The amino acid (AA) sequence identity between FP-2, FP-2' and FP-3 is high (97.5% and 66.6% in their catalytic domains respectively; Goh & Sim, 2004). All three enzymes have been cloned, overexpressed, biochemically characterized (Shenai et al., 2000; Sijwali et al., 2001a; Singh et al., 2006; Sim et al., 2001; Chan et al., 2005) and shown to be involved in degradation of hemoglobin. Gene disruption studies suggest that FP-2 and FP-3 are the major trophozoite-stage cysteine proteases and FP-2' cannot fully compensate the loss of FP-2 or FP-3 (Sijwali et al., 2006).

**1.4.3 Metalloprotease present in the food vacuole** - The metalloprotease falcilysin is likely to be involved in downstream processing of the peptide fragments generated by plasmepsins and falcipains (Eggleson et al., 1999). Falcilysin prefers peptide fragments of 10 - 20 amino acids in length and charged residues at P1 and P4' positions with polar or charged residues appearing frequently at P2 (Eggleson et al., 1999). It has been suggested that these peptides may be exported into the cytoplasm where degradation to amino acids by aminopeptidases takes place (Kolakovich et al., 1997).

**1.5 FP-2 as promising drug target** - Since hemoglobin degradation is the main source of nutrition for *Plasmodia* during their intra-erythrocytic development, interference with this process holds great promise as a new approach to chemotherapy of malaria. Due to their central function in hemoglobin catabolism, the falcipains are suitable and promising targets for development of potential



antimalarials. Availability of biochemical data for FP-2 establishes it as an important drug target (Shenai et al., 2000; Sijwali et al., 2001a; Singh et al., 2006).

**1.5.1 FP-2 biochemical properties** - FP-2 is a multi-functional enzyme: It degrades hemoglobin at acidic pH within the digestive food vacuole and cleaves the membrane-associated cytoskeletal proteins ankyrin and protein 4.1 at neutral pH (Dua et al., 2001). This activity displays a pH-optimum in the range of 7.0-7.5 and is thought to contribute to destabilization of the erythrocyte membrane, leading to host-cell rupture and release of the mature merozoites. The autoproteolytic processing of its own precursor at neutral pH has been suggested as a third function of FP-2 (Shennai et al., 2000; Dahl and Rosenthal, 2005). Recently, it has been shown that FP-2 is responsible for the proteolytic activation of PM II (Drew et al., 2008).

FP-2 is synthesized during the trophozoite stage as a membrane-bound proenzyme comprising 484 amino-acid residues (Shenai et al., 2000; Pandey et al., 2004). The proenzyme is transported to the food vacuole through the endoplasmatic reticulum/Golgi system, and during this process the N-terminal 243 residues containing the membrane anchor are proteolytically removed (Fig. 1.5). An autoproteolytic processing mechanism was suggested on the basis of inhibitor studies (Dahl and Rosenthal, 2005) and from the observation that the recombinant proenzyme undergoes spontaneous processing during *in vitro* refolding (Shenai et al., 2000; Sijwali et al., 2002).

```

      10      20      30      40      50      60
1  MDYNMDYAPH EVISQQGERF VDKYVDRKIL KNKKSLLVII SLSVLSVVGF VLFYFTPNSR 60
                                Predicted membrane anchor
61  KSDLFKNSSV ENNNDYYIIN SLLKSPNGKK FIVSKIDEAL SFYDSKKNDI NKYNEGNNNN 120

121  NADFKGLSLF KENTPSNNFI HNKDYFINFF DNKFLMNNAE HINQFYMFIF TNNKQYNSPN 180
                                M R G S H H H H H H G S G   Truncated precursor sequence
181  EMKERFQVFL QNAHKVNMHN NNKNSLYKKE LNRFPADLTTH EFKNKYISLR SSKPLKNSKY 240
      ↓ FpC285aM
241  LLDQMNYEEV IKKYRGEENF DHAAYDWRLE SGVTPVKDQK NCGSCWAFSS IGSVESQYAI 300
                                Mature protein
301  RKNKLITLSE QELVDCSFKN YGCNGGLINN AFEDMIELGG ICPDGDYPYV SDAPNICNID 360

361  RCTEKYGIKN YLSVPDNKIK EALRFLGPIS ISVAVSDDFA FYKEGIFDGE CGDELNHAVM 420

421  LVGFGMKEIV NPLTKKGEKH VYVIKNSWC QQWGERGFIN IETDESGLMR KCGIGTDAFI 480

481  PLIE

```

**Fig 1.5:** Full-length FP-2 sequence. The whole protein is 484 AA in length and has a molecular weight of 56 kDa. The truncated precursor sequence used in the expression strains is shown (coloured in red). The AA residue Gln 1 is the start of the mature protein, wild-type FP-2. The active-site mutant FpC285aM was expressed and the mature protein starts with Met 2 as its N-terminal residue and lacks Gln 1. The predicted membrane anchor sequence is shown in pale green.

It remains unclear how FP-2 can perform the diverse activities of self-processing, hemoglobin degradation and cytoskeletal degradation at different pH optima during the various stages of parasite development. Disruption of the FP-2 gene results in reduced hemoglobin degradation in the trophozoite stage and accumulation of undegraded hemoglobin within the parasite food vacuole, but seems to be compensated for by overexpression of FP-2' and FP-3 in later stages of the *Plasmodium* life cycle (Sijwali and Rosenthal, 2004). This suggests an overlapping function and the need for an effective inhibitor that simultaneously inhibits the falcipains.

**1.5.2 Uncertainty in native hemoglobin cleavage** - There is still some controversy in the literature concerning the identity of the enzyme(s) responsible for the initial cleavage of hemoglobin. Based on *in-vitro* studies using purified proteases, it has

been proposed that aspartic proteases alone are capable of degrading intact hemoglobin and FP-2 cleaves hemoglobin only secondarily, suggesting a degradation pathway initiated by PM I (Goldberg et al., 1991; Gluzman et al., 1994). On the contrary, studies by Salas et al., 1995, showed that FP-2 is capable of degrading native hemoglobin. Further studies on FP-2 (Shenai et al., 2000), and FP-3 (Sijwali et al., 2001b) proved the ability of these enzymes to digest native hemoglobin. Previous studies by Gamboa de Domínguez and Rosenthal. (1996) demonstrated that cysteine protease inhibitors but not the aspartic proteinase inhibitor pepstatin cause a swelling of the food vacuole due to the accumulation of unprocessed hemoglobin, thereby blocking further development of the malarial parasites. This suggests an important role of FP-2 in initial degradation of hemoglobin, release of heme, and further hydrolysis of globin.

**1.5.3 FP-2 : Hemoglobin binding models** - Due to the unavailability of complex structural data two different, hemoglobin : FP-2 binding models have been proposed. Pandey et al. (2005) reported that FP-2 lacking ten amino-acid residues (185 - 194) in the C-terminal region of the catalytic domain is incapable of degrading hemoglobin. It has been proposed that FP-2 captures hemoglobin via this unique loop or hairpin prior to the subsequent cleavage of the substrate (Pandey et al., 2005). Details of the complex formation between FP-2 and hemoglobin are not well understood, leading to contrasting models with stoichiometric ratios of FP-2 to hemoglobin of 4:1 and 2:1, i.e. four or two molecules of FP-2 bind to each hemoglobin tetramer respectively (Pandey et al., 2005, Wang et al., 2006).

**1.6 Scope and objectives of this work** - Several previous reports established the function of FP-2 in hemoglobin degradation and the importance of this process as an essential requirement for parasite growth and development (Goldberg et al., 1991, Gluzman et al., 1994, Shenai et al., 2000). Interference with hemoglobin degradation holds great promise as possible mode of action for future anti-plasmodial drugs (Lecaille et al., 2002; Leung-Toung et al., 2006). Towards this goal, a structure-based drug design approach is undertaken for better understanding of the enzyme's functional properties. The X-ray structure of FP-2, crystallized in the presence of the irreversible cysteine-protease inhibitor iodoacetamide, at 3.1 Å resolution is presented. Since this work is first of its kind to address the three-dimensional structure of any cysteine protease in the plasmodial family, the structure reveals unique features of the FP-2 subfamily of cysteine proteases. The four copies of the enzyme in the crystallographic asymmetric unit provide individual snapshots of the dynamic and adaptable unique motif and the flexible amino-terminal helical extension. Proteolytic processing of the FP-2 precursor *in trans* at a neutral to weakly alkaline pH, and the binding of hemoglobin to FP-2 is dependent on the pH of the medium and the oxidation state of the heme iron are discussed. Taking these results into account, methemoglobin is proposed as the preferred substrate for FP-2 within the acidic food vacuole. Based on the crystal structure, various mutants of FP-2 were constructed. Biochemical studies with the mutants question the involvement of the unique loop or hairpin in hemoglobin capturing. Detailed study of degradation and cleavage specificity of FP-2 with myoglobin address the mechanism of FP-2 binding to hemoglobin and the systematic release of the heme group. Finally, the studies with natural and artificial inhibitors pave way for new approaches for the design of novel antimalarials.

## 2. MATERIALS AND METHODS

### 2.1 Materials -

**2.1.1 Equipments** - A list of laboratory instruments and devices used for this work is shown in Table 2.1.

**Table 2.1:** General laboratory devices and their manufacturers

Equipment	Manufacturer
<b>Protein purification/concentration/analysis</b>	
Analytical balance - Sartorius BP 210 D	Sartorius (Göttingen)
Cell shaker - Clim-O-shake (IRC-1-V)	Adolf Kühner (Switzerland)
Centrifuge - Heraeus Biofuge plus	Heraeus Instruments (Hanau)
Centrifuge - RC3B	Thermo Electron (Freiburg)
Centrifuge - Sorvall RC-5B	DuPont Instruments (USA)
Concentration	Millipore (Eschborn)
FPLC systems - Äkta Prime Plus	GE Healthcare (Freiburg)
French pressure cell - B15	Thermo IEC (Karlsruhe)
Gel electrophoresis system	GE Healthcare (Freiburg)
Mass spectrophotometer (MALDI-TOF) - Biflex II	Bruker (Karlsruhe)
pH meter - TR 156 Schott	Schott (Mainz)
Sonicator - Branson	Branson (USA)
Spectrophotometer - Cary 50 UV/Vis	Varian (Darmstadt)
Surface plasmon resonance - Biacore 3000	Biacore (Freiburg)
Table balance - Mettler PB3002	Mettler (Giessen)
Ultracentrifuge - Centrikon T-2070	Centrikon (Neufarn)
Water purification system - Milli-Qplus 185	Millipore (Eschborn)
<b>Crystallization</b>	
Microscope - Leica MZ8 binocular	Leica (Solms)

## MATERIALS AND METHODS

Microscope - Olympus SZX12 binocular	Olympus (Hamburg)
<b>Data collection</b>	
Air stream cooler – ONYX	Oxford Cryosystems (UK)
Cryostat- Oxford controller 600 series	Oxford Cryosystems (UK)
Image plate detector - ONYX	Oxford Diffraction (UK)
X-ray generator - Xcalibur PX Ultra	Oxford Diffraction (UK)
Synchrotron - X13 Beamline (DESY)	DESY (Hamburg)
Image plate detector - Mar 165	Mar Research (Hamburg)
<b>Computing</b>	
DELL work station - Precision	DELL (Frankfurt am Main)
DELL work station - Trinitron	DELL (Frankfurt am Main)

**2.1.2 Chemicals** - A list of chemicals used for this work is given in Table 2.2 below.

**Table 2.2:** Chemical items and their manufacturers

<b><i>Chemical item</i></b>	<b><i>Manufacturer*</i></b>	<b><i>Chemical item</i></b>	<b><i>Manufacturer*</i></b>
Acetonitrile	Millipore	Lithium sulfate	Fluka
Ammonium sulfate	Gerbu	Methanol	Merck
Crystal screen	Hampton	NaCl	Roth
Deionized water	Millipore	PEG 3350, 10K	Fluka
DMSO	Merck	SDS	Serva
DTT	Gerbu	Sodium acetate	Merck
Ethylene glycol	Merck	TCEP	Sigma
Glycerol	Gerbu	TFA	Merck
HEPES	Fluka	Tris	Gerbu
Iziti <sup>TM</sup>	Hampton	Triton X-100	Gerbu
L-Arginine	Serva	Urea	Gerbu

\* More details are given in the text

**2.1.3 Crystallization materials and cryo-tools** - Dialysis buttons and dialysis membranes, Al's oil, magnetic base crystal caps, mounted cryoloops, Crystal Screen 1 and 2, crystal storage vials, cryo canes, magnetic crystal wands, curved vial clamps

and micro tools: Hampton Research (Laguna Niguel, USA). 24-well limbro screw cap Nextal plates were purchased from Qiagen (Hilden, Germany). Other basic chemicals were obtained from Sigma-Aldrich Chemie GmbH (Munich, Germany) and Fluka Chemie GmbH (Buchs, Switzerland).

**2.1.4 Buffers and solutions** - All buffers and solutions were prepared using deionised water at 20°C (Millipore water purification system). Stock solutions of buffers were filtered and sterilized using 0.20 µm nonpyrogenic filters (Millipore), and kept at 4°C and 20°C. High-concentration PEG stock solutions were not filtered because of the high viscosity.

## **2.2 Methods** -

**2.2.1 Large-scale expression, purification and refolding** - The gene constructs encoding the *P. falciparum* FP-2 wild-type protein, inactive FP-2 (FPc285aM - active site Cys42Ala mutant), the FP-2 $\Delta$ 185-194 unique motif-deletion mutant, and the FP-2 $\Delta$ KK/G mutant were expressed and purified. Bacteria producing the wild-type protein, inactive FP-2 (FPc285aM), FP-2 $\Delta$ 185-194, and FP-2 $\Delta$ KK/G mutant (FP-2 lysine deletion and lysine/glycine mutant, where lysine 192 has been deleted and lysine 193 has been mutated to a glycine (K193G)) were grown in 2 - 8 L batches at 37°C on YT- or LB-medium containing 100 µg/ml carbenicillin and 25 µg/ml kanamycin. Recombinant gene expression was induced by addition of 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) to the liquid cultures at an optical density of 0.6 - 0.7 (measured at 546 nm). The cultures were incubated for another 3 h at 37°C and the cells harvested by centrifugation. The cells were washed with 1/10<sup>th</sup> of the culture volume of cold 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, resuspended in 1/100<sup>th</sup> of the culture volume of the same buffer and either stored at

## MATERIALS AND METHODS

–20°C, or immediately broken by sonication. Recovery of recombinant wild-type FP-2 and the mutants, all of which were produced in form of inclusion bodies, was achieved by collecting the insoluble fraction from the cell lysate after centrifugation for 20 min at 20,000 x *g*. The inclusion bodies were washed several times with various wash buffers, redissolved, and the recombinant protein was purified by Ni-NTA chromatography essentially as described previously (Sijwali et al., 2001a). Purified protein was refolded as described. Protein concentration in the Ni-NTA elution fractions was estimated using a theoretical extinction coefficient of 44,640 M<sup>-1</sup>·cm<sup>-1</sup> at 279 nm (Gasteiger et al., 2003) and the ratio of protein solution to refolding buffer was reduced to 1:50. The refolded protein was desalted on a HighTrap™ column (GE Healthcare) equilibrated with 20 mM Tris/HCl, pH 7.5, and eluted with the same buffer. The desalted protein was further purified on a Porous 20HQ 4.6/10 anion exchange column (PerSeptive Biosystems) equilibrated with 20 mM Tris/HCl pH 7.5 and eluted with a gradient of 0 - 500 mM NaCl.

**Table 2.3:** Expression and Purification of various constructs of FP-2

<b>1. Expression</b>	
- Expression Plasmid: Origami M15	
- 8 x 1litre pre-warmed YT medium (100 µg/ml carbenicillin; 25 µg/ml kanamycin) inoculated with over night culture (1:20), 37°C, 180 rpm	
- induction with 0.5 mM IPTG at OD 546 = 0.7-0.8 (Freshly prepared 1.90 g/ 32 ml H <sub>2</sub> O and 8 ml for each 2 litres culture)	
- Expression at 37°C for 4 hrs	
- pelleted at 5020 x <i>g</i> for 20 min, resuspend pellets in 200 ml 50 mM Tris, 1 mM EDTA pH 7.5 and again pelleted at 5020 x <i>g</i> for 20 min. Once again resuspend pellets in 40 ml, 50 mM Tris, 1 mM EDTA pH 7.5 and stored @ -20°C	
<b>2. Cell Lysis</b>	
- lyse cells by sonification for 6 min (20 sec pulse, 20 sec cooling), power level 60	
- pellet inclusion bodies, 30 min, ~27000 x <i>g</i> (SS-34: 18000 rpm), 4°C	
<b>3. Washing of Inclusion Bodies</b>	
<b>Buffers:</b>	
<b>Buffer A</b>	(2 M urea, 2.5% Tx-100, 20 mM Tris pH 8.0), 2x washing



<b>Buffer B</b>	(20% sucrose, 20 mM Tris pH 8.0), 2x washing
<b>Buffer C</b>	(6M guanidine-HCL, 20 mM Tris, 500 mM NaCl, 10 mM Imidazole pH 8.0)
<i>Buffer C is also the nickel-chromatography start buffer</i>	
<b>Buffer D</b>	(8 M urea, 20 mM Tris, 1 M Imidazole pH 8.0), elution Buffer
<b>Method:</b>	
- remove soft, upper layer which mostly contains membranes.	
- resuspend inclusion bodies by using a homogenizer in 15 ml <b>Buffer A</b> and add up the volume to 40 ml	
- centrifugation for 45 min, 18000 rpm, 4°C	
- discard supernatant and resuspend inclusion bodies by using a homogenizer in 15 ml <b>Buffer A</b> and add up the volume to 40 ml	
- centrifugation for 45 min, 18000 rpm, 4°C	
- discard supernatant and resuspend inclusion bodies by using a homogenizer in 15 ml <b>Buffer B</b> and add up the volume to 40 ml	
- centrifugation for 45 min, 18000 rpm, 4°C	
- discard supernatant and resuspend inclusion bodies by using a homogenizer in 15 ml <b>Buffer B</b> and add up the volume to 40 ml	
- centrifugation for 45 min, 18000 rpm, 4°C	
- discard supernatant and resuspend inclusion bodies by using a homogenizer in 15 ml <b>Buffer B</b> and add up the volume to 40 ml	
- centrifugation for 45 min, 18000 rpm, 4°C	
- discard supernatant and resuspend inclusion bodies by using a homogenizer in 15 ml <b>Buffer C</b> and constantly stir at 4°C for 1 hr	
- centrifugation for 45 min, 18000 rpm, 4°C	
- supernatant filtered through 0.22 µm filter and loaded directly to nickel column	
<b>4. Column Purification</b>	
<i>Load the protein to the nickel-column (fraction size 1.5 ml and flow speed 1.5 ml) and once the conductivity line reaches 0.04 millisiemens (mS), the protein is eluted by increasing steps of linear gradient.</i>	
- <b>Buffer D</b> 0 - 3% volume 15 ml	
- <b>Buffer D</b> 3% - 30% volume 30 ml	
- <b>Buffer D</b> 30% - 60% volume 30 ml	

- <b>Buffer D</b> 60% - 100% volume 10 ml (mostly the protein is eluted in the previous gradient steps itself)	
<i>Small amount of supernatant from the washing steps and fractions from the Ni-column containing the FP-2 protein were loaded to 15% SDS gel for verification.</i>	
<i>The protein samples were pooled and the concentration measured. Once the concentration of the protein reaches 6 mg/ml, the sample is divided into two or three separate batches, each of a volume 10 ml. At high concentrations the protein tends to precipitate during refolding.</i>	
<b>5. Solubilization and Refolding</b>	
- Add the protein solution drop by drop to ice cold stirring refolding <b>Buffer E</b>	
- refolding by 1/50 dilution in <b>Buffer E</b>	100 mM Tris
	30% glycerol
	250 mM L-arginine
	1 mM EDTA
	1 mM GSH
	1 mM GSSG
	pH 9.0
- incubated for 20 hrs at 4°C without stirring	
- precipitated protein removed by filtering (0.22 µm)	
- concentrated to 25 - 30 ml with Amicon concentrator (GE Amersham), 10 kDa cut-off, room temperature	
(solution clear, no visible precipitation)	
- 1/10 dilution in Millipore H <sub>2</sub> O	
- concentrated to 2 - 3 ml with Amicon, 10 kDa cut-off, room temperature,	
(solution clear, no visible precipitation)	
<i>The same principle (Step 5) was repeated for the other batches. Once verified by SDS PAGE, the protein batches were pooled. To verify whether the wild-type FP-2 is active or not and for further activity tests, natural and artificial substrates were used.</i>	

6. Ion Exchange Chromatography	
<i>In this step, for purification of the wild-type FP-2 protein, 10 mM iodoacetamide was included in the desalting and ion-exchange buffers to inhibit the activity.</i>	
- protein sample loaded on 5 ml HighTrap (GE Healthcare) desalting column, equilibrated and eluted in 20 mM Tris/HCl, pH 7.5 (flow and collection rate of 2 ml/min)	
- desalted protein was further purified on a Porous 20HQ 4.6/10 anion exchange column (PerSeptive Biosystems)	
- eluted with linear gradient of <b>Buffer F/G</b> (0-500 mM NaCl) at a flow and collection rate of 2 ml/min	
<b>Buffer F</b>	<b>Buffer G</b>
20 mM Tris pH 7.5	20 mM Tris pH 7.5
50 mM NaCl	500 mM NaCl (Protein elutes around 250 mM NaCl)
- SDS-gel with fractions (16 µl/lane)	
- fractions containing FP-2 are pooled	
- concentrated according to the yield, Amicon concentrator 10 kDa cut-off, room temperature	

Expression and purification of the following FP-2 mutants (FP-2 $\Delta$ 185-194, FP-2 $\Delta$ KK/G) were carried out in a similar fashion as the wild-type FP-2. The procedure for the mature, inactive FP-2 mutant (FPc285aM) was the same as described above; however, since this protein could not be purified by metal-affinity chromatography (as it does not contain a His-Tag), the inclusion bodies were more intensely washed: twice with buffer A (2 M urea, 2.5% (w/v) Triton X-100, 20 mM Tris/HCl, pH 8.0) and twice with buffer B (20% (w/v) sucrose, 20 mM Tris/HCl, pH 8.0). After each centrifugation step, the inclusion bodies were resuspended by sonication. Inclusion bodies from a 6 L culture were resuspended in 5 ml of buffer B plus 10 mM MgCl<sub>2</sub> and 125 U benzonase (Merck). The suspension was stirred overnight at 4°C followed by addition of 25 ml buffer B, centrifugation, and solubilization in 17 ml of denaturing

## MATERIALS AND METHODS

buffer D (8 M urea, 1 M imidazole, 20 mM Tris/HCl pH 8.0). Protein concentration was measured using a theoretical extinction coefficient of  $40,605 \text{ M}^{-1}\cdot\text{cm}^{-1}$  at  $\lambda = 279 \text{ nm}$  (Gasteiger et al., 2003). This solution was either stored at  $-20^{\circ}\text{C}$  or directly subjected to refolding and ion-exchange purification.

**2.2.2 General activity assay** - The chromogenic substrate Z-Phe-Arg-pNA (Bachem AG) was used for routine activity assays. The assay mixture contained the protein sample in 50  $\mu\text{M}$  Z-Phe-Arg-pNA, 10 mM DTT, and 100 mM sodium acetate (pH 5.5) in a final volume of 700  $\mu\text{l}$ . The absorbance at 405 nm was measured against a blank consisting of protein-free assay mixture. The activity was calculated using an extinction coefficient of  $10,500 \text{ M}^{-1}\cdot\text{cm}^{-1}$ . The pH-dependence of the reaction was tested using Bis-Tris-propane/HCl buffer at pH values 7.0 and 8.0 instead of sodium acetate.

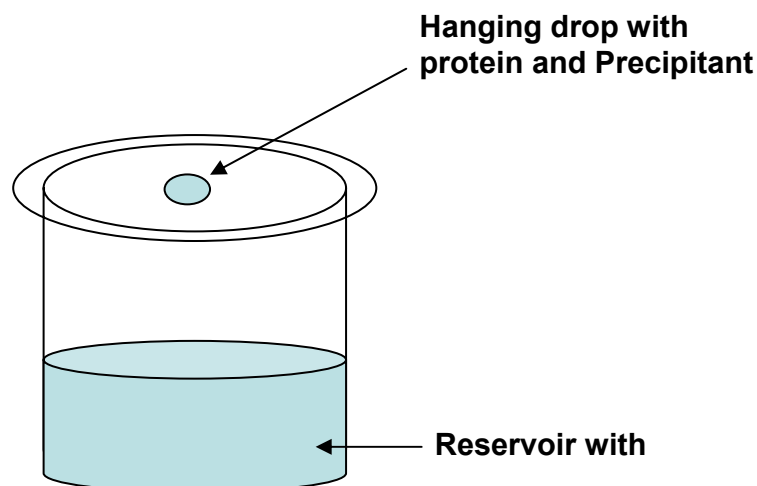
**2.2.3 Self-processing of the FP-2 prodomain** - To study the processing of inactive, truncated FP-2 precursor (FPc285aM) by active, refolded FP-2, 0.2 mg/ml of refolded FPc285aM was incubated with varying concentrations of wild-type FP-2. The assay mixture was incubated for different time intervals over a total time of 120 minutes at  $37^{\circ}\text{C}$ . 20  $\mu\text{l}$  aliquots were extracted at fixed time points, mixed immediately with 10  $\mu\text{l}$  denaturing SDS-PAGE loading buffer, heated for 5 minutes at  $95^{\circ}\text{C}$  and analyzed on a 15% Tris-Tricine gel (Schägger & von Jagow, 1987). To test the pH-dependence of the degradation, the sodium acetate buffer was replaced by 100 mM Bis-Tris-propane/HCl at higher pH values (7.0 and 8.0).

**2.2.4 Self-degradation of wild-type active mature FP-2** - Self-degrading activity was assayed by incubating active mature FP-2 with 100 mM sodium acetate (pH 5.0

and pH 6.0), and 10 mM DTT. At higher pH values, sodium acetate buffer was replaced with 100 mM Bis-Tris phosphate (pH 7.0, 8.0 & 9.0). The assay mixture was incubated for varying intervals over a total time of 120 minutes at room temperature. 20  $\mu$ l aliquots were extracted at fixed time points, mixed immediately with 10  $\mu$ l denaturing SDS-PAGE loading buffer, heated for 5 minutes at 95°C and analyzed on a 15% Tris-Tricine gel (Schägger and von Jagow, 1987).

**2.2.5 Crystallization experiments** - Protein crystallization occurs when the purified protein undergoes slow precipitation from an aqueous solution. As a result, individual protein molecules align themselves in a repeating series of "unit cells" by adopting a consistent orientation. Once the protein of interest is purified, crystals may be obtained by carefully searching for suitable crystallization conditions, which include a number of variables such as pH, buffer condition(s), temperature, precipitant(s), concentration of the protein solution etc. (McPherson, 2004). There are a number of ways to obtain crystals but the most common method to grow protein crystals is the hanging-drop vapor diffusion technique.

Crystals of FP-2 were obtained using this technique. The method is based on vapor diffusion of water. A droplet containing purified protein, buffer, and precipitant was applied to a siliconized cover slip, inverted and allowed to equilibrate with a larger reservoir (~1 ml) containing similar buffers, and precipitants in higher concentrations (Fig. 2.1). Initially, the precipitant concentration in the droplet of protein solution was insufficient to obtain crystals, but as water vaporizes from the drop and transfers to the reservoir, the precipitant concentration increases to a level optimal for crystallization. Since the system is in equilibrium, these optimum conditions are maintained until the crystallization is complete (Bergfors, 1999).



**Fig 2.1:** Schematic diagram of Hanging drop vapor diffusion technique

**2.2.5.1 Crystallization of FP-2** - Preliminary crystallization trials were conducted using Crystal Screens I and II (Hampton Research). Wild-type, mature FP-2 was crystallized at 4°C in the presence of the irreversible inhibitor iodoacetamide. Using the hanging-drop vapor diffusion method, showers of crystals were obtained in the beginning. Typically, a 2 µl aliquot of the FP-2 protein solution in 20 mM Tris, 250 mM NaCl at pH 7.5 and 10 mM of iodoacetamide was mixed with an equal volume of the reservoir solution (0.4 - 2.4 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM sodium citrate, pH 4.5 - 7.5) and allowed to equilibrate against 1 ml of reservoir solution. After further optimization, high-quality crystals were obtained by equilibrating a mixture containing 5 µl of protein solution (2 mg/ml) and an equivolume of precipitant solution (0.4 - 0.8 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM sodium citrate, pH 4.5 - 5.0) over reservoirs containing excess precipitant. Crystals grew in about a week and reached maximum dimensions of 250 x 50 x 50 µm. Crystallization plates set up using freshly purified protein gave better-quality crystals than using aged protein solution stored over a week at 4°C (about 3 weeks after purification).

**2.2.5.2 Crystallization of PbICP C-** PbICP (*Plasmodium berghei* inhibitor of cysteine proteases) was crystallized at 4°C. Typically, a 2 µl aliquot of the PbICP protein solution (20 mM Tris pH 7.5, 150 mM NaCl) was mixed with an equal volume of reservoir solutions (100 mM sodium citrate pH 5.5, 22.5% PEG 3350) and (500 mM lithium sulfate, 100 mM HEPES pH 7.5, 15% PEG 10000). The droplets were allowed to equilibrate over the same. Crystals were obtained in about 2 to 3 days. Standard protein characterization procedures like, cracking the crystals with glass rod and addition of a dye (discussed below) confirmed that they were protein crystals.

### **2.2.6 Characterization protein crystals and optimization of cryo-conditions –**

Once crystals were obtained, there were quite a few techniques to distinguish a salt crystal from protein. (i) Cracking the crystal with a very fine glass rod or glass pipette. This is the simplest and probably the best test. If it is a protein crystal, it should disintegrate very easily, whereas salt crystals are harder to break. (ii) Examining the well to see if the precipitant contains crystals. (iii) Setting up a "no protein" control drops with exactly the same conditions (buffer + additives) to see if the crystals grow without the presence of the protein. (iv) Analyzing the crystals on SDS gel (Lämmli, 1970) with silver staining or doing a western blot if an antibody is available. A 0.3 x 0.3 x 0.3 µm crystal has about 15 µg of protein in it. (v) Another common and fast technique is using a dye. Hampton's IZIT dye (C<sub>16</sub>H<sub>18</sub>ClN<sub>3</sub>S) or Methylene Green will stain protein but not salt crystals as the solvent channels in the protein are large enough to allow the dye to bind to the protein. These large solvent channels are absent in salt crystals and they stay unstained. Both the cracking (i) and use of a dye (v) method were used to identify crystals of FP-2. When poked with a sharp needle the crystals cracked easily. Similarly, the crystals turned blue by adding 1 µl of IZIT

dye to the protein drop and after a wait for 12 hrs, the crystals were coloured blue due to the binding of the dye.

Determination of optimal cryo-conditions involved many trials with conventional cryoprotectants, such as glycerol, Al's oil (Hampton Research), ethylene glycol etc. Protein plus precipitant solution containing 25% (v/v) ethylene glycol worked the best. Crystals were immediately cooled to 100 K in a nitrogen-gas stream.

**2.2.7 Diffraction data collection** - X-ray diffraction data were collected at 100 K using an incident wavelength of 0.8045 Å at the Joint University of Hamburg – University of Lübeck - EMBL Beamline X13, DESY (Hamburg). Native data were collected to a Bragg spacing of 3.1 Å (Table 3.2). The C-centered orthorhombic cell had dimensions  $a = 145.83$  Å,  $b = 168.28$  Å and  $c = 178.09$  Å. Crystal indexing and integration of reflection intensities were performed with Mosflm 6.2.4 (Steller et al., 1997; Leslie, 1992), with subsequent data scaling and merging carried out with Scala 3.2.5 of the CCP4 software suite (Evans, 1997; Collaborative Computational Project, Number 4, 1994).

**2.2.8 Structure determination and refinement** – FP-2 structure determination and refinement were performed along with Dr. Tanis Hogg (University of Lübeck). Structure determination by molecular replacement phasing was carried out using Phaser 1.3 (Storoni et al., 2004) using a 'mixed-model' (Schwarzenbacher et al., 2004) derived from residues 4-218 of the crystal structure of KDEL-tailed cysteine endopeptidase (Than et al., 2004). The correct solution, consisting of four FP-2 monomers in the crystallographic asymmetric unit (ca. 76 % solvent content), was found in space group C222<sub>1</sub> and yielded a log(likelihood) gain (The log (likelihood)



gain is the difference between the likelihood of the model and the likelihood calculated from a Wilson distribution, so it measures how much better the data can be predicted with your model than with a random distribution of the same atoms) of 927 in the 40.0 - 3.5 Å resolution range. An initial  $\sigma_A$ -weighted electron-density map (Read, 1986) calculated from the molecular-replacement phases and measured amplitudes was subjected to prime-and-switch phasing with 4-fold non-crystallographic symmetry (NCS) averaging, as implemented in Resolve 2.05 (Terwilliger, 2004), in order to reduce bias arising from the model-based phases (Prime-and-switch phasing primes density modification with biased model phases, then switches entirely to an independent source of phase information (the probability of the map) to remove the model bias). The NCS-averaged prime-and-switch electron-density map was used to construct an initial model of FP-2 with the molecular-graphics program *Coot* (Emsley and Cowtan, 2004). The initial models obtained as a result of the molecular replacement were subjected to maximum-likelihood refinement using REFMAC5 (Murshudov et al., 1997) as incorporated into the CCP4 program suite (Collaborative Computational Project, Number 4, 1994). Model improvement was accomplished by rigid-body refinement, followed by many rounds of refinement and manual model building. Initially the monomer D was built manually into the electron density map, and then the other monomers (A, B, C) were generated by NCS symmetry operations. The final rounds of crystallographic refinement were performed in CNS 1.1 (Brünger et al., 1998) using anisotropic bulk solvent correction; torsion-angle simulated annealing, positional and grouped B-factor protocols with strict 4-fold NCS constraints. The model was subjected to iterative rounds of building into  $\sigma_A$ -weighted  $3F_o - 2F_c$  and  $F_o - F_c$  maps, and crystallographic refinement with relaxed 4-fold NCS restraints was initiated at later stages once structural differences between the individual monomers were detected. Composite

simulated-annealing omit maps were used regularly during the building process in order to verify and correct the model. The final FP-2 model, which exhibits good stereochemistry (Table 3.3), consists of 962 amino-acid residues and 40 water molecules. Final  $R$  and  $R_{free}$  values are 23.1% and 27.4% for data in the 40.0-3.1 Å range (with 5% of the observed data reserved for calculation of  $R_{free}$ ).

**2.2.9 Electrostatic potential map calculations** - Continuum electrostatics play an important role in several areas of biomolecular simulation, including: (i) Simulation of diffusional processes to determine ligand-protein and protein-protein binding kinetics. (ii) Implicit-solvent molecular dynamics of biomolecules. (iii) Solvation and binding energy calculations to determine ligand-protein and protein-protein equilibrium binding constants and aid in rational drug design. (iv) Biomolecular titration studies.

Electrostatic potential maps were calculated for FP-2 at pH 5.0 using APBS tools (Baker et al., 2001). The program uses the Poisson-Boltzmann equation (PBE), one of the most popular continuum models for describing electrostatic interactions between molecular solutes in salty, aqueous media. The default values were used for grid points, solvent radius, di-electric constant, ion radius, ion concentration etc. The input PQR file (A PQR file is a PDB file with the temperature and occupancy columns replaced by columns containing the per-atom charge (Q) and radius (R)) for potential calculations was generated using the server PDB2PQR (Dolinsky et al., 2004). The server enables a user to convert PDB files into PQR files. Molecular figures were generated using the program PyMOL (DeLano, 2002).

**2.2.10 Homology modeling of various FP-2 constructs** - Homology modeling is a method used for protein structure prediction where structural data is unavailable. An atomic-resolution model of a protein is constructed from its amino acid sequence (the

## MATERIALS AND METHODS

"query sequence" or "target"). Homology modeling techniques rely on the identification of one or more closely related proteins, whose three-dimensional structures are available to the target protein. An alignment is made between the residues in the query sequence to residues in the template sequence. One or more template sequences can be in alignment for a more accurate three-dimensional model of the query protein. The sequence alignment and template structure(s) are then used to produce a structural model of the target based on the template structure(s). As protein structures are more conserved than DNA sequences, detectable levels of sequence similarity usually imply significant structural similarity (Marti-Renom et al., 2000). The validity of the homology model depends on the quality of the sequence alignment and template structure. A high-quality structural model can be produced if the target and template are closely related.

Homology models based on the wild-type FP-2 X-ray structure were created for various FP-2 constructs (FP-2 $\Delta$ 185-194, FP-2 $\Delta$ KK/G). The Swiss PDB-viewer (SPDBV v3.7) software was used to create the models (Guex and Peitsch, 1997). The sequence alignment was carried out manually using the interactive tool available within the program. The final sequence alignment was submitted to the Swiss model server (an Automated Comparative Protein Modelling Server) and the corresponding PDB file was obtained.

A similar procedure was followed for homology modeling of PblCP C-terminal. The NMR structure of the *T. cruzi* cysteine protease inhibitor Chagasin (Salmon et al., 2006) was used as a template (PDB-ID: 2FO8). The sequence alignment was carried out using ClustalX (Thompson et al., 1997) and further adjustments were done manually. The final sequence alignment was submitted to the Swiss model server and the results were obtained via their online server.

**2.2.11 Hemoglobinase and myoglobinase assay** - The same procedure as described for the self-processing of the FP-2 prodomain (Section 2.2.3) was used to study the processing of hemoglobin and myoglobin by active, refolded wild-type FP-2 and by the other mutant constructs (FPc285aM, FP-2 $\Delta$ 185-194, and FP-2 $\Delta$ KK/G). Hemoglobin (Hb)- and myoglobin (Mb)- degrading activity was assayed by incubating human hemoglobin (HHb), rat hemoglobin (RHb) and horse-skeletal muscle myoglobin (Sigma-Aldrich) with 0.02 mg/ml of various FP-2 constructs. The assay mixture consisted of 0.2 mg/ml hemoglobin or myoglobin, 100 mM sodium acetate (pH 5.5), and 10 mM DTT (Dithiothreitol). The reaction mixture was incubated with FP-2 for varying intervals over a total time of 120 minutes at 37°C. 20  $\mu$ l aliquots were extracted at fixed time points, mixed immediately with 10  $\mu$ l denaturing SDS-PAGE loading buffer, heated for 5 minutes at 95°C, and analyzed on a 15% Tris-Tricine gel (Schägger & von Jagow, 1987). To test the pH-dependent degradation activity of FP-2 on hemoglobin and myoglobin, the sodium acetate buffer was replaced by 100 mM Bis-Tris-propane/HCl at pH values 7.0 and 8.0. In one of the experiments with the FPc285aM mutant, instead of DTT, 10 mM TCEP (Tris(2-carboxyethyl)phosphine) was used as a reducing agent.

Similar to the assays with hemoglobin and myoglobin, carbonic anhydrase isoform 1 (CA1) was used as a substrate for wild-type FP-2. Purified CA1 was kindly provided by Dr. Tanis Hogg (University of Lübeck). The assay mixture consisted of 0.2 mg/ml of CA1 and 0.02 mg/ml of FP-2. The pH-dependent activity assay was carried out at pH 5.5 and 6.0 in sodium acetate buffer.

**2.2.12 Inhibition assay studies** - Hemoglobin, myoglobin degradation studies were conducted in the presence of various inhibitors. In the case of protein inhibitors such as PbICP C-terminal and CA1, 0.2 mg/ml of the inhibitors were pre-incubated with

0.02 mg/ml of wild-type FP-2 at room temperature for 15 min before the addition of the substrates. Simultaneous addition of the inhibitor and substrate to the reaction mixture was also carried out. Similar assays were also performed with inhibitor compound X (provided by Prof. Dr. François Diederich, ETH Zurich) and a set of inhibitors (kindly provided by Dr. Xu Shen, Shanghai Institute of Materia Medica). Typically, 1 mM of compound X in (100% methanol stock) was used in the assay mixture (<1% of methanol in final reaction mixture). In the case of the inhibitors from Shanghai, inhibitor concentrations of five times the  $K_d$  values were used ( $K_d$  values obtained by our collaborators, Table 3.4). All the inhibitors were pre-incubated with FP-2 for 15 min before the start of the reaction.

**2.2.13 Surface plasmon resonance (SPR)** - SPR is a phenomenon that occurs when polarized light (under conditions of total internal reflection) strikes an electrically conducting planar metal surface (typically a gold or silver layer) at the interface between media of different refractive index: the glass of a sensor surface (high refractive index) and a buffer (low refractive index). As molecules (bait ligand or protein) are immobilized on a sensor surface (SPR crystal), the refractive index at the interface between the surface and a solution (containing the prey analyte) flowing over the surface changes, altering the angle at which reduced-intensity polarized light is reflected from a supporting glass plane. The change in angle, caused by binding or dissociation of molecules from the sensor surface, is proportional to the mass of bound material and is recorded in a sensorgram. When sample is passed over the sensor surface, the sensorgram shows an increasing response as molecules interact. The response remains constant if the interaction reaches equilibrium. When sample is replaced by a buffer (typically NaOH) the response decreases as the interaction partners dissociate. Complete profiles of recognition, binding and dissociation are

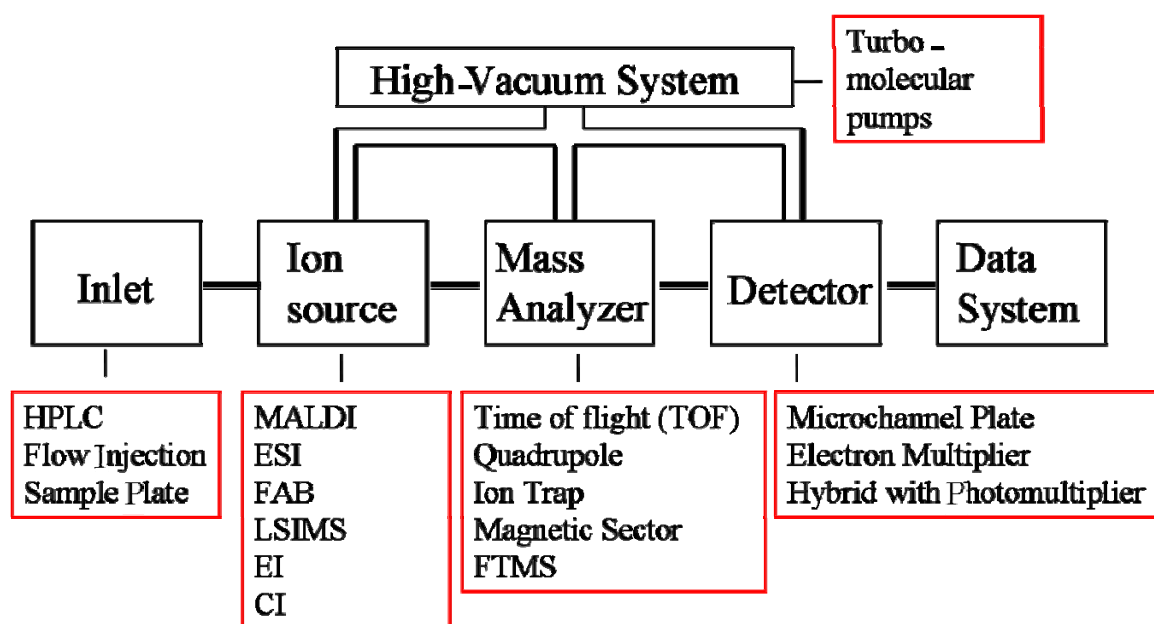
generated in real time. From these profiles, data such as specificity, affinity, kinetic behaviour and sample concentration can be determined (Aslan et al., 2005). For most applications, a dextran matrix covering the gold layer enables molecules to be immobilized to a sensor surface and provides a hydrophilic environment for interactions. Surface specificity is determined by the nature of the immobilized molecule. The main advantage of this method is that accurate results can be obtained even with low concentrations of samples. Since light does not penetrate the sample, interactions can be followed in coloured, turbid or opaque samples. No labels are required and detection is instantaneous.

**2.2.13.1 Surface plasmon resonance studies** - The binding of hemoglobin and myoglobin to mature, inactive FP-2 (FPc285aM) was studied by surface plasmon resonance using a Biacore<sup>®</sup> 3000 system at room temperature at pH values 5.0 to 8.0. Binding studies for hemoglobin were performed by our collaborators (Shanghai institute of Materia Medica) and myoglobin binding studies were conducted under the supervision of Dr. Thomas Weimar (University of Lübeck). Inactive FP-2 was immobilized to the carboxy-methylated dextran matrix of a CM5 research-grade sensor chip (Biacore-GE Healthcare) by the standard primary-amine coupling reaction in HBS-EP buffer (10 mM HEPES/NaOH, pH 7.5, 150 mM NaCl, 3 mM EDTA). The flow-cell was activated for 7 min with a 1:1 mixture of 0.2 M N-hydroxysuccinimide and 50 mM N-ethyl-N'-(3-diethylaminopropyl)-carbodiimide) at a flow rate of 5 µl/min at 25°C. Inactive, mature FP-2 (FPc285aM) at a concentration of 0.066 mg/ml in 10 mM sodium acetate (pH 3.9) was injected for 7 min at 5 µl/min, yielding a final immobilization level of 600 response units (RU). The surface of the flow-cell was subsequently blocked with a 7-min injection of 1 M ethanolamine (pH 8.5). Finally, 50 mM NaOH was injected for 10 s to remove any non-covalently bound

## MATERIALS AND METHODS

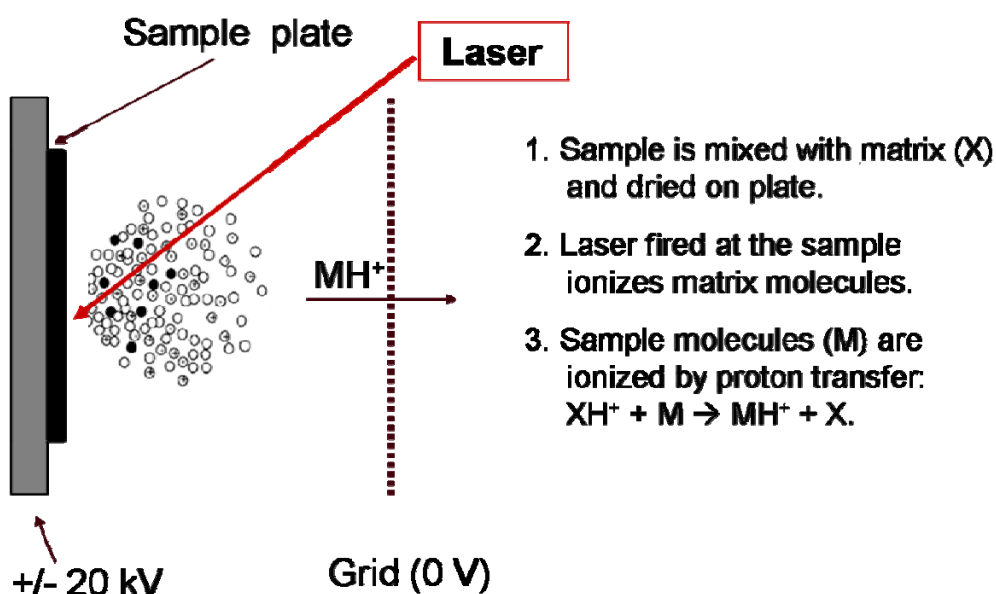
protein. The control flow-cell was activated and immediately blocked without immobilizing any protein. Equilibration of the baseline was carried out by a continuous flow of HBS-EP buffer (pH 5.0) through the chip for at least 3 h. Horse skeletal muscle myoglobin (Sigma-Aldrich) in 20 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ , 150 mM NaCl, 3 mM EDTA of the indicated pH was injected at a flow rate of 30  $\mu\text{l}/\text{min}$  for 120 seconds at 25°C. Following each myoglobin injection, the chip was regenerated by two injections of 50 mM NaOH at a flow rate of 50  $\mu\text{l}/\text{min}$  for 6 seconds. The experiment was repeated at pH values 6.0, 7.0 and 8.0. The results were analyzed using the Biacore software. Alternatively, the equilibrium dissociation constants were determined in the case of hemoglobin. This was done by plotting the intensity of the steady-state response (RU) against the hemoglobin concentration. The data were fitted to the standard formula of the absorption isotherm using the program KaleidaGraph 3.6 (Synergy Software).

**2.2.14 MALDI-TOF mass spectrometry** - Mass spectrometry (MS) is an analytical technique that identifies the chemical composition of a compound or sample based on the mass-to-charge ratio ( $m/z$ ) of charged particles. A mass spectrometer typically consists of an inlet where the samples can be loaded. Depending upon the technique, the mode of sample loading varies. Once samples are loaded, an ion source generates ions which are then analyzed with a mass analyzer and a detector captures those ions. The ion source, mass analyzer, and the detector are completely sealed and function under high vacuum conditions. Using any kind of data system (computers) the results can be viewed.



**Fig 2.2:** A typical working flow chart of mass spectrometry (MS).

Matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF) is an ionization technique used in mass spectrometry (MS). The samples are mixed with the matrix and subjected to a sample plate. Once the samples are air-dried they are loaded into the spectrometer. A laser beam is fired at the samples which are then ionized due to proton transfer.



**Fig 2.3:** Schematic diagram representing the working principle of MALDI-TOF.



MALDI-TOF is one of the most common techniques widely used today for characterization of peptides and proteins. Used in a variety of modes, MALDI-TOF MS provides information such as the molecular mass of an intact protein, peptide mass mapping from proteolytically digested fragments (usually tryptic digest), and peptide sequencing. For optimum results, frequent calibration and tuning of the MALDI-TOF mass spectrometer are necessary (Lee et al., 2002).

**2.2.14.1 MALDI-TOF analysis of substrate cleavage products** - The cleavage products of myoglobin and hemoglobin were analyzed using MALDI-TOF mass spectrometry. 8  $\mu$ l of 100% acetonitrile (ACN) and 130  $\mu$ l of 0.1% trifluoroacetic acid (TFA) were added to 15  $\mu$ l sample and the mixture loaded onto a Zip plate (C18-media immobilized; ZPC1 800 10, Millipore). After applying vacuum (20 mbar) for 5 minutes the samples were washed with 100  $\mu$ l 0.1% TFA and eluted in 60% ACN and 0.1% TFA. 2  $\mu$ l of each sample were applied to an MTP Anchor Chip 600/384 (Bruker Daltonics) and air-dried. For lower ranges of masses (< 4000), the matrix was prepared using a saturated solution of  $\alpha$ -cyano cinnamic acid in 50% ACN (Sigma-Aldrich) and for higher mass ranges (>4000), a saturated matrix solution of  $\alpha$ -cyano sinapic acid in 50% ACN (Sigma-Aldrich) was used. MALDI-TOF grade peptide and protein markers (MS-CAL-1) for calibration purposes were also purchased from Sigma-Aldrich. Depending on the experiment, either  $\alpha$ -cyano cinnamic acid or  $\alpha$ -cyano sinapic acid in 50% ACN was mixed with the samples at ratio of 1:1 and allowed again to air-dry. Mass spectroscopy was performed using a Bruker MALDI-TOF instrument (Bruker Daltonics) at BNI (Bernard Nocht Institute, Hamburg) under the supervision of Dr. Joachim Clos. The average molecular mass of the cleavage products were determined. All experiments were done in duplicate. A similar procedure was followed for the experiments with the FPc285aM mutant.

Annotation of the peaks was carried out using the software X-TOF (Bruker Daltonics) and moverZ software (Genomic Solutions). The programs, FindPept tool (Gattiker et al., 2002) and ProtParam tool (Gasteiger et al., 2003) was used to identify the peptide fragments and to map the cleavage products to the substrate protein.

**2.2.15 *NH<sub>2</sub>-terminal sequence analysis*** - Automated Edman degradation chemistry was used to determine the NH<sub>2</sub>-terminal sequence of myoglobin cleavage products. The cleavage products were separated on a 15% Tris-Tricine gel (Schägger & von Jagow, 1987) and transferred onto an Immobilon-P PVDF membrane (Millipore, Germany). The protein bands were visualized by staining with a freshly prepared solution of Coomassie Blue, followed by destaining with a fresh solution containing 15% methanol and 10% acetic acid. The membrane was washed with distilled water to remove any traces of acetic acid and air-dried. Using a sharp scalpel, the bands were carefully cut out from the membrane, labeled and subjected to NH<sub>2</sub>-terminal sequencing. The sample sequencing was performed by Dr. Arnd Petersen (Forschungszentrum Borstel).

**2.2.16 *Protein-protein docking studies*** - Docking studies for the FP-2 : myoglobin complex was performed under the supervision of Dr. Domantas Motiejunas and Dr. Rebecca Wade (EML Institute, Heidelberg). The hydrogens were added to the crystal structures of FP-2 (PDB-ID: 2GHU) and myoglobin (PDB-ID: 1YMB) using the program WHAT IF (Vriend, 1990). The electrostatic potential maps were calculated using the program UHBD (Madura et al., 1995). The output file from UHBD was subjected to rigid body docking using the in-house program SDA (Motiejunas et al., 2008). Further molecular dynamics simulations were carried out on blind docking

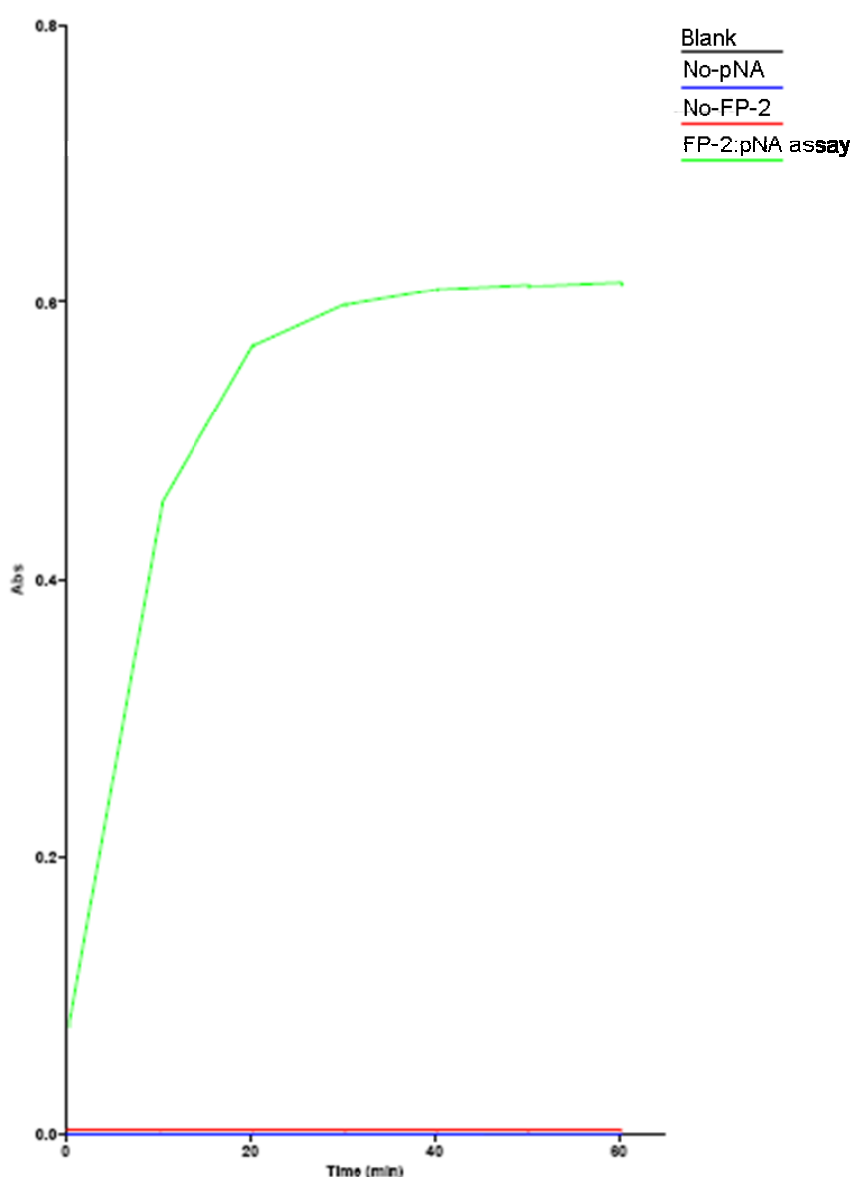
results using AMBER 9 (Ponder and Case, 2003). Visualization and generation of figures were done using PyMOL (DeLano, 2002).

**2.2.17 *PbICP* C-terminal sequence alignment** - The sequence of chagasin, the cysteine protease inhibitor of *Trypanosoma cruzi* (Salmon et al., 2006), was used for sequence alignment studies. The alignment was derived using the program ClustalX 1.83 (Thomson et al., 1997) using the identity matrix with a gap opening penalty of 10 and a gap extension penalty of 0.1 for the pair-wise comparison. Further gap opening was manually carried out to align the loop regions of the proteins.

## 3. RESULTS AND DISCUSSION

### 3.1 FP-2 general activity assays

**3.1.1 FP-2 enzyme activity determination assays** - In order to verify the activity of the enzyme, the standard artificial substrate Z-Phe-Arg-pNA was used. The UV/Vis spectrum measurement was carried out and the substrate cleavage by the enzyme was monitored at 405nm. The protein refolded in Tris buffer exhibited efficient hydrolysis of the substrate (Fig. 3.1).

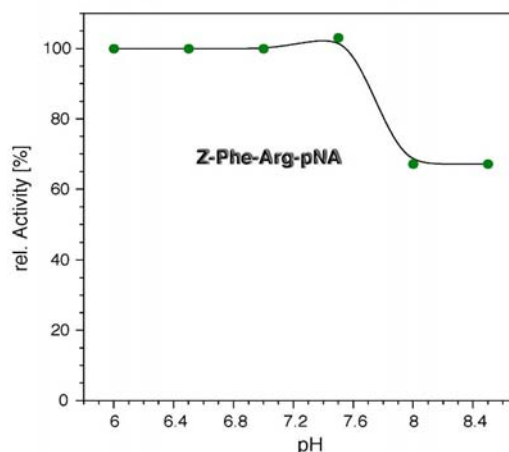


**Fig. 3.1:** FP-2 activity with the artificial substrate Z-Phe-Arg-pNA. The activity was measured over a time range of 60 min. Data points were collected after each minute for the first 25 min and later every 5 min until 60 min.

## RESULT AND DISCUSSION

Four different samples were incubated for a total of 60 min and the measurement was carried out with an interval of 1 min between each run. The samples (buffer + pNA without FP-2 (red); buffer + FP-2 without pNA (blue)) showed no significant activity. The absorbance of these samples could be compared to that of the blank sample containing just the buffer (sodium acetate, pH 5.5) which is exactly the same as the base line (black). Only the sample containing the FP-2 : pNA mixture (green) showed considerable activity (Fig. 3.1). The reaction is efficient and immediately activated, once the enzyme is in contact with the substrate. The reaction continued at a similar rate for the next 15 min and the substrate was efficiently degraded by FP-2. As Z-Phe-Arg-pNA is a chromogenic substrate, the colour change in the reaction mixture was clearly visible with time as the hydrolysis proceeded. The speed of the reaction gradually started to decline after 20 min. No further prominent activity could be observed from 20 to 60 min indicating that the reaction was more or less complete within the initial 15 min of incubation.

**3.1.2 pH-dependence of the activity with artificial substrate** - The peptidic substrate Z-Phe-Arg-pNA was incubated with FP-2 at different pH values. The substrate was cleaved with relatively high efficiency over the entire pH range studied (pH 6.0 – 8.4); however, the proteolytic activity decreased by 30% above pH 7.5 (Fig. 3.2). Even minor substrate modifications can apparently result in significantly different pH-dependent activity profiles: Shenai and colleagues (Shenai et al., 2000) observed no activity at pH 8.0 with Z-Phe-Arg-AMC (AMC=7-Amino-4-methyl coumarin) as a substrate, whereas in this work 60% of the maximal activity was still observed at this pH with Z-Phe-Arg-pNA (Fig.3.2). The exact reason for this discrepancy remains to be established. Apparently, pNA is a better leaving group than AMC at this elevated pH.

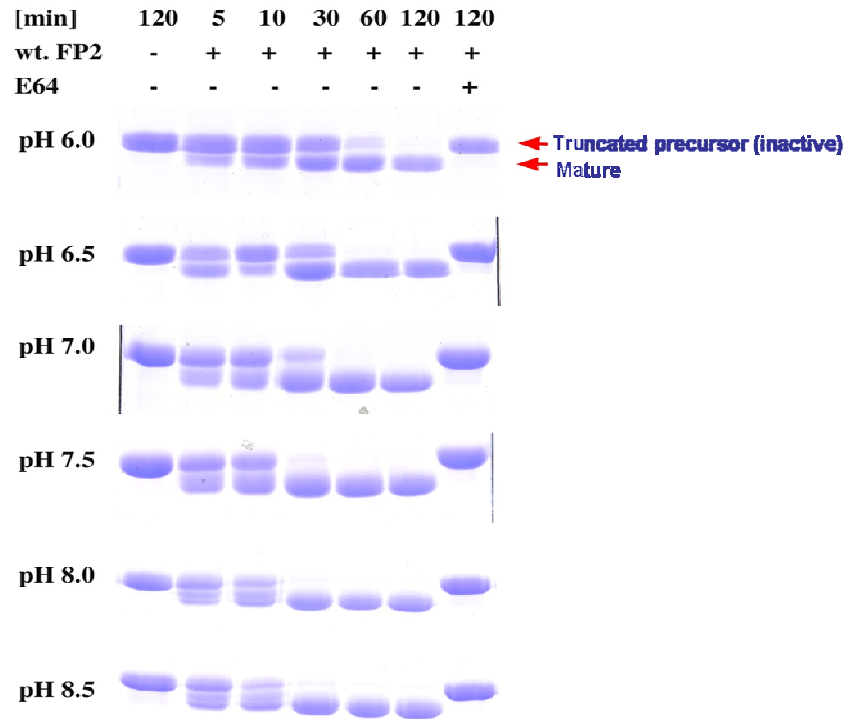


**Fig. 3.2:** pH-dependent activity of FP-2 over the pH range 6.0-8.4. The relative activity of the enzyme is plotted against the pH range.

## 3.2 FP-2 self processing assays

**3.2.1 Auto-processing of FP-2 prodomain** - The auto-processing assay shown in Fig. 3.3 demonstrates that mature, refolded wild-type FP-2 can process the inactive, truncated proenzyme *in trans*. Like the Z-Phe-Arg-pNA activity, *trans* processing was detectable over a wide pH range. The FP-2 wild-type was incubated with the inactive FPc285a at different pH values. The samples were collected at various time intervals as shown in Fig. 3.3. Maximum activity was detected at pH 8.5, where the processing could be clearly observed within 10 min of incubation. The precursor was completely processed within 30 min. The FP-2 processing activity gradually decreased with the lowering of pH. For e.g. at pH 6.0, the precursor is visible even after 30 min of incubation but significantly processed within 60 min. Overall, the reaction rate is reduced in the samples at lower pH values compared to the corresponding samples at higher pH values. This self-processing was fully sensitive to the cysteine-protease inhibitor E-64 (L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane). The samples containing E64 completely lacked activity suggesting that the processing is not due to the presence of any other non-cysteine protease impurities.

The control experiment, where no active FP-2 was present, showed no activity even after prolonged incubation for 120 min.



**Fig. 3.3:** The assay mixtures contain 0.2 mg/ml of the refolded, truncated and inactive FP-2 precursor (FPc285a). In a control experiment, FPc285a was incubated in the absence of the active FP-2 to verify the absence of protease activity in the FPc285a-preparation. The cysteine protease inhibitor E64 was added to a final concentration of 10  $\mu$ M to one of the assay mixtures to confirm that the observed processing was performed by FP-2.

Recombinant FP-2 was reported to be proteolytically processed during refolding (Shenai et al., 2000). This result shown in Fig. 3.3 demonstrate for the first time that wild-type FP-2 is capable of self-processing *in trans*, with an optimum at a weakly alkaline pH. The N-terminal sequencing of the refolded FP-2 revealed the sequence of the native mature protein. The auto-processing site of wild-type FP-2 was revealed to be YLLD<sub>-1</sub>↓Q<sub>+1</sub>MNY and the mature protein starts with a glutamine as its N-terminal residue.

Auto-processing was not observed during the refolding of the inactive, truncated FP-2 precursor (FPc285a). This result showed that the prodomain processing was not an artifact due to trace amounts of a cysteine protease being present from the expression host and further supports that only FP-2 is responsible for auto-processing. Based on inhibitor studies *in vivo*, Dahl & Rosenthal (2005)

previously concluded that FP-2 and -3 are autohydrolytically processed at a neutral pH prior to their secretion into the food vacuole.

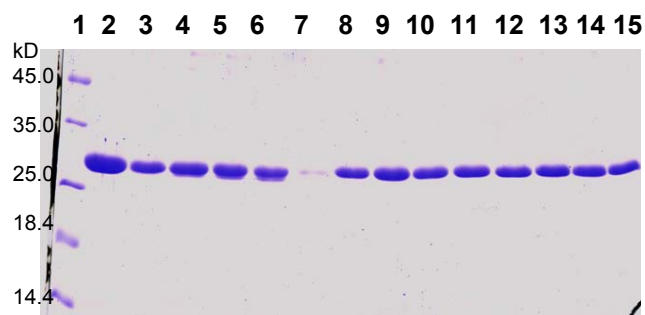
The broad pH optima of Z-Phe-Arg-pNA hydrolysis and of self-processing indicate that the active site of the protein is in a catalytically active state at least in the range from pH 4.5 (as shown in (Shenai et al., 2000)) to 8.5. Since the self-processing is still observable under acidic conditions, FP-2 may also contribute to the processing of other plasmodial proteins within the food vacuole. For instance, it has been shown that an as yet unidentified cysteine protease is involved in the processing of the plasmepsins (Francis et al., 1997; Banerjee et al., 2003). However, whose results showed that the protein in question was E-64 insensitive. This would argue against a significant contribution of FP-2 to the processing of these particular proteins. But recent studies by Drew and colleagues (Drew et al., 2008) showed that FP-2 is involved in the prodomain processing of PMII and the processing site has been revealed to be NYLG<sub>124</sub>↓S<sub>125</sub>SND.

**3.2.2 Self-degradation assay for FP-2 at pH 5.0 to 9.0** - The same procedure as described for the self-processing of prodomain was used to study the degradation of the active mature protein. The active mature wild-type FP-2 was incubated at various pH range at room temperature for various time intervals. Samples were collected at time intervals as indicated in Fig. 3.4 & 3.5. At pH 5.0 and 6.0, sodium acetate buffer and at pH 7.0 to 9.0 Bis-Tris-phosphate was used. FP-2 in H<sub>2</sub>O was used as control (ctrl) experiment in both the gels. There was no or minimal degradation observed at all the pH ranges even after 120 min of incubation, suggesting that the mature FP-2 is relatively stable over all pH ranges at room temperature. The protein sample at pH 5.0 incubated overnight (Ov.N) was exhibiting considerable self-degradation.



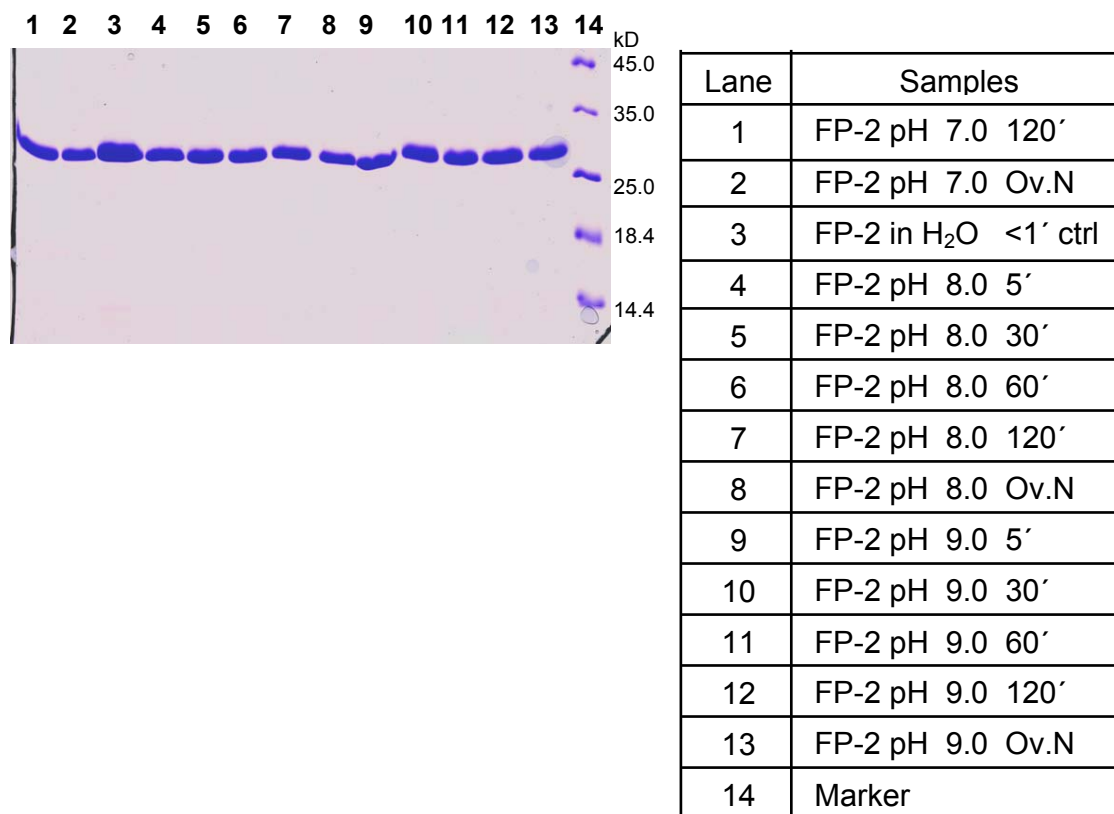
## RESULT AND DISCUSSION

However, all the other samples from various pH ranges were relatively stable and showed no visible degradation even after over night incubation.



Lane	Samples
1	Marker
2	FP-2 in H <sub>2</sub> O <1' ctrl
3	FP-2 pH 5.0 5'
4	FP-2 pH 5.0 30'
5	FP-2 pH 5.0 60'
6	FP-2 pH 5.0 120'
7	FP-2 pH 5.0 Ov.N
8	FP-2 pH 6.0 5'
9	FP-2 pH 6.0 30'
10	FP-2 pH 6.0 60'
11	FP-2 pH 6.0 120'
12	FP-2 pH 6.0 Ov.N
13	FP-2 pH 7.0 5'
14	FP-2 pH 7.0 30'
15	FP-2 pH 7.0 60'

## RESULT AND DISCUSSION



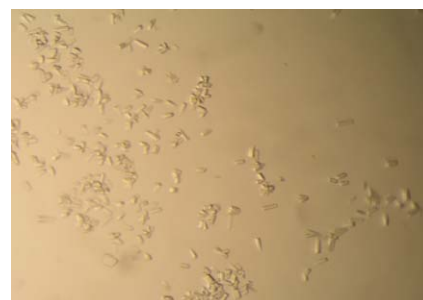
**Fig. 3.4 & 3.5:** 1.0 mg/ml purified active mature FP-2 was incubated with buffers at various pH values ranging from pH 5.0 to 9.0. The samples were incubated at room temperature for the indicated times (5' to 120'). FP-2 in H<sub>2</sub>O (<1 min sample) was used as a control experiment. Samples were subjected to SDS polyacrylamide gel electrophoresis (Shagger et al., 1987)

### **3.3 Crystallization and Structure determination of FP-2**

**3.3.1 Crystallization of recombinant wild-type FP-2** - The protein solution containing the FP-2 in buffer (section 2.2.1) was screened for potential crystallization conditions using the commercially available kits (Crystal Screens I and II from Hampton Research). The screening was carried out at 4°C, 10°C, and at room temperature. The experiments were performed using the well known hanging drop vapour diffusion method (Section 2.2.5). Purified wild-type, mature FP-2 was mixed with an equal volume (2  $\mu$ l) of precipitant solution and allowed to equilibrate over 1 ml of reservoir solution (Table 3.1). Showers of crystals were obtained in the beginning (within one or two days) using the hanging-drop vapor diffusion method. Only at 4°C, trials produced relatively less number of crystals (Fig 3.6). The crystals were extremely small and fragile. When poked with a thin needle the crystals fell apart as thin 2D plates suggesting loose packing atleast along one axis. After further optimization, rapid nucleation was absent and larger 4-5 big crystals (Fig. 3.7) were obtained by equilibrating a mixture containing 5  $\mu$ l of protein solution (2 mg/ml) and an equivolume of precipitant solution (Table 3.2) over reservoirs containing excess precipitant. Crystals grew in about a week when stored at 4°C. A precipitant solution containing 25% (v/v) ethylene glycol was used as cryo-protectant and the crystals were soaked for a short time and immediately cooled to 100 K in a nitrogen-gas stream.

**Table 3.1:** Crystallization conditions for initial trials

		$(\text{NH}_4)_2\text{SO}_4$ →					
100 mM sodium citrate ↓		0.4M	0.6M	1.2M	1.6M	2.0M	2.4M
	pH 4.5						
	pH 5.5						
	pH 8.5						
	pH 7.5						



**Fig 3.6:** Initial shower of crystals.

**Table 3.2:** Optimized crystallization conditions

		100 mM sodium citrate →					
$(\text{NH}_4)_2\text{SO}_4$ ↓		pH 4.5	pH 4.6	pH 4.7	pH 4.8	pH 4.9	pH 5.0
	0.4M						
	0.5M						
	0.6M						
	0.8M						



**Fig 3.7:** Native crystals of FP-2.

**3.3.2 X-ray diffraction data** - Crystals displayed space group  $C222_1$  with unit cell dimensions  $a = 145.83 \text{ \AA}$ ,  $b = 168.28 \text{ \AA}$ ,  $c = 178.09 \text{ \AA}$ , and four proteinase monomers in the asymmetric unit. Diffraction data were collected at the Joint University of Hamburg – University of Lübeck - EMBL Beamline X13, DESY (Hamburg) source equipped with a MAR CCD 165 mm detector and at an incident wavelength of  $0.8045 \text{ \AA}$ . Diffraction data collected gave an average value for  $I/\sigma(I)$  of 11.5 for all reflections (resolution range  $40.0$  to  $3.04 \text{ \AA}$ ) and 2.7 in the highest resolution shell ( $3.20$  to  $3.04 \text{ \AA}$ ). The data was more than 95% complete. A total of 262,326 reflections were collected, representing 40,840 unique reflections. Data processing gave  $R_{\text{merge}} = 15.2\%$ . The Matthews coefficient for 4 molecules per asymmetric unit (a.u) was  $5.02 \text{ \AA}^3/\text{Da}$  and the solvent content was 75.5% (Matthews, 1968). Diffraction data are summarized in Table 3.3.

**Table 3.3:** Summary of X-ray diffraction data

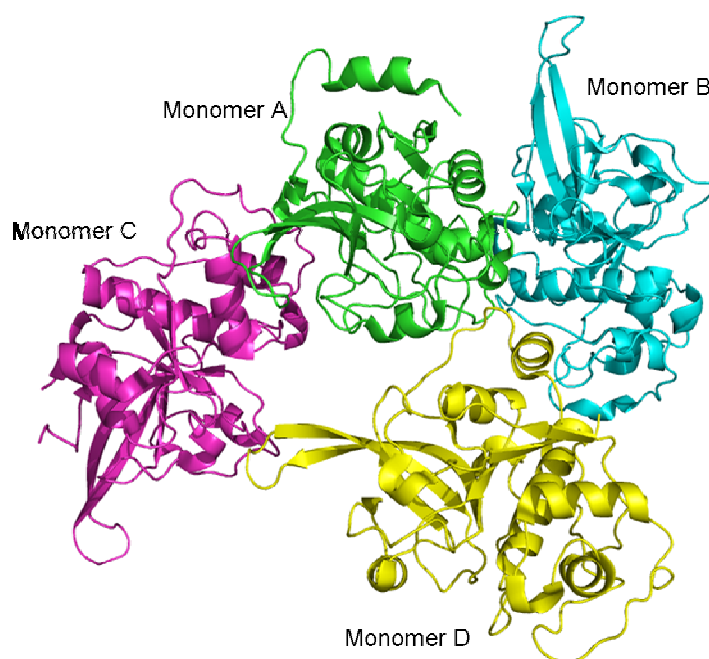
<b>Data collection</b>	
Wavelength (Å)	0.8045
Resolution (Å)	40.0-3.04 (3.20 - 3.04)
Space group	C222 <sub>1</sub>
Unit-cell Parameters	
a (Å)	145.83
b (Å)	168.28
c (Å)	178.09
Overall reflections	262,326 (23,804)
Unique reflections	40,840 (4,411)
Multiplicity	6.4 (5.4)
Completeness (%)	95.8 (71.9)
<sup>1</sup> R <sub>merge</sub> (%)	15.2 (76.4)
I/σ(I)	11.5 (2.7)
<b>Refinement</b>	
Resolution (Å)	40.0-3.1 (3.29 - 3.10)
No. reflections	39,960 (6,561)
Overall completeness	99.9 (100.0)
<sup>2</sup> R <sub>cryst</sub>	0.231 (0.397)
<sup>2</sup> R <sub>free</sub>	0.274 (0.405)
r.m.s.d. from ideal geometry	
bonds (Å)	0.009
angles (°)	1.5
Protein atoms	7,614
Solvent atoms	40
Ramachandran plot	
<sup>3</sup> most favoured (%)	75.3
<sup>3</sup> additionally allowed (%)	22.8
<sup>3</sup> generously allowed (%)	1.8
<sup>3</sup> disallowed regions (%)	0

<sup>1</sup>R<sub>merge</sub> =  $\sum_{hkl} \sum_i |I(hkl)_i - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I(hkl)_i$ , where  $I(hkl)$  is the intensity of reflection  $hkl$ , and  $\langle I(hkl) \rangle$  is the average intensity over all equivalent reflections.

<sup>2</sup>R<sub>cryst</sub> =  $\sum_{hkl} |F_o(hkl) - F_c(hkl)| / \sum_{hkl} F_o(hkl)$ .  $R_{free}$  was calculated for a test set of reflections (5%) omitted from the refinement.

<sup>3</sup>As defined by Morris and colleagues (Morris et al., 1992).

**3.3.3 Overall crystal structure** - To determine the three-dimensional structure of wild-type mature FP-2 by Molecular Replacement, the amino-acid sequence was submitted for a BLAST search (URL: <http://www.ncbi.nlm.nih.gov/BLAST>) to identify promising homologous search models. As FP-2 is a member of the C1A papain family proteases, several hits were obtained. The best hit was KDEL-tailed cysteine endopeptidase, KDEL-CysEP, from *Ricinus communis* (castor bean) with a score of 43% sequence identity. The structure was determined by the Molecular-Replacement technique using a modified set of coordinates of the KDEL-tailed cysteine endopeptidase (Than et al., 2004). The crystallographic asymmetric unit consists of four almost identical FP-2 monomers (Fig 3.8), with each exhibiting good stereochemistry, and 40 water molecules (Table 3.3). The atomic coordinates and structure factors for the crystal structure of FP-2 have been deposited with the Research Collaboratory for Structural Bioinformatics Protein Databank (PDB) (<http://www.rcsb.org/pdb>) (Berman et al., 2000) with accession code 2GHU.



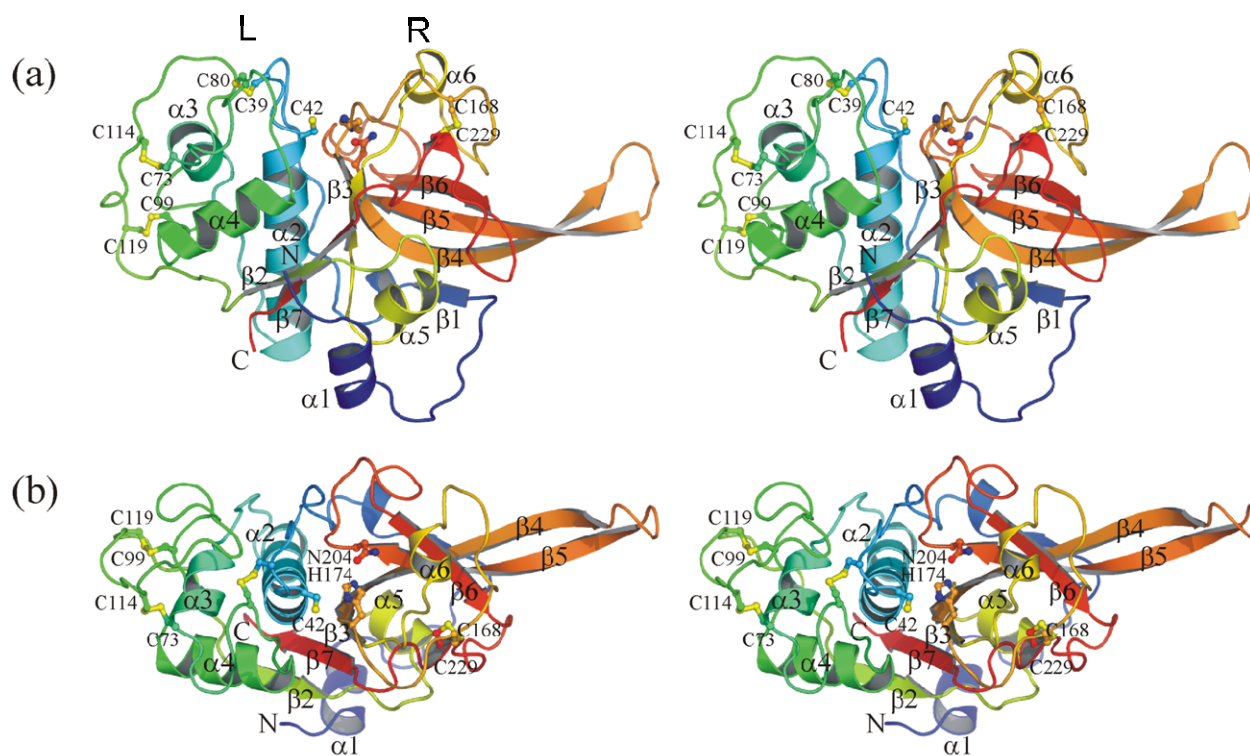
**Fig 3.8:** Cartoon representation of the four almost identical monomers of FP-2 present in the asymmetric unit. Each individual monomer is labeled and coloured differently.

## RESULT AND DISCUSSION

Despite a limiting resolution of 3.1 Å, the  $3F_o - 2F_c$  electron density maps were of good quality, especially for monomer D. Initially, the C $\alpha$  backbone atoms of monomer D was built manually into the electron density map, and then the other monomers (A, B, and C) were generated by NCS symmetry operations and most of the individual monomer side chains were built afterwards. This allowed unambiguous tracing of the polypeptide chain for 958 out of 964 amino-acid residues. The Pro189-Leu190 segment of monomers A and C was modeled with zero occupancy and the Glu14-Glu15 segment in monomer C was not modeled due to poor density in these regions. Unless otherwise indicated, the discussion of the structural aspects of FP-2 is based on the slightly better defined monomer D of the crystallographic asymmetric unit.

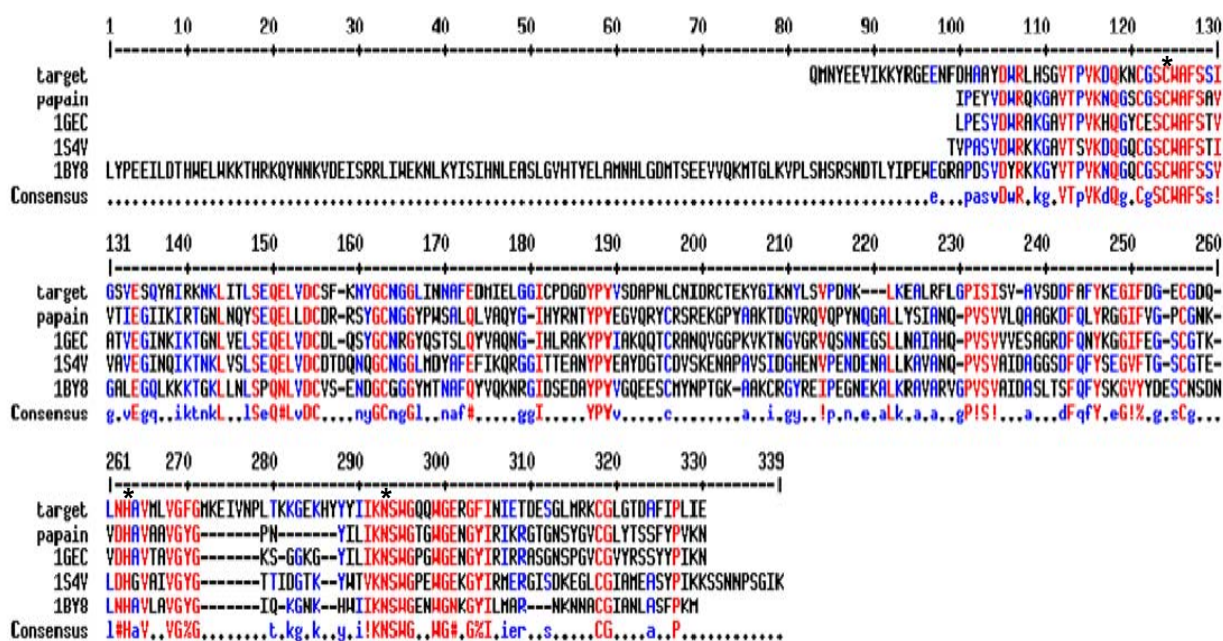
The overall fold of monomer D of FP-2 is illustrated in Fig. 3.9, showing a core structure with a distorted ellipsoidal shape that is characteristic of the C1A class of proteases. The polypeptide chain of mature FP-2 folds into the prototypical structure of papain-like proteases (Fig. 3.9), with two distinct domains separated by a long central substrate-binding cleft containing the active site (Barrett and Rawlings, 2001). The left domain (L-domain) is predominantly  $\alpha$ -helical ( $\alpha 2$ ,  $\alpha 3$ ,  $\alpha 4$ ), with several long disulphide-stabilized segments lacking regular secondary structure, while the right domain (R-domain) contains a large antiparallel  $\beta$ -sheet ( $\beta 2$ - $\beta 7$ - $\beta 3$ - $\beta 4$ - $\beta 5$ - $\beta 6$ ) that is extended at the central part of  $\beta 4$  by the very short  $\beta 1$  and is decorated by peripheral helices  $\alpha 1$ ,  $\alpha 5$  and  $\alpha 6$ . The residue numbering scheme chosen for the crystal structure reflects the sequence of mature FP-2 rather than its full-length precursor (Fig. 3.15), i.e., residue 1 has been assigned to the Gln of the new N-terminus liberated after self-processing (KYLLD↓QMNYF).





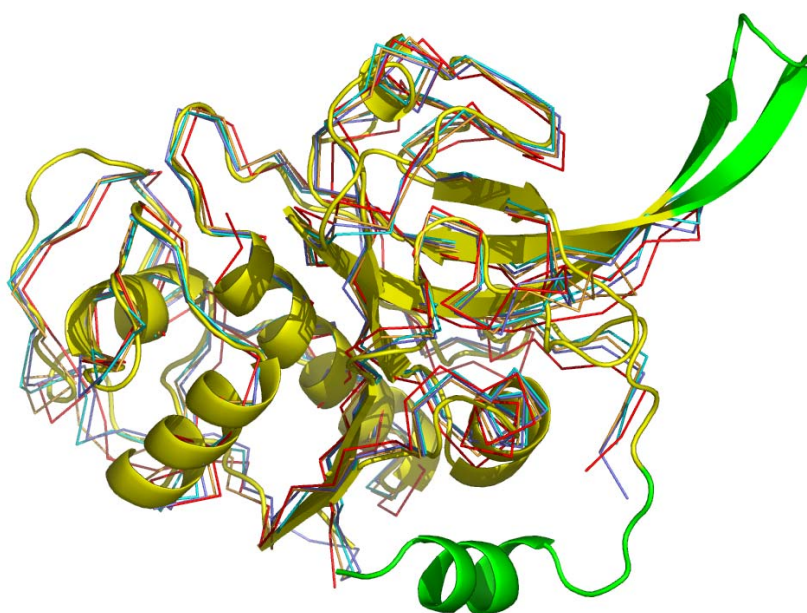
**Fig 3.9:** (a) Stereoscopic cartoon representation of FP-2 with the L-domain on the left, the R-domain on the right. The spectral colour-coding is according to sequence, from dark blue (amino terminus) to dark red (carboxy terminus). Disulphide bonds and active-site residues are shown in ball-and-stick. Secondary structure elements, amino- and carboxy-termini, the active-site cysteine (C42) and cysteine residues involved in disulphide bonds are labeled. (b) Stereoscopic orthogonal view relative to panel (a) generated by a ~90° rotation towards the viewer. Active-site residues C42, H174 and N204 are labeled.

**3.4 Sequential and structural comparison with C1A proteases** - Phyre v 0.2, a fold-recognition program (sequence to structure alignment) was used (Bennett-Lovsey et al., 2008) to find the homologous structures of FP-2. As expected, the hits were quite a few papain-family proteases which are structurally closely related to FP-2. The top four homologous proteins (9PAP: papain (Kamphuis et al., 1984), 1GEC: glycyl endopeptidase from *Carica papaya* (O'Hara et al., 1995), 1S4V: KDEL-tailed cysteine endopeptidase (Than et al., 2004), and 1BY8: human procathepsin K (Lalonde et al., 1999)) were used for multiple sequence alignment, carried out using the program Multalin 5.4.1 (Corpet, 1988) as shown in Fig. 3.10.



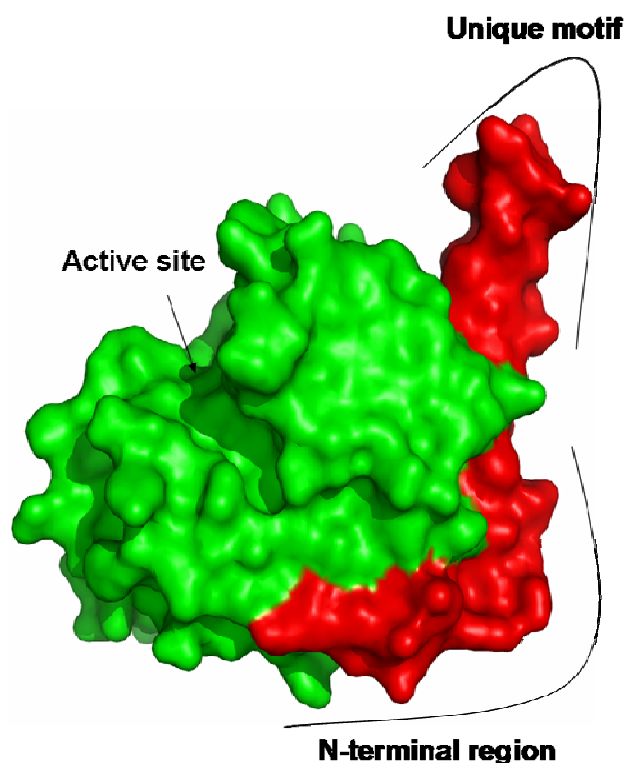
**Fig 3.10:** Multiple sequence alignment between FP-2 (target) and structurally closely related homologues proteins. The identical residues coloured in red and conservative replacements are coloured blue. A gap weight of 12 and gap length weight of 2 was used.

Fig. 3.10 demonstrates the high sequence similarity between various papain-like proteases. The active-site region comprising the catalytic triad C42, H174 and N204 (in case of FP-2) is especially highly conserved. These regions are represented by the sequence numbering 120-130 (C42), 261-270 (H174), and 290-300 (N204) respectively (residues marked with an asterix). Based on this sequence alignment, the X-ray structures were superimposed (Fig. 3.11).



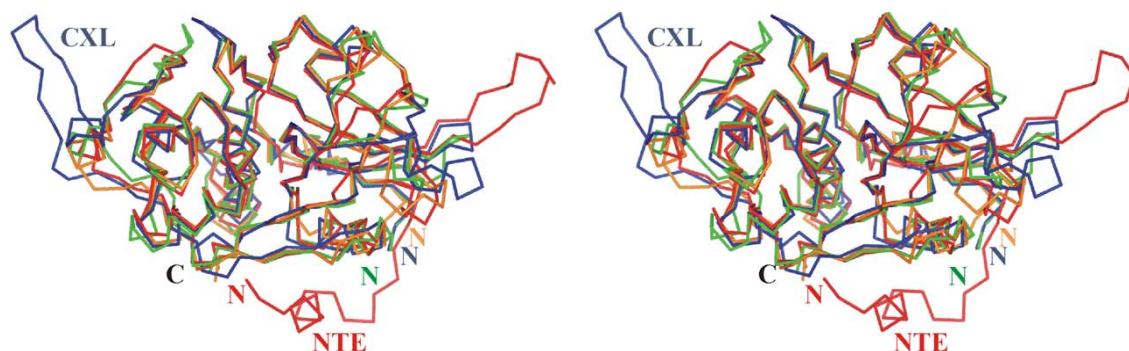
**Fig. 3.11:** Superposition of FP-2 and papain family proteases. FP-2 is represented as a cartoon and homologues structures as ribbons (9PAP: red, 1GEC: marine blue, 1S4V: violet, and 1BY8: orange). The regions that are unique to FP-2 are coloured green.

The superimposed structures and the sequence alignment (Fig. 3.10 and 3.11) revealed that the overall structure of FP-2 compliment the classical sequence and structure of C1A family proteases. The structures superimposed quite well, with an overall backbone (C $\alpha$ ) root-mean square deviation (r.m.s.d.) of 1.33 Å for KDEL-tailed cysteine endopeptidase, followed by papain (1.41 Å), glycyl endopeptidase (1.53 Å), and human procathepsin K (1.89 Å). The long N-terminal region of human procathepsin K (1BY8) was manually removed from the PDB file for better alignment. Interestingly, quite a few structural regions were unique to FP-2 and absent in other papain-family proteases. Presence of a long N-terminal extension and a rather unusual loop or unique motif (Fig. 3.11, coloured green) was instantaneously noticed. The surface model representation of FP-2 demonstrates the regions in detail (Fig. 3.12).



**Fig 3.12:** Surface model of FP-2 coloured in green. N-terminal region and unique structural motif are highlighted in red. The active-site cleft is also denoted.

The surface representation of FP-2 depicts the distinctive features of FP-2. The long N-terminal region and the unique motif are highlighted in red. This motif is a classical feature attributed only to plasmodial proteases (Pandey et al., 2005). In order to probe the entire Protein Data Bank for the closest known structural homolog of FP-2 exhibiting similar features, a comprehensive VAST search (Vector Alignment Search Tool) (Gibrat et al., 1996) was carried out. Interestingly, this procedure identified cruzipain (also known as cruzain), the major cysteine protease of the blood parasite *Trypanosoma cruzi* (the causative agent of Chagas' disease) (McGrath et al., 1995; Brinen et al., 2000; Huang et al., 2003), as the most similar structural relative to FP-2, with a backbone r.m.s.d. of 1.7 Å over 207 homologous residues. Using the same search algorithm, monoclinic papain (Pickersgill et al., 1992) was found to exhibit a backbone r.m.s.d. of 1.8 Å over 204 residues. The overall sequence identity between FP-2 and cruzipain is ~39 %, compared to ~38 % for papain.



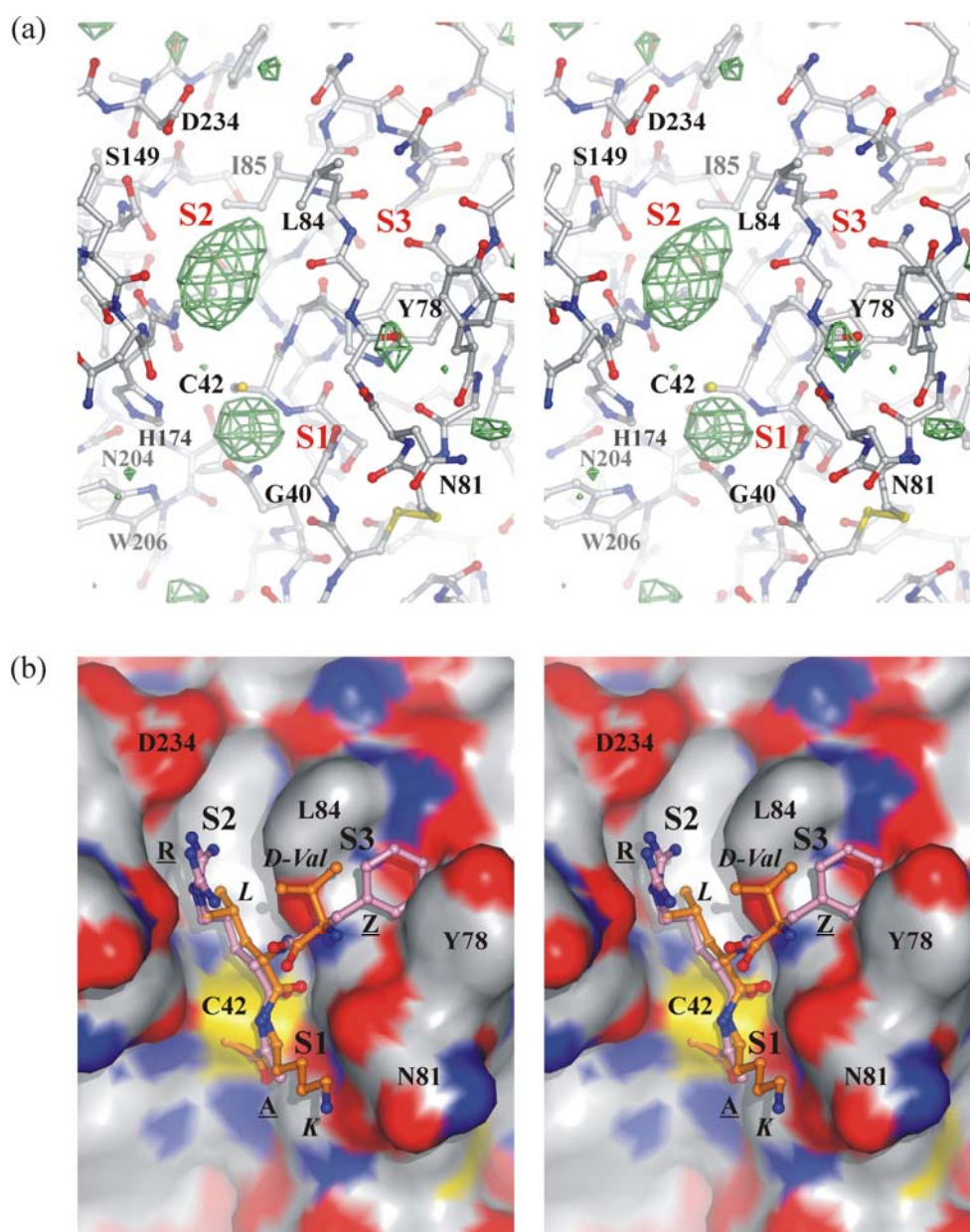
**Fig. 3.13:** Stereoscopic view of C $\alpha$  superposition comparing FP-2 (red), human cathepsin X (blue; PDB-ID: 1EF7) (Guncar et al., 2000), papain (orange; PDB-ID: 9PAP) (Kamphuis et al., 1984) and cruzipain (green; PDB-ID 1ME3) (Huang et al., 2003). Amino- and carboxy-termini, N-terminal extension (NTE) of FP-2, and the 110-123 loop (CXL) of cathepsin X are labeled.

Interestingly, cathepsin X, though not as well superimposed as the other structures, revealed the presence of a loop (CXL) almost of similar length to that of the FP-2 unique loop. But the loops are almost exactly 180° apart. The function of this loop is unknown (Guncar et al., 2000).



### 3.5 Three-dimensional structure insights of FP-2

**3.5.1 The active-site cleft** - The active site of FP-2 is located at the bottom of an elongated surface depression that has a slightly altered surface topology when compared to that of cruzipain and papain. The subsites S1, S2, and S3 are marked in red. The active-site Cys42 exhibits strong residual  $F_o - F_c$  electron density closely flanking the side chain sulphur atom, indicating the presence of an attached acetamide moiety, although the density was not strong enough to facilitate modeling of the inhibitor (Fig. 3.14a). Most striking was the presence of a large ellipsoidal  $F_o - F_c$  electron density peak residing within the S2 subsite (which is formed by the side chains of Trp43 ( $\alpha_2$ ), Leu84 ( $\alpha_3/\alpha_4$  loop), Ile85 (N-terminus of  $\alpha_4$ ), Ser149 (C-terminus of  $\beta_3$ ), Ala175 (N-terminus of  $\beta_5$ ), as well as the main chain carbonyl group of Gly83 ( $\alpha_3/\alpha_4$  loop)) (Fig. 3.14a). When the structures of various C1A-protease : inhibitor complexes were superimposed with FP-2, the difference density overlapped perfectly with the side-chain atoms of residues in the inhibitor P2 position. The binding mode of Z-R-A-FMK (benzoyl-arginine-alanine fluoromethylketone) in complex with cruzipain (Gillmor et al., 1997) is overlaid in pink ball-and-stick. D-V-L-K-CMK (D-valine-leucine-lysine-chloromethylketone) in complex with the KDEL-tailed cysteine endopeptidase (Than et al., 2004) is overlaid in orange ball-and-stick. (Fig. 3.14b), suggesting the density may have arisen from S2-subsite binding of Tris (a component of the crystallization buffer, section 2.2.5.1) or ethylene glycol (the cryoprotectant), or possibly iodide (a product of iodoacetamide treatment). Computation of an anomalous-difference Fourier did not support the presence of an anomalous scatterer such as iodine however, the residual  $F_o - F_c$  density lacks characteristic features which would allow an unambiguous assignment to any particular buffer molecule.



**Fig 3.14:** (a) Stereoscopic representations of the FP-2 catalytic cleft. Residues are coloured according to atom type (blue = nitrogen; grey = carbon; red = oxygen; yellow = sulphur). (a) Ball-and-stick model of the catalytic cleft overlaid with the final difference density map ( $F_o - F_c$ ;  $3\sigma$ ; green). Prominent difference peaks are visible within the S2 pocket and next to the sulphur atom of the catalytic Cys42. Substrate-specificity pockets S1, S2 and S3 are labeled in red; selected FP-2 residues are labeled in grey/black. (b) Surface representation of the catalytic cleft in an identical orientation as for panel (a). Binding modes of Z-R-A-FMK (benzoyl-arginine-alanine fluoromethylketone) in complex with cruzipain (Gillmor et al., 1997) are overlaid in pink ball-and-stick. Binding mode of *D*-V-L-K-CMK (*D*-valine-leucine-lysine-chloromethylketone) in complex with the KDEL-tailed cysteine endopeptidase (Than et al., 2004) is overlaid in orange ball-and-stick. Substrate specificity pockets (S1, S2 and S3), inhibitor residues, and selected FP-2 residues are labeled. Labels for Z-R-A-FMK residues are underlined; those for *D*-V-L-K-CMK are italicized.

**3.5.2 Subsite-specificity of FP-2** - The S2 pocket is the major determinant of specificity for most cysteine proteases (Brömme et al., 1994; Bode & Huber, 2000) and is predominantly hydrophobic in nature. All the characterized FP-2 cleavage motifs of its specific intracellular protein targets, including ankyrin (NVS**AR**↓FWLSD) (Dhawan et al., 2003), band 4.1 (SQEEIK↓KHHASI) (Hanspal et al., 2002), and the FP-2 precursor (KYLLD↓QMNYF) (Shenai et al., 2000) agree well with the observation. They all contain a hydrophobic residue at the P2 position (marked in bold in the preceding sequences). Experiments with fluorogenic substrates have also shown that FP-2 has a strong preference for substrates with a hydrophobic residue, particularly a leucine, at the P2 position (Shenai et al., 2000; Ramjee et al., 2006).

The crystal structure of FP-2 reveals that although the S2 subsite is generally hydrophobic, the electronegative side chain of Asp234 is positioned very deep in the S2 pocket (Fig. 3.14). Previous X-ray structures of cruzipain and cathepsin B have uncovered a Glu side chain in the homologous position (Glu205 in cruzipain; Glu245 in cathepsin B) (McGrath et al., 1995; Musil et al., 1991). It has been demonstrated for cruzipain that the binding of substrates containing a basic side chain in the P2 position is mediated by salt-bridge formation between the P2 side chain and the carboxylate moiety of Glu205 in the S2 pocket, thereby explaining the dual specificity of cruzipain and cathepsin B for substrates containing either a hydrophobic or basic residue at the P2 position (Choe et al., 2005). In FP-2, however the S2 pocket is slightly deeper due to the presence of the shorter Asp234 side chain at this position and it is therefore questionable whether an Arg or Lys in P2 could make a similar salt-bridge interaction (Fig. 3.14b). From the analysis of the active-site cleft, it might be feasible to exploit this feature of the S2 pocket in designing selective inhibitors specific against FP-2. A good start could be to engineer a pseudopeptide inhibitor containing a longer basic side chain in the P2 position, such as



homoarginine, which would presumably have the capacity to form a salt bridge with Asp234.

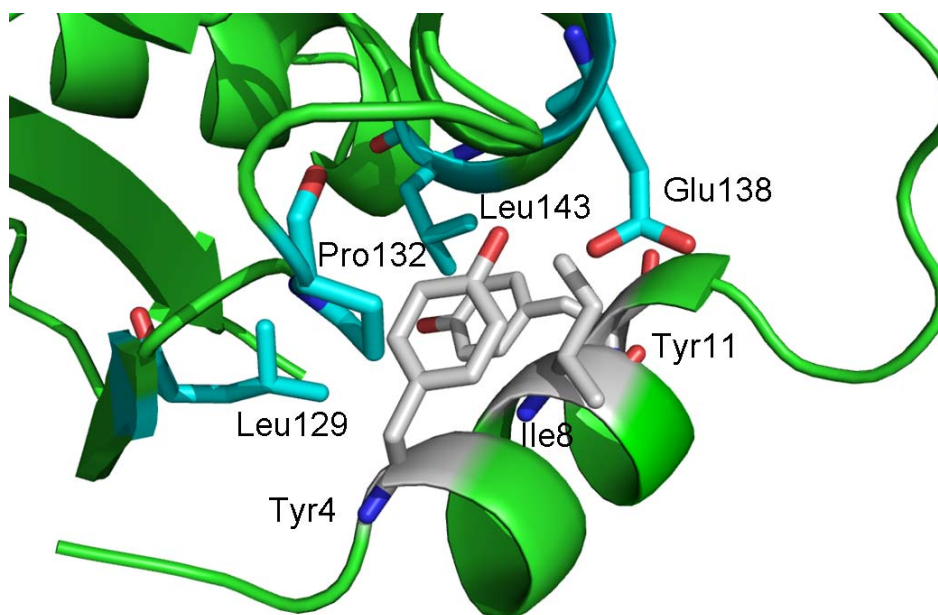
**3.5.3 An additional Cys99-Cys119 disulphide bond in FP-2** - Apart from the catalytic Cys42 (corresponding to Cys285 of the full-length precursor), there are a total of eight cysteine residues in FP-2 forming four disulphide bonds (Cys39-Cys80, Cys73-Cys114, Cys99-Cys119 and Cys168-Cys229) as shown in Fig. 3.9. The disulphides Cys39-Cys80 and Cys73-Cys114 are well conserved among the papain-like enzyme structures. Like papain and the majority of C1A proteases, FP-2 also has the disulphide bond Cys168-Cys229. This bond is involved in the fixing the upper loops that define the S2 and S1' substrate-binding sites (Fig. 3.9). This is in contrast to some of the cathepsins which lack this stabilizing element. An additional Cys99-Cys119 disulphide bond unique to the FP-2 family of proteases was predicted on the basis of homology modeling studies (Sabnis et al., 2002), and mutations of either Cys99 or Cys119 have been shown to abolish enzyme activity, probably by disrupting protein folding and leading to a catalytically inactive conformation (Goh & Sim, 2004). The FP-2 crystal structure confirms the presence of a Cys99-Cys119 disulphide which appears to play an important role together with the Cys73-Cys114 disulphide in stabilizing the long  $\alpha 4/\beta 2$  loop (~30 residues) that covers the lateral surface of the L-domain (Fig. 3.9).

**3.5.4 Structural features unique to the FP-2 subfamily** - Additional structural features that are unique to the FP-2 subfamily (with members such as the vivapain, vinckepain and berghepain homologues) include a 17-residue N-terminal extension, and a 14-residue  $\beta$ -hairpin protuberance extending from  $\beta 4$  and  $\beta 5$  that comprises the unique motif (Fig. 3.9, 3.11, 3.12, and Fig. 3.15).

59

Fig. 3.15 demonstrates the high sequence similarity between various FP-2-subfamily proteases, such as the human pathogens *Plasmodium falciparum* and *P. vivax*, and the rodent pathogens *P. vinckei*, *P. chabaudi*, and *P. berghei*. The boxed insert shown in the figure highlights the N-terminal extension and the unique motif. The biological roles of these insertions are explained in detail in section 3.5.4.3. These regions, though sequentially dissimilar, are structurally highly conserved within the subfamily. The only exception to it is FP-1. Structural differences were clearly evident between the four FP-2 monomers in both of these regions, specifically in  $\alpha 1/\beta 1$ -loop residues 13 - 16, and in residues 186 - 195 ( $\beta 5$ -loop- $\beta 6$ ) of the unique motif. These segments were therefore exclusively omitted from the 4-fold NCS restraints implemented during crystallographic refinement.

**3.5.4.1 The N-terminal extension** - The presence of a 17-residue N-terminal extension is one of the major distinctive structural features of the FP-2 subfamily when compared to other C1A proteases. Previous functional studies have shown that the N-terminal extension plays a crucial role in folding of the mature protein into its active conformation (Pandey et al., 2004). The crystal structure of FP-2 reveals that the N-terminal extension comprises a short  $\alpha$ -helix ( $\alpha 1$ , residues 4-11) followed by a 10-residue loop (Fig. 3.9 and 3.11). The  $\alpha 1$  helix is stabilized by the highly conserved Tyr4, Ile8, and Tyr11, which pack tightly against residues of the R-domain elements  $\beta 2$  (Leu129),  $\alpha 5$  (Glu138, Leu143), and the intervening  $\beta 2/\alpha 5$  loop (Pro132) as shown in Fig 3.16. The long loop (residues 12-21) of the N-terminal extension passes through the lower surface of the R-domain, before entering the domain's core nearly 20 Å away at the base of the unique extended  $\beta$ -hairpin motif. The loop exhibits elevated conformational flexibility as evidenced by relatively high B factors and diffuse electron density in some of the monomers (A and C).



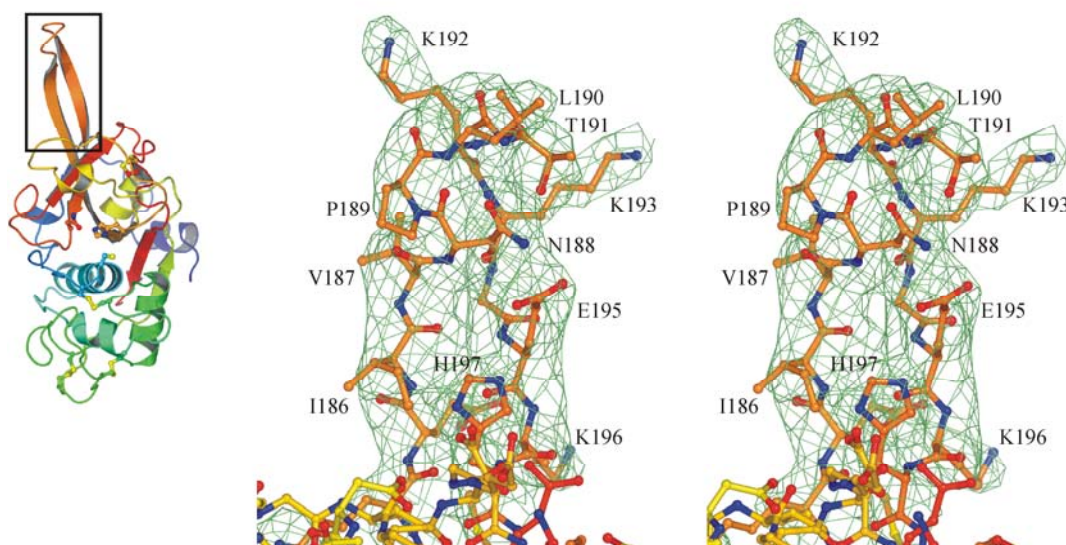
**Fig 3.16:** The N-terminal helix is stabilized by the highly conserved residues Tyr4, Ile8, and Tyr11 shown in white, which pack tightly against residues of the R-domain elements  $\beta 2$  (Leu129),  $\alpha 5$  (Glu138, Leu143), and the intervening  $\beta 2/\alpha 5$  loop (Pro132) shown in cyan.

**3.5.4.2 N-terminus as a structural determinant of protein folding** - Although the exact involvement of the 17-residue N-terminal segment in mediating correct folding of mature FP-2 is not clearly understood, previous mutational studies have suggested a critical role for the highly conserved Tyr4 (Pandey et al., 2004). The X-ray structure of FP-2 reveals this residue to be situated at the N-terminus of helix  $\alpha 1$  where it is involved in multiple side chain interactions with  $\beta 2$ ,  $\alpha 5$  and the  $\beta 2/\alpha 5$ -bridging loop of the R-domain core (Fig. 3.16). This observation and the high sequence conservation of other residues mediating packing interactions between helix  $\alpha 1$  and the bulk of the R-domain would spontaneously suggest that positioning of  $\alpha 1$  in this region is critical for overall folding of the polypeptide, and that the long  $\alpha 1/\beta 1$  loop may act simply as a spacer to allow the N-terminal helix  $\alpha 1$  to interact with the lower surface of the R-domain. An essential role for the N-terminal extension in protein folding cannot be easily substantiated by a structural comparison of FP-2 with other C1A proteases

lacking this structural element. However, the underlying  $\beta$ 2-loop- $\alpha$ 5 structure does not exhibit any pronounced dissimilarities between the different proteins (Fig. 3.9).

**3.5.4.3 The unique motif / loop** - A 14-residue insertion located in the C-terminal half of FP-2 has been recently shown to fulfill an important function by capturing hemoglobin (Pandey et al., 2005). This unique loop is also known as FP-2<sub>arm</sub> (Wang et al., 2006). This insert adopts an extended antiparallel  $\beta$ -hairpin conformation which causes the  $\beta$ 4/ $\beta$ 5 strands to protrude out from the surface of the R-domain by ~20 Å (Fig. 3.9, 3.12, & Fig. 3.17). The high conformational flexibility of the  $\beta$ -hairpin was responsible for the diminished electron density in the distal direction along the long axis of this segment in monomers A and C. This is not surprising since the  $\beta$ -hairpin of monomers A and C extends into bulk solvent and does not make contact with any adjacent monomers in the crystal (this crystal form has a solvent content of 75.5 %). As a result, the distally positioned residues of the  $\beta$ -hairpin were difficult to model and refine in monomers A and C, hence some of these residues exhibit rather strained main chain torsion angles, and Pro189 and Leu190 at the tip of the hairpin in these monomers were modeled with zero occupancy. In monomers B and D, however, the  $\beta$ -hairpin is stabilized by intermolecular contacts, allowing for confident modeling and refinement of the entire insert including residues at the tip of the hairpin (Fig. 3.17).





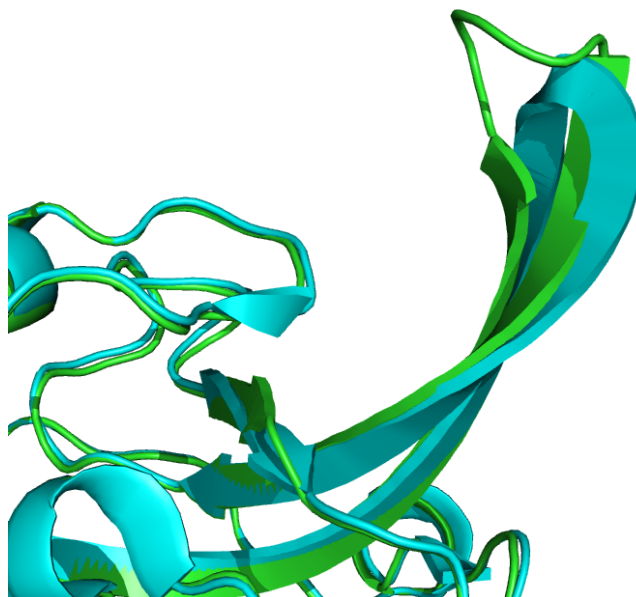
**Fig 3.17:** Stereoscopic view of the FP-2 hemoglobin-binding  $\beta$ -hairpin. A ball-and-stick model for the hemoglobin-binding  $\beta$ -hairpin and neighbouring residues of monomer D is shown (boxed inset highlights the area of detail). A simulated-annealing omit map ( $F_o - F_c$ ;  $3\sigma$ ) calculated for residues 184-197 of the  $\beta$ -hairpin is superimposed, and residues of the  $\beta$ -hairpin are labeled.

The 14-residue hemoglobin-binding  $\beta$ -hairpin (residues 183-196) exhibits a high degree of conformational flexibility when freely exposed to bulk solvent. The large positional deviations exhibited by the  $\beta$ -hairpin when comparing this structure and the structure of the FP-2 : cystatin complex (Wang et al., 2006) further underscores an inherent flexibility and plasticity of this element that may have functional implications in hemoglobin binding. The sequence comparison studies conducted for different falcipains and related homologues indicated conservation of this insert length (14 residues) but with very low sequence identity (Fig. 3.15). Containing a much larger insertion in this region, FP-1 is an exception to this rule. FP-1 has a much larger  $\alpha 1/\beta 1$  loop, but it has also been shown to be incapable of degrading hemoglobin (Greenbaum et al., 2002).

**3.5.4.4 Comparison with the FP-2 : cystatin complex** - One day after the experimental data for FP-2 were submitted to the Protein Data Bank (PDB), the 2.7 Å crystal structure of FP-2 in complex with egg-white cystatin (a macromolecular

cysteine protease inhibitor) appeared in the PDB (1YVB; Wang et al., 2006). A brief comparison between the two structures showed that the overall C $\alpha$  r.m.s.d. between monomer D of our FP-2 crystal structure and their FP-2 in the cystatin complex is 0.63 Å. As expected, the largest C $\alpha$  deviations between the two structures occur in the hemoglobin binding insert (Leu190 at the tip of the  $\beta$ -hairpin exhibits the most prominent C $\alpha$  r.m.s.d (3.0 Å, Fig. 3.18)). The structural dissimilarity does not appear to be a direct consequence of cystatin binding to FP-2 as this  $\beta$ -hairpin is positioned quite far from the cystatin in the FP-2: cystatin complex.

From both structures of FP-2, it is clear that the proximity of the N-terminal  $\alpha$ 1/ $\beta$ 1 loop region to the so-called unique motif does raise speculations for a possible function in hemoglobin capture (Fig. 3.12). This is supported by sequence comparison studies of the different plasmodial FP-2 subfamily proteases, which confirmed that the  $\alpha$ 1/ $\beta$ 1 loop is restricted to a length of 10 - 11 residues (Fig. 3.15). The evidence directs towards a possible hemoglobin-binding role for the N-terminal extension, particularly the  $\alpha$ 1/ $\beta$ 1 loop. Whether this structural element does indeed contribute to hemoglobin capture requires further examination. Similarly for the hemoglobin-binding hairpin or ‘‘unique motif’’, a sequence analysis of 607 family C1A proteases has identified 40 members with insertions of longer than 12 residues at this location, with 18 of these members representing falcipains and FP-2 homologues from other species (Pandey et al., 2005). Other C1A proteases, such as cathepsin X (Guncar et al., 2000), contain prominent insertions at other locations (Fig. 3.13), although the biological role of these insertions remains less explored.

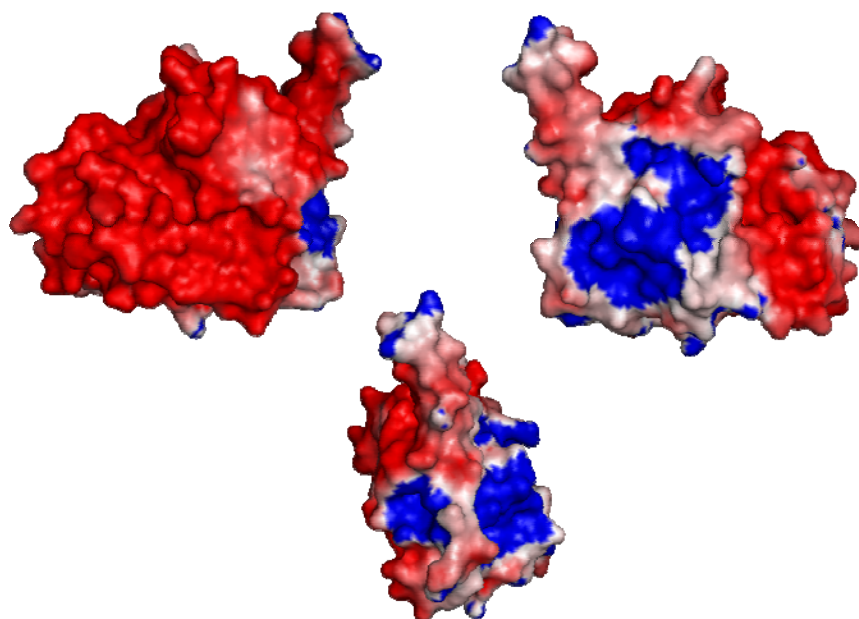


**Fig. 3.18:** superimposition of the unique loop of wild-type FP-2 (green) and Wang et al., 2006 structure (cyan) where the largest C $\alpha$  deviations occur.

**3.6 Mechanism of hemoglobin degradation** - Though the function of this insert has been established (Pandey et al., 2005), the mechanism of hemoglobin binding or how the degradation of substrate proceeds is still unclear. To understand the underlying mechanism of hemoglobin binding and degradation by FP-2 and to verify the involvement of this motif, a series of experiments were conducted.

**3.6.1 Electrostatic potential map calculation of FP-2** - Using the software APBS tools, electrostatic potential maps were calculated for FP-2 at pH 5.0. The charge distribution showed that the L-domain is totally acidic in nature and patches of positive charges could be observed in parts of the R-domain and also in the unique motif. These positive charges were due to the contributions from residues Lys192 and Lys193 in the unique motif. Furthermore the N-terminus also exhibited high positive surface potential. In Fig. 3.19, blue colour denotes positive potential, red negative, with transparent white being neutral (zero potential energy).

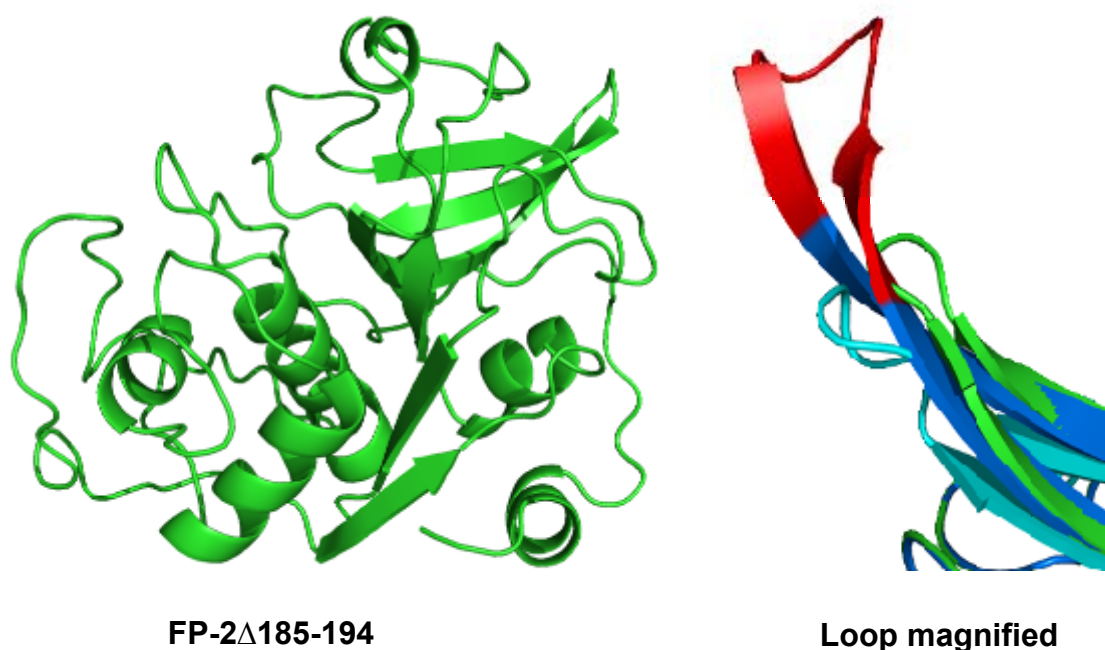




**Fig. 3.19:** Electrostatic potential surface representation of FP-2 at pH 5.0. The side view and the front view clearly depict the charge distribution. Blue colour denotes positive potential, red negative, with transparent white being neutral (zero potential energy). The electrostatic potential isocontour (kT/e) level was set at 0.1 for neutral and -4.0 and 0.1 for charged residues.

**3.6.2 Homology modeling of different FP-2 constructs** - Various FP-2 constructs were designed based on the studies of the hemoglobin-binding hairpin and the results from electrostatic potential calculations. One of the constructs, FP-2 $\Delta$ 185-194, is a  $\Delta$ 10 loop deletion mutant (deleted amino-acid sequence EIVNPLTKKG). This deletion removes large parts of the extended anti-parallel  $\beta$ -hairpin unique motif without influencing the overall fold of wild-type FP-2. This mutant was designed to mimic the papain structure which does not contain the unique substrate binding loop. The mutant was designed based on the studies by Pandey and co-workers (Pandey et al., 2005). The plasmid encoding the protein was expressed and purified as described in section 2.2.1.

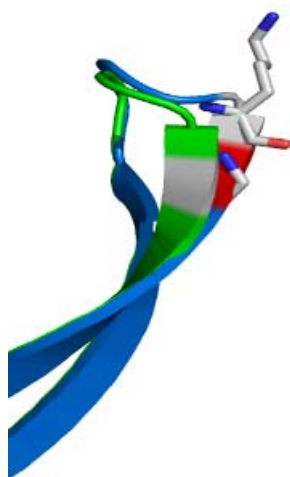
The structural model of this construct was based on the wild-type FP-2 crystal structure. The modeling was pretty straightforward and Swiss model was used to generate the PDB file.



**Fig 3.20:** The model of FP-2Δ185-194 (green) based on wild-type FP-2. The loop region of the structure is magnified and superimposed onto wild-type FP2 (blue) and KDEL-tailed cysteine endopeptidase (magenta) (Than et al., 2004).

The FP-2Δ185-194 model was pretty much the same as the wild-type FP-2 except that the unique loop region has been shortened. Fig. 3.20 shows the superimposition of the motif onto wild-type FP2 and KDEL-tailed cysteine endopeptidase (Than et al., 2004). The extra region in wild-type FP-2 is indicated in red and the motif is quite structurally similar to that of the KDEL-tailed cysteine endopeptidase (Than et al., 2004) with a backbone Cα r.m.s.d. of 1.31 Å.

From the electrostatic potential calculations, it was clear that residues Lys192 and Lys193 if at all play a role in hemoglobin binding. In order to exploit this interaction and to cutoff the interactions, a mutant (FP-2ΔKK/G) was prepared that lacks one of the lysine residues and had the other lysine replaced by a glycine residue. The introduction of glycine residue gave stability to the loop. A structural model was built also for this construct based on the wild-type FP-2 (Fig 3.21).

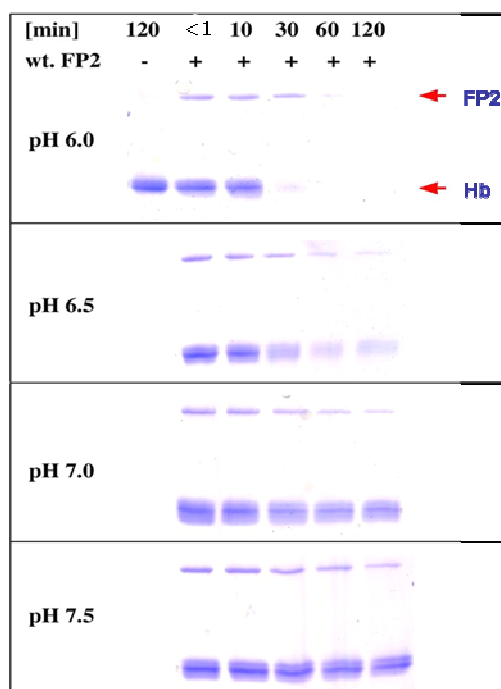


**Fig 3.21:** Superposition of the loop region of wild-type FP-2 and FP-2 $\Delta$ KK/G mutant.

The plasmid encoding the protein was expressed and purified as described in section 2.2.1. To establish the importance of the unique loop in the involvement of hemoglobin binding, the mutant proteins FP-2 $\Delta$ 185-194, FP-2 $\Delta$ KK/G were used alongside wild-type FP-2 in hemoglobin degradation assays.

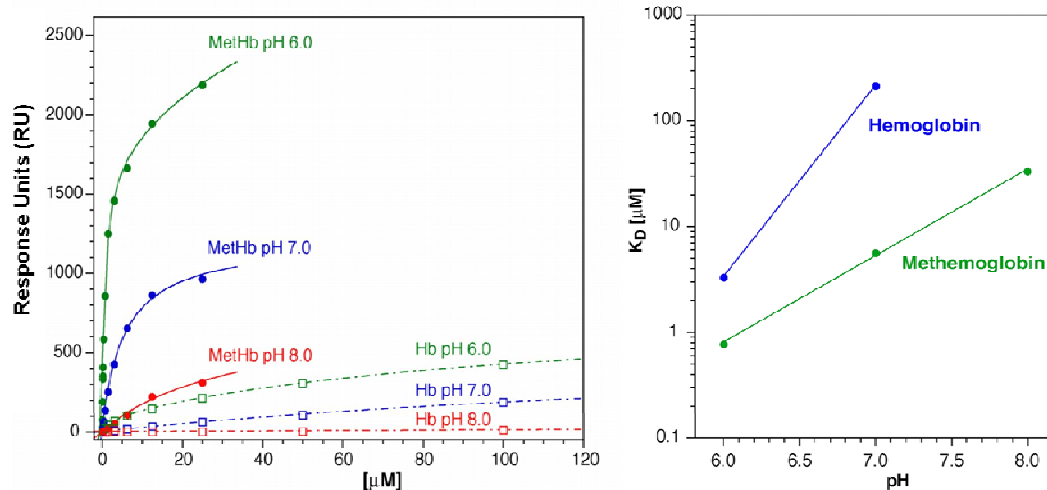
### **3.7 Hemoglobin degradation assays -**

**3.7.1 The pH-dependent activity of FP-2 with hemoglobin** - The pH-dependent proteolytic activity of recombinant, active FP-2 was measured with human methemoglobin at different pH ranges (Fig. 3.22). This assay was carried out at 37°C and similar to the procedure as described for the auto-processing of FP-2 prodomain (Fig. 3.3). The hemoglobinase activity (Fig. 3.22) displayed a clear pH optimum in the acidic range as has been previously described (Shenai et al., 2000). Hemoglobin was completely digested within 30 min at pH 6.0. Also at pH 6.5, FP-2 exhibited efficient but diminished proteolysis activity. No hemoglobin band could be observed within 60 min of incubation, whereas only minimal or no proteolysis was observed at pH 7.0 and 7.5, respectively. This result is also in contrast to the autoprocessing assay shown in Fig. 3.3, where the *trans* processing exhibited detectable activity over a wide pH range and showed a maximum at pH 8.5. This may be due to the fact that the *in-vivo* degradation of hemoglobin by FP-2 is carried out inside the food vacuole, where the pH is also in the acidic range (Yayon et al., 1984; Krogstad et al., 1985). It is also interesting to note that FP-2 undergoes some self-degradation at pH 6.0 and 6.5 (Fig 3.21, 60 & 120 min) whereas almost no or minimal degradation was observed at pH 7.0 and 7.5.



**Fig 3.22:** pH-dependent hemoglobinase activity of wild-type FP-2. Hb = hemoglobin. The protein is completely degraded at acidic pH values and little or no degradation was observed at higher pH values.

To further investigate the basis for the different pH dependencies of the various catalytic activities of wild-type FP- 2, the binding of hemoglobin to mature inactive FP-2 (FPc285aM) by surface plasmon resonance was investigated using a Biacore<sup>®</sup> 3000 system. These experiments also complemented the previous results that the binding of hemoglobin to FP-2 decreased with increasing pH (Fig. 3.22). The results demonstrated that the observed reduction in hemoglobinase activity of FP-2 at higher pH values is due to decreased substrate binding under these conditions. Interestingly, very significant differences were observed between the binding of FP-2 to methemoglobin and hemoglobin, with methemoglobin exhibiting much lower  $K_d$  values (Fig. 3.23).



**Fig 3.23:** pH-dependent binding of human hemoglobin and methemoglobin to inactive FP-2 as measured by surface plasmon resonance. The intensity of the steady-state response (RU) was plotted against hemoglobin concentration. The insert shows the pH-dependence of the dissociation-constants for the Hb-FP-2 and metHb-FP-2 complexes. MetHb = methemoglobin.

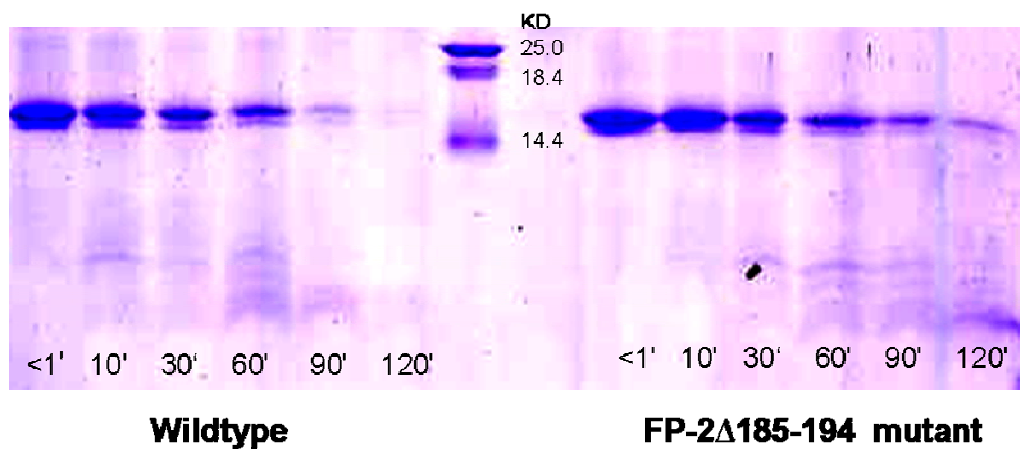
**3.7.2 Preference for methemoglobin over hemoglobin** - Affinity of wild-type FP-2 for methemoglobin is significantly higher than that for hemoglobin was completely unexpected. Previous studies have shown that several factors contribute to the formation of methemoglobin during plasmodial infection, including the acidic pH of the plasmodial food vacuole, oxidative damage within infected erythrocytes (Akompong et al., 2000; Müller, 2004) and the reduced activity of NADH-methemoglobin reductase (Stocker et al., 1985). Significantly, increased methemoglobin content in the range of 20- 42 % has been detected in the plasmodial food vacuole compared to 0.6- 1.0 % in uninfected erythrocytes (Akompong et al., 2000). Thus, the higher affinity of FP-2 for methemoglobin might be due to abundant natural availability and seems to reflect a natural adaptation by the parasites to the conditions within *Plasmodia*-infected erythrocytes. This hypothesis is supported by further experimental data, where Sobolewski and colleagues (Sobolewski et al., 2005) demonstrated in *ex-vivo* experiments using murine erythrocytes that the conversion of 95 % of the hemoglobin into methemoglobin by addition of NO (nitrous oxide) had no detectable impact on the viability of *P. berghei*. The authors suggested

that under these conditions, the parasites had either resorted to an alternative food source or were able to utilize methemoglobin. Akompong et al. (2000) further demonstrated that a reduction of the methemoglobin content by the addition of riboflavin resulted in a reduced size of the *P. falciparum* food vacuole and blocked parasite proliferation in erythrocyte cultures. Consequently, it seems that the formation of methemoglobin is not an unwanted side effect of a plasmodial infection, but a prerequisite for the efficient utilization of hemoglobin as a food source by the parasite. Thus it may be concluded that the selectivity of FP-2 for methemoglobin may represent an evolutionary adaptation to oxidative stress conditions within the host cell and these results indicate that FP-2 is perfectly suited to perform this function.

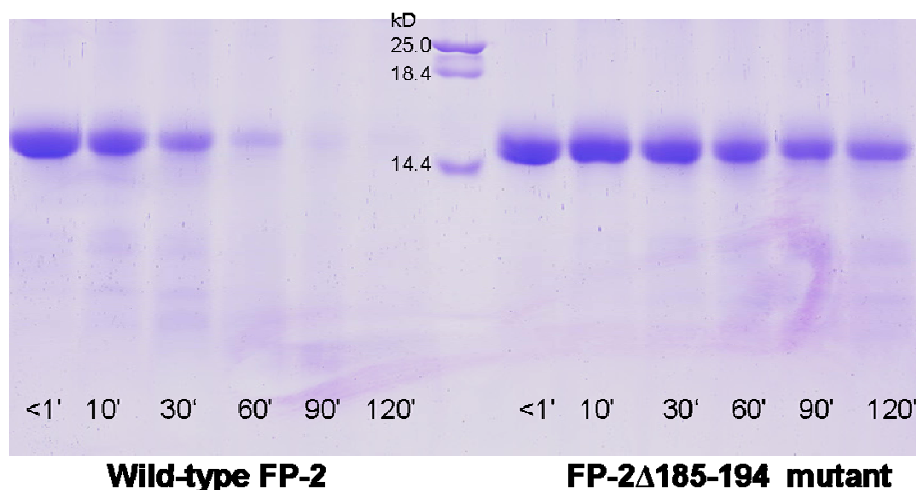
**3.7.3 FP-2 $\Delta$ 185-194 mutant cleavage assays** - During the pH-dependent degradation assay of hemoglobin, several intermediate degradation bands were observed. In order to verify this and also to establish the importance of the unique loop, the FP-2 $\Delta$ 185-194 protein was introduced into hemoglobin assays along with wild-type FP-2. The proteolytic activity of the mutant and wildtype FP-2 was measured with human hemoglobin at pH 5.5, mimicking the pH within the digestive food vacuole (All the following experiments were carried out at pH 5.5 unless otherwise indicated). For the first time, an activity of the FP-2 $\Delta$ 185-194 mutant is reported (Fig. 3.23). Surprisingly, the mutant was readily active against hemoglobin. In the case of wild-type FP-2, hemoglobin was completely degraded within 120 min of incubation. However, a faint band of hemoglobin was still visible (Fig. 3.23, 120 min) in the mutant activity assay indicating that the efficiency of the protein is slightly diminished. The reason for this diminished activity is yet to be established and only a three-dimensional structure of the mutant protein can provide the solution. Fig. 3.24

## RESULT AND DISCUSSION

demonstrates that the degradation proceeded in a step-wise manner, producing relatively stable intermediates. Both forms of FP-2 produced comparable results representing a similar cleavage pattern. The intermediate bands were visible within 10 minutes of incubation and their intensities were at a maximum at 60 min incubation time. With time, these intermediate cleavage bands are further processed into much smaller peptide fragments and in the case of wild-type FP-2, the protein is completely degraded within 120 min of incubation.



**Fig 3.24:** Human hemoglobin (HHb) cleavage by wild-type FP-2 and FP-2Δ185-194 mutant. The cleavage patterns are very much comparable. The marker and the incubation times are denoted.



**Fig 3.25:** Degradation of rat hemoglobin (RHb) by wild-type and mutant FP-2. The degradation is reduced in the case of the mutant, in comparison to processing of the human hemoglobin.



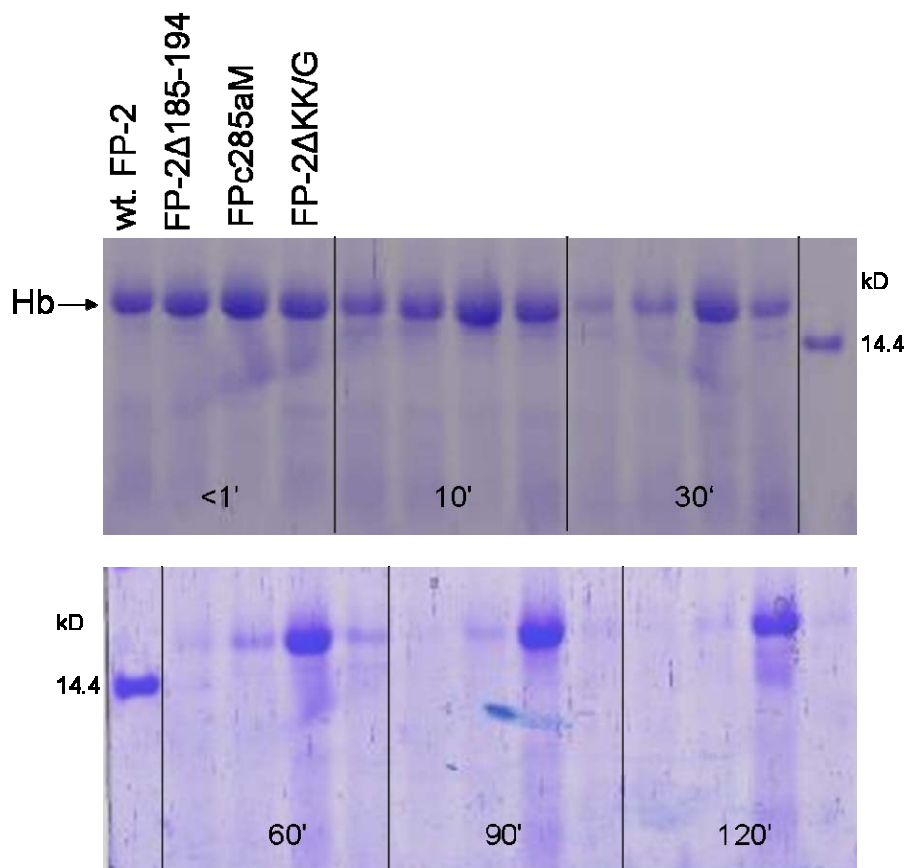
A similar kind of assay was carried out using rat hemoglobin (RHb) as a substrate (Fig. 3.25). Cleavage proceeds in a similar manner as that of human hemoglobin producing stable intermediates. The intermediate bands of hemoglobin due to cleavage by both the wild-type and mutant the protease follow the same pattern. The wild-type FP-2 completely degrades the substrate within 90 min of incubation and no visible rat hemoglobin band could be observed after that. The activity of the FP-2 $\Delta$ 185-194 mutant is lower when compared to that of wild-type.

From the above results, it is pretty clear that both the wild-type and the mutant FP-2 can degrade hemoglobin without previous intervention by plasmepsins. The protease activity of the FP-2 $\Delta$ 185-194 mutant is somewhat reduced compared to that of the active mature wild-type FP-2. Hemoglobins (HHb, RHb) are shown to be degraded in a well-ordered and specific manner. The activity of the mutant on hemoglobin questions the role of the unique loop in capturing the substrate through its arm-like structure and contradicts the earlier reports (Pandey et al., 2005; Wang et al., 2006).

**3.7.4 Hemoglobin cleavage assay by FP-2 $\Delta$ KK/G** - The FP-2 $\Delta$ KK/G mutant (FP-2 lysine deletion and lysine/glycine mutant, where lysine 192 has been deleted and lysine 193 has been mutated to a glycine (K193G)) was purified and refolded in optimized refolding buffer (Section 2.2.1). The protein was fully activated upon refolding. The assay included the following FP-2 constructs: wild-type FP-2 sample as in lane 1 in all time intervals, FP-2 $\Delta$ 185-194 (lane 2), FPc285aM (lane 3), and FP-2 $\Delta$ KK/G (lane 4) (Fig 3.26). In this assay, FP-2 $\Delta$ KK/G efficiently hydrolyzed human hemoglobin at acidic pH (5.5) with almost complete turnover of substrate after 120 minutes incubation at 37°C (Fig 3.26).

## RESULT AND DISCUSSION

Interestingly, samples obtained after 1, 10, 30 minutes incubation showed metastable degradation intermediates. The degradation proceeded at more or less the same rate as that by the wild-type FP-2. Complete substrate processing was achieved within 90 min. FP-2 $\Delta$ 185-194 also showed efficient hydrolysis. The inactive FPc285aM mutant was used as a control.



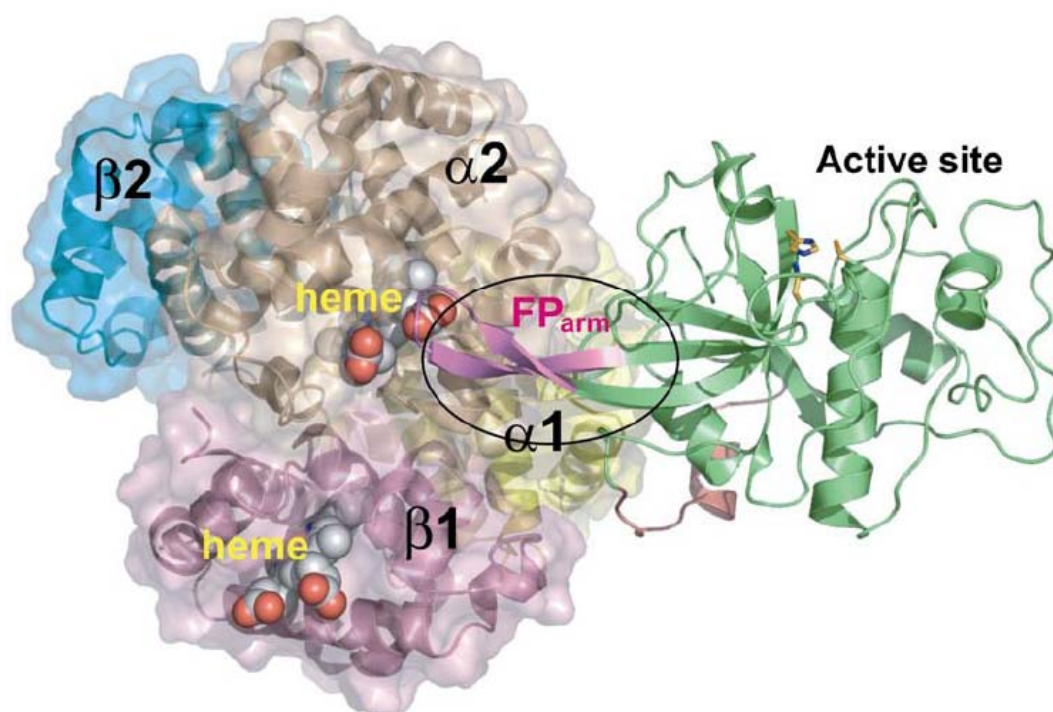
**Fig 3.26:** Degradation of hemoglobin by various FP-2 constructs (wild-type FP-2 (lane 1), FP-2 $\Delta$ 185-194 (lane 2), FPc285aM (lane 3), FP-2 $\Delta$ KK/G (lane 4). Samples were collected at different time intervals. The incubation time and the lysozyme as marker are denoted.

The distribution of cleavage intermediates of both human hemoglobin and rat hemoglobin generated after various time intervals seemed to be identical for wildtype and the mutants. Efficient hemoglobin degradation with metastable intermediates indicated that the FP-2 cleavage of hemoglobin is not a random process rather proceeds in a highly specific and sequential manner. This result again affirms that the unique loop may not be an essential requirement in hemoglobin degradation. The

outcome of these experiments further leads to the speculation that the N-terminus of the protease might play a critical role in hemoglobin capture. In this way, the slightly reduced activity of the mutants in comparison to the wild-type can be explained.

### 3.8 FP-2 studies with myoglobin as a substrate -

**3.8.1 Hemoglobin binding models** - Two contradictory models for hemoglobin binding to FP-2 have been proposed. Pandey et al. (2005) reported that the binding of FP-2 to hemoglobin is of 4:1 ratio, such that four molecules of FP-2 for each hemoglobin tetramer. Later in 2006, the results of protein-protein docking from the same group (Wang et al., 2006) showed that the unique loop or FP-2<sub>arm</sub> of FP-2 binds in the interface between  $\alpha 1$  and  $\beta 2$  globins, with its tip near the heme moiety in the  $\alpha 2$  globin. This would correspond then to a 2:1 model i.e. two molecules of FP-2 for each hemoglobin tetramer (Fig. 3.27).

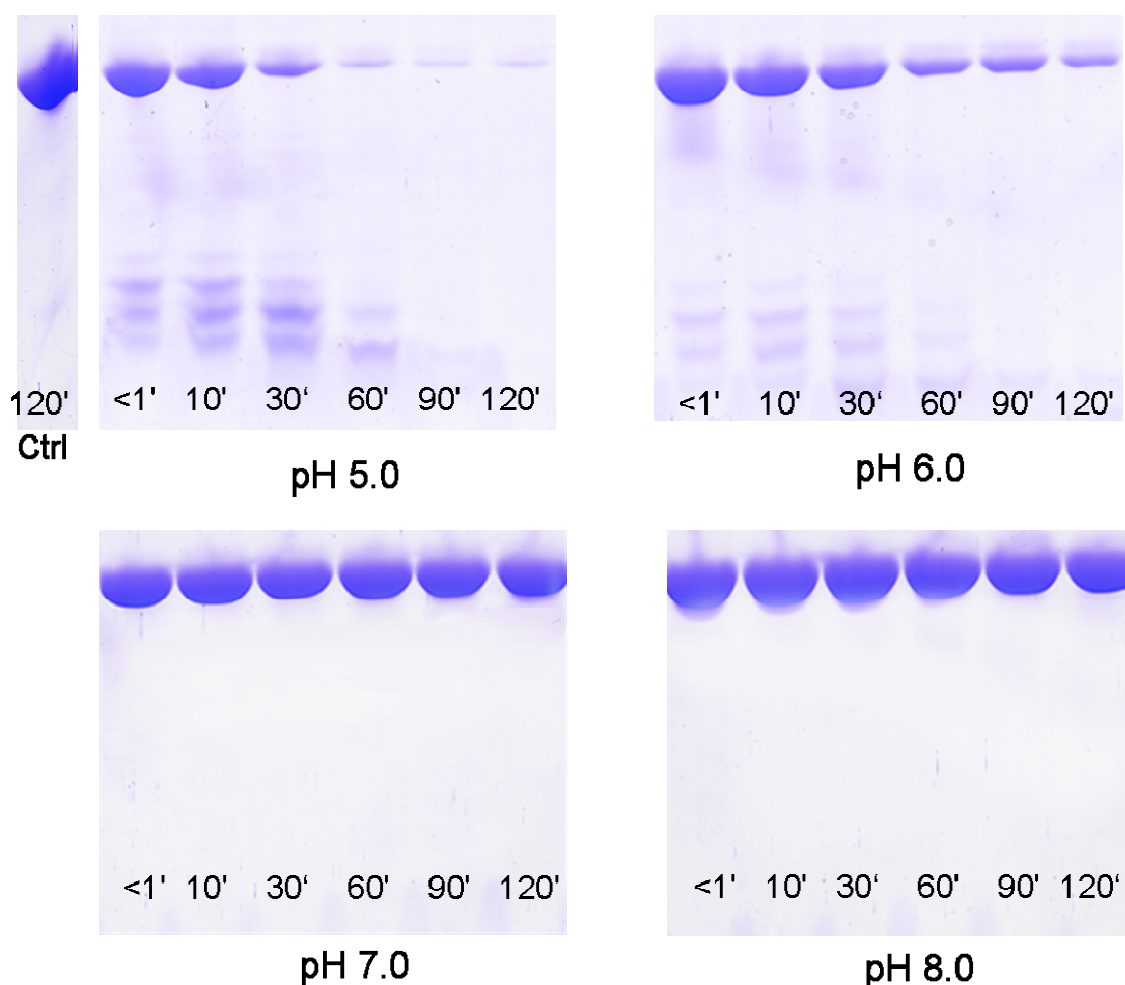


**Fig 3.27:** A putative model of an FP-2-hemoglobin complex. Semi-transparent surface representation of the hemoglobin tetramer (PDB-ID: 1BZ0), with  $\alpha$  and  $\beta$  subunits labeled and coloured differently. The heme groups are displayed as spheres and coloured in white. FP-2 is displayed by cartoon presentation only in green. The FP-2<sub>arm</sub> or unique motif is highlighted in pink. Picture taken from Wang et al. (2006).

In addition, they suggested that FP-2 may bind hemoglobin in a two-step sequential manner: first through the unique motif or FP-2<sub>arm</sub> and then at the active site. Based on the previous results (Section 3.7), FP-2 is active even without the involvement of this motif. To obtain deeper knowledge about the mechanism of hemoglobin binding and to unravel the ordered degradation by FP-2, myoglobin was used in this study as a simplified model for hemoglobin.

**3.8.2 Proteolytic cleavage of myoglobin by FP-2** - The proteolytic activity of recombinant, active FP-2 with myoglobin as substrate was tested at different pH values. FP-2 efficiently hydrolyzed myoglobin at pH 5.0 with almost complete turnover of the substrate within 120 minutes incubation at 37° (Fig. 3.28). The cleavage rate was slightly reduced at pH 6.0 (Fig. 3.27), whereas minimal or no proteolysis was observed at pH 7 or 8 and in control experiments without enzyme (Fig. 3.28). This pH optimum corresponds to the pH reported for the food vacuole (pH 5.0-5.4) (Yayon et al., 1984; Krogstad et al., 1985) and the pH optimum of myoglobin was very much similar to that of hemoglobin degradation as reported for FP-2 (Shenai et al., 2000; Hogg et al., 2006, or Fig. 3.22). The proteolysis of myoglobin proceeds via distinct intermediates (Fig. 3.28, <1 to 60 minutes incubation time). Four to five discrete myoglobin fragments could be identified immediately after the start of the reaction (Fig. 3.28, <1 min). These primary cleavage products accumulate over time caused by ongoing proteolysis of full-length myoglobin by FP-2 (Fig. 3.28, 10 to 30 min). Further degradation of primary cleavage products is evident only after most of the full-length substrate has been processed (Fig. 3.28, 60 to 120 min), suggesting that FP-2 also cleaves myoglobin in a sequentially ordered way.

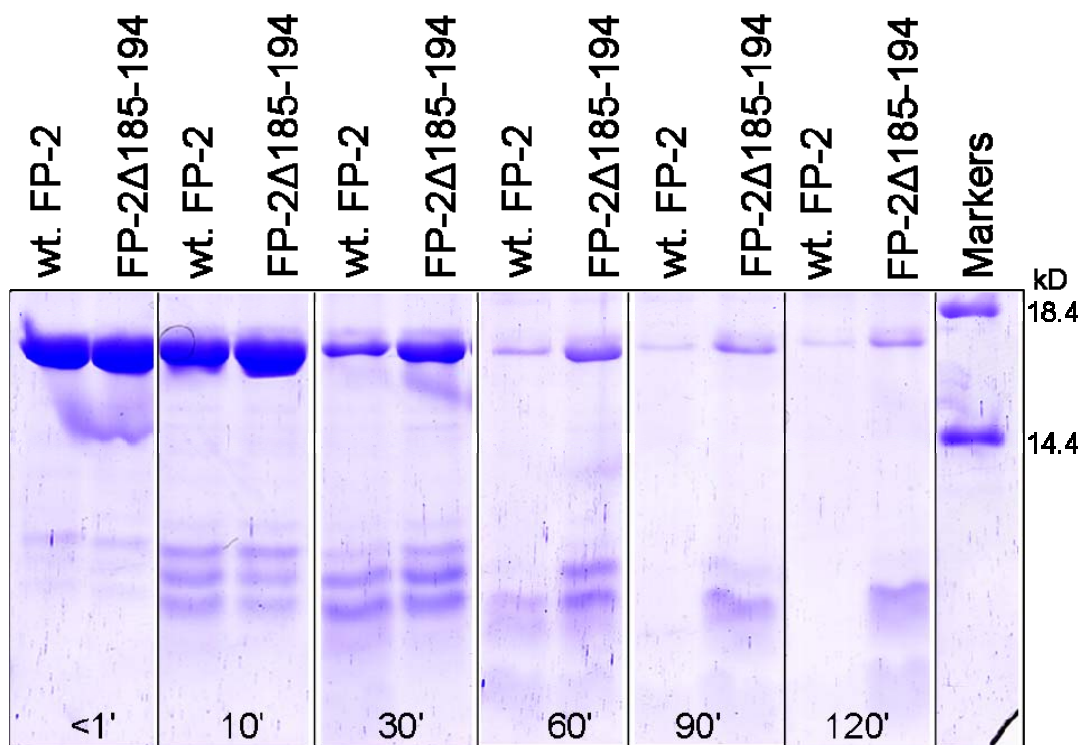
## RESULT AND DISCUSSION



**Fig 3.28:** pH-dependent catalytic activity of FP-2 with myoglobin. 0.02 mg/ml of refolded and purified, active, mature FP-2 was incubated with 0.2 mg/ml of myoglobin at various pH values ranging from pH 5.0 to 8.0. The samples were incubated at 37 °C for the indicated times (<1' to 120'). Samples were subjected to SDS polyacrylamide gel electrophoresis. In a control experiment, myoglobin was incubated at pH 5.0 in the absence of the active FP-2. No activity was observed with samples incubated at pH 7.0 and 8.0.

To further investigate the interaction of myoglobin with FP-2, surface plasmon resonance studies have been carried out using inactive mature FPc285aM to prevent degradation of the substrate. This analysis clearly indicated a specific interaction between FP-2 and immobilized myoglobin with an optimum in the acidic pH range (pH 5.0 to pH 6.0). An association constant could not be determined because of apparent unusual kinetics of FP-2 : myoglobin binding. Nevertheless, the results clearly indicated that the reduced proteolytic activity of FP-2 with myoglobin as substrate at higher pH is due to decreased substrate binding, as shown earlier for hemoglobin (Fig. 3.23).

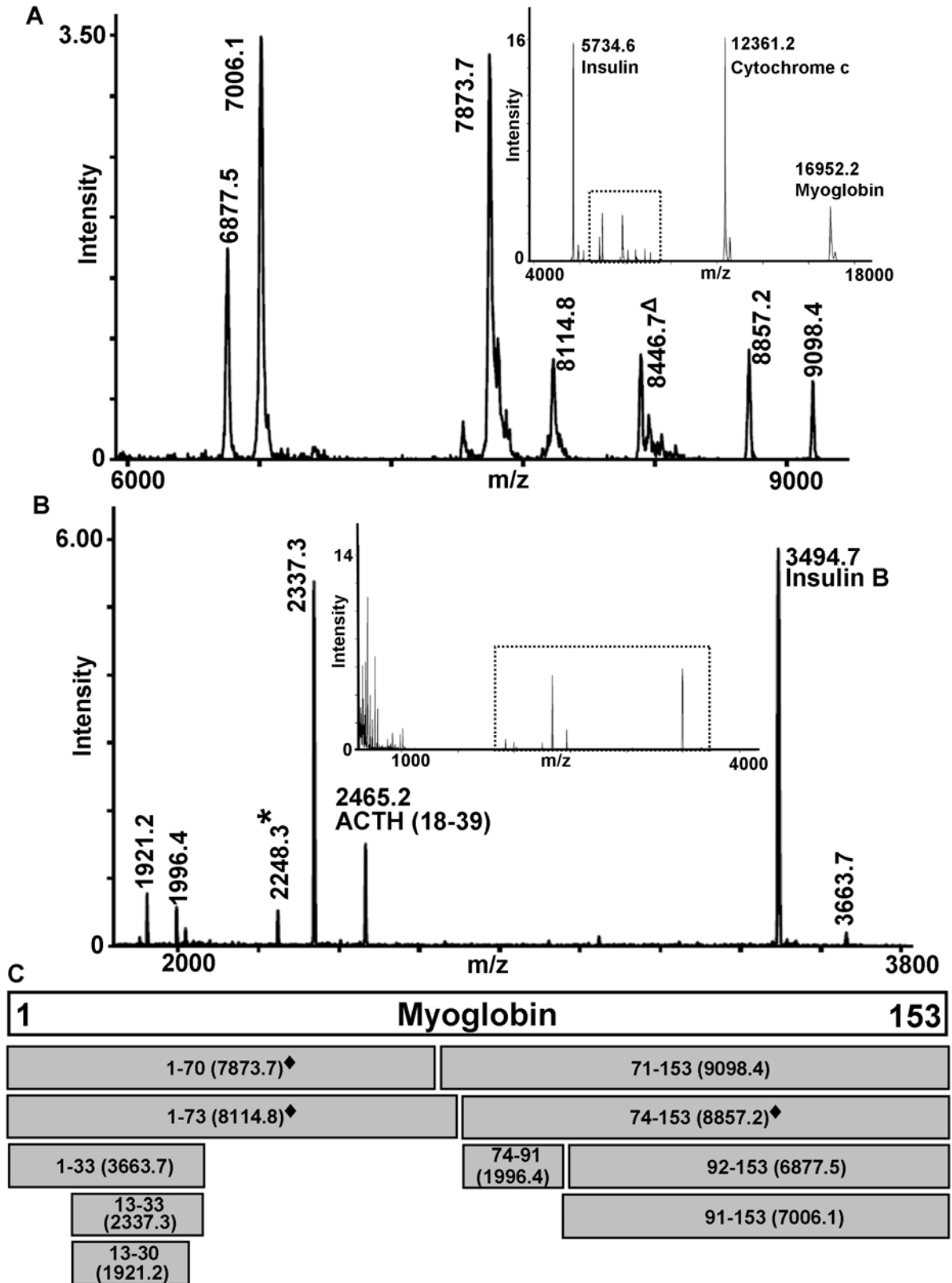
**3.8.3 Proteolytic cleavage of myoglobin by FP-2 $\Delta$ 185-194** - The myoglobin degradation assay by the mutant further supports the idea that the unique motif is not involved in binding. Like the activity shown with hemoglobin (Fig. 3.24), FP-2 $\Delta$ 185-194 was readily active against myoglobin. The different time-interval samples from wild-type and the mutant were loaded in adjacent lanes (Fig. 3.29). The protease activity of the mutant is very much comparable to that of the active mature wild-type FP-2 with marginal differences. Alike the wild-type, hydrolysis advances in a step-wise manner producing relatively stable intermediates (Fig. 3.29, 10 min). Further processing of these metastable intermediates is clearly visible (Fig. 3.29, 60 min). Within 120 minutes of incubation, the wild-type completely hydrolyzed myoglobin, while the intermediate bands were visible to some extent in the case of the mutant.



**Fig. 3.29:** Degradation of myoglobin by wild-type and FP-2 $\Delta$ 185-194 mutant. The cleavage is highly specific and occurs in an ordered fashion with metastable intermediates.

**3.8.4 Analysis of myoglobin primary cleavage products** - Cleavage products of myoglobin obtained at various time points were subjected to MALDI-TOF mass-spectrometric analysis. When measured in higher molecular weight range (4000 to 18000), multiple cleavage products with masses 6877.5, 7006.1, 7873.7, 8114.8, 8446.7, 8857.2, and 9098.4 were detected in the MALDI-TOF spectra recorded from samples that had been incubated for 10 min or less (Fig. 3.30A). These cleavage products compare well to the results obtained from the SDS-PAGE experiments (Fig. 3.28), allowing for the fact that small mass differences observable by MALDI-TOF mass spectrometry cannot be resolved in the gel experiments. All fragments generated by these primary cleavages could be annotated (Fig. 3.30C). The peak corresponding to the mass 8446.7 represents the doubly ionized myoglobin, whereas the other six peaks can be attributed to peptides produced by the cleavage of myoglobin by FP-2.

Two pairs of these initial peptide fragments can be arranged in-line, covering the whole sequence of myoglobin as shown in Fig. 3.30C: One of the primary cleavage sites is located in helix E between residues Thr70 and Ala71, producing the peptide fragment pair 1-70 (7873.7) and 71-153 (9098.4) (Fig. 3.30C). A second primary cleavage site between Gly73 and Gly74 is located in the same helix just three residues apart from the first site. Cleavage at this site produces a second peptide fragment pair 1-73 (8114.8) and 74-153 (8857.2) (Fig. 3.30C). Additional primary cleavages occur between Ala90 and Gln91, Gln91 and Ser92 in helix F, producing the peptide fragments 92-153 (6877.5) and 91-153 (7006.1) respectively. Due to efficient degradation of peptides derived from the N-terminus of myoglobin, continuous fragments corresponding to segments 1-90 or 1-91 could not be detected by mass-spectrometry.



**Fig. 3.30:** MALDI-TOF spectrum of primary myoglobin degradation products obtained from samples incubated for 10 min or less. The marker protein masses are also denoted **(A)** MALDI-TOF analysis was carried out in the molecular mass range 4000 to 18000 (inset). Boxed region of the inset highlights the area of the spectrum in detail. Double ionized myoglobin peak is indicated by  $\Delta$ . **(B)** MALDI-TOF analysis carried out in <4000 molecular mass range (inset). The peak mass that could not be mapped to one specific peptide fragment is denoted with an asterisk. **(C)** Boxed representation of primary peptide fragments. The peptide fragments identified by  $\text{NH}_2$ -terminal sequencing are denoted by ♦.

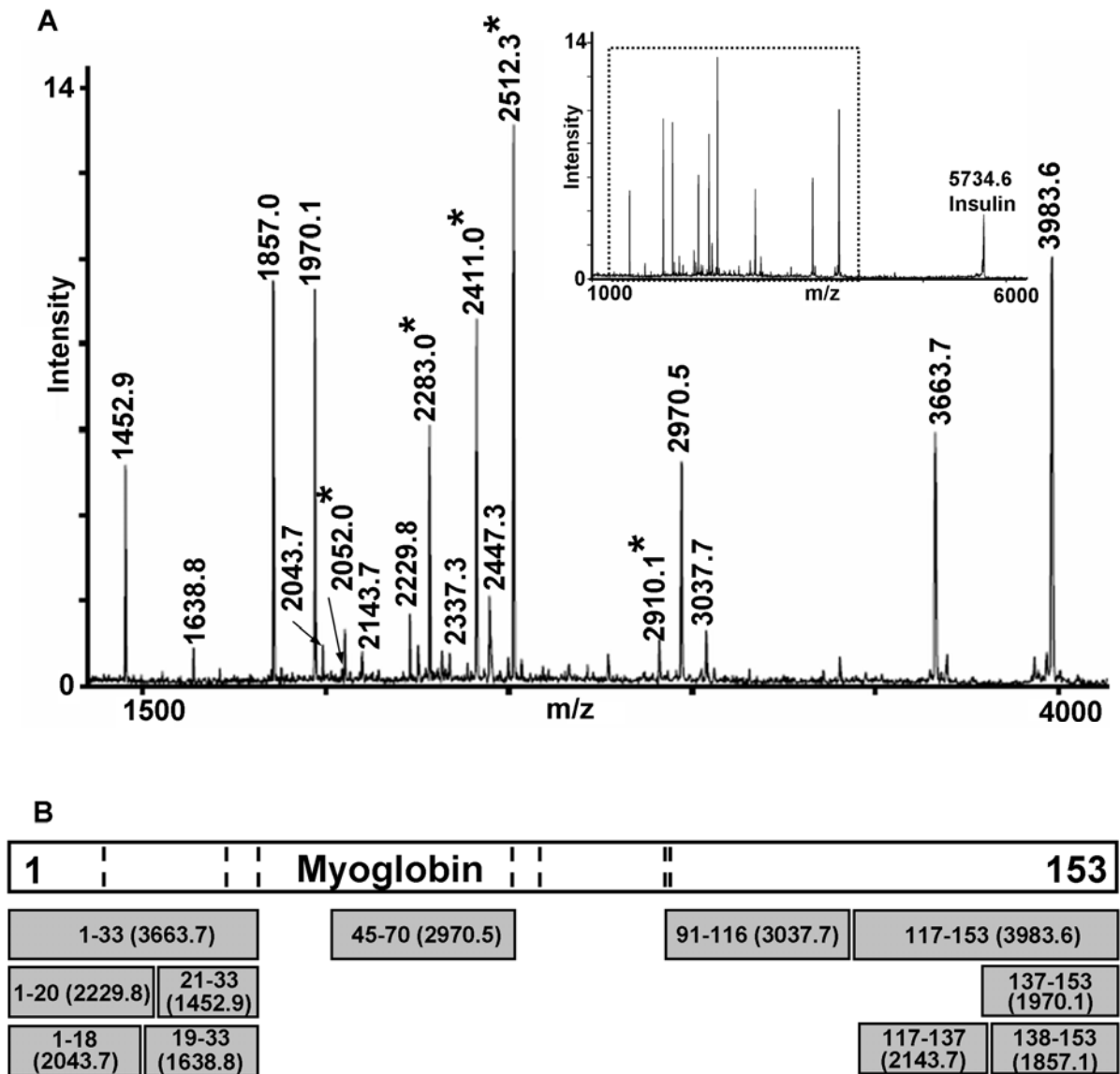


## RESULT AND DISCUSSION

However, a MALDI-TOF spectrum recorded in the lower molecular mass range (<4000; Fig. 3.30B) revealed additional peaks with masses corresponding to 1921.2, 1996.4, 2248.3, 2337.3, and 3663.7. These masses were mapped to the myoglobin sequence and as expected could be attributed to further degradation of the large primary cleavage products within 10 min incubation with FP-2. These cleavages occur mostly in the N-terminal region of myoglobin. The mass 3663.7 corresponds to the peptide fragment 1-33, which is generated by cleavage between residues Phe33 and Thr34 in helix B. Subsequently, this peptide is subjected to further cleavage between residues Asn12 and Val13 in helix A, producing the peptide fragment 13-33 (2337.3), which in turn undergoes further cleavage between Ile 30 and Arg 31 in helix B producing the peptide fragment 13-30 (1921.2). In contrast, only one additional cleavage site in peptides derived from the C-terminus of myoglobin could be identified. The product with a mass of 1996.4 corresponds to the peptide fragment 74-91, which was generated by further processing of the peptide fragment 74-153 (cleavages between Gly73 and Gly74; Gln91 and Ser92). The peak with the mass of 2248.3 (marked by an asterisk in Fig. 3.30B) could not be annotated since too many potential peptide fragments of myoglobin with very little differences in their masses potentially correspond to this signal.

A MALDI-TOF spectrum including the low-mass range (<1000) showed many additional peaks (Fig. 3.30B, inset) corresponding to peptide fragments of less than 10 amino-acid residues. Unfortunately, these small degradation products could not be unambiguously assigned to myoglobin cleavage products and therefore the identification of the corresponding cleavage sites was not possible. However, it is likely that these small peptide fragments are derived from further degradation of the N-terminal fragments 1-12 and 31-73 of myoglobin by FP-2.

**3.8.5 Analysis of secondary cleavage products** - To characterize myoglobin degradation by FP-2 at later stages, samples obtained after 30 minutes incubation of enzyme and substrate, were subjected to MALDI-TOF mass spectrometry. The samples contained many cleavage products in the lower mass range (<4000) corresponding to peptides generated by further processing of primary cleavage products (Fig. 3.31A). The full-range spectrum is shown as an inset (Fig. 3.31A). Almost all the mass products could be unambiguously assigned to myoglobin cleavage products, except peaks corresponding to the masses 2052.0, 2283.0, 2411.0, 2512.3, 2910.1 (Fig. 3.31A, marked with asterisks). As expected, peaks in the very low molecular mass range (< 1000) could not be annotated due to the fact that too many possible peptide fragments of myoglobin could be matched to the observed masses. Fig. 3.31B shows the mapping of the secondary cleavage products to their respective myoglobin fragments.



**Fig. 3.31:** MALDI-TOF spectrum of secondary myoglobin degradation products obtained from samples incubated for 30 min. **(A)** MALDI-TOF analysis was carried out in the molecular mass range <4000 (inset). Boxed region of the inset highlights the area of the spectrum in detail. The peak mass that could not be mapped to one specific peptide fragment is denoted with an asterisk. **(B)** Boxed representation of secondary peptide fragments.

It is interesting to note that the peak corresponding to 3663.7 (peptide fragment 1-33) is much more prominent in the spectrum after 30 min incubation than in the spectrum recorded from samples after 10 min incubation (Fig. 3.30B), suggesting that this degradation product could be of late origin and accumulating with time. This peptide was subjected to further cleavages at two specific sites: cleavage between residues

Asp20 and Ile21 in helix B, leading to peptide fragments 1-20 (2229.8) and 21-33 (1452.9), and cleavage between Glu18 and Ala19, leading to peptide fragments 1-18 (2043.7) and 19-33 (1638.8). Further secondary cleavage sites are observed between residues Asp44 and Lys45, producing peptide fragment 45-70 (2970.5), as well as in primary peptide fragments originating from the C-terminal region of myoglobin (Fig. 3.31B). The primary cleavage product corresponding to residues 91-153 (7006.1 Da) was degraded into several smaller peptides. Cleavage between His116 and Ser117 located at the C-terminal end of helix G produced the two peptides 91-116 (3037.7) and 117-153 (3983.6). The fragment 117-153 is further cleaved between residues Lys137 and Phe138, producing the peptides 117-137 (2143.7) and 138-153 (1857.0). Cleavage between Glu 136 and Lys 137 produced the fragment 137-153 (1970.1) as shown in Fig. 3.31B. The corresponding cleavage product covering the sequence range 117-136 was not detected.

**3.8.6 *NH<sub>2</sub>-terminal amino-acid analysis*** - Automated Edman degradation was used to determine the N-terminal sequences of the cleavage products extracted from SDS gels. Two N-terminal sequences could be identified (the corresponding peptide fragments are highlighted in Fig. 3.30C). One of the sequences (Gly1-Leu2-Ser3) corresponds to the N-terminus of native myoglobin and possibly originates from the primary cleavage products (1-70, 1-73, and 1-33). In the gel experiments (Fig. 3.28), cleavage products in the mass range 7000-8000 do occur which might well correspond to peptide fragments 1-70 and 1-73. Unlike MALDI-TOF, individual bands of the peptide fragments cannot be observed in gels due to the small mass differences. The second N-terminal sequence (Gly74-Ile75-Leu76) corresponds to one of the intermediate peptide fragments covering the sequence range 74-153 (8857.2 Da). All the cleavage sites identified using N-terminal sequencing were also

detected by MALDI-TOF mass spectrometry. Thus, both lines of investigation produced mutually supporting results.

**3.8.7 Analysis of hemoglobin cleavage products** - Hemoglobin primary and secondary degradation products obtained at various time points were also subjected to mass-spectrometric analysis. The samples that had been incubated for 10 min or less produced multiple cleavage products when measured in higher molecular mass range (4000 to 18000) with masses of 6293.3, 7478.7, 8406.1, 8850.9 (Fig. 3.32A). Some of the primary cleavage products (6182.3, 6566.8) could not be unambiguously annotated (Fig. 3.32A, marked with asterisks). Two pairs of initial peptide fragments can be arranged in-line, covering the whole sequence of hemoglobin  $\alpha$ - and  $\beta$ -chains, as shown in Fig. 3.32C: One of the primary cleavage sites is between residues Ser84 and Asp85 ( $\alpha$ -chain), producing the peptide fragment pair 1-84 (8850.9) and 85-141 (6293.3) (Fig. 3.32C). The other primary cleavage site is located in the  $\beta$ -chain between residues Gly69 and Ala70, producing the peptide fragment pair 1-69 (7478.7) and 70-146 (8406.1) (Fig. 3.32C).

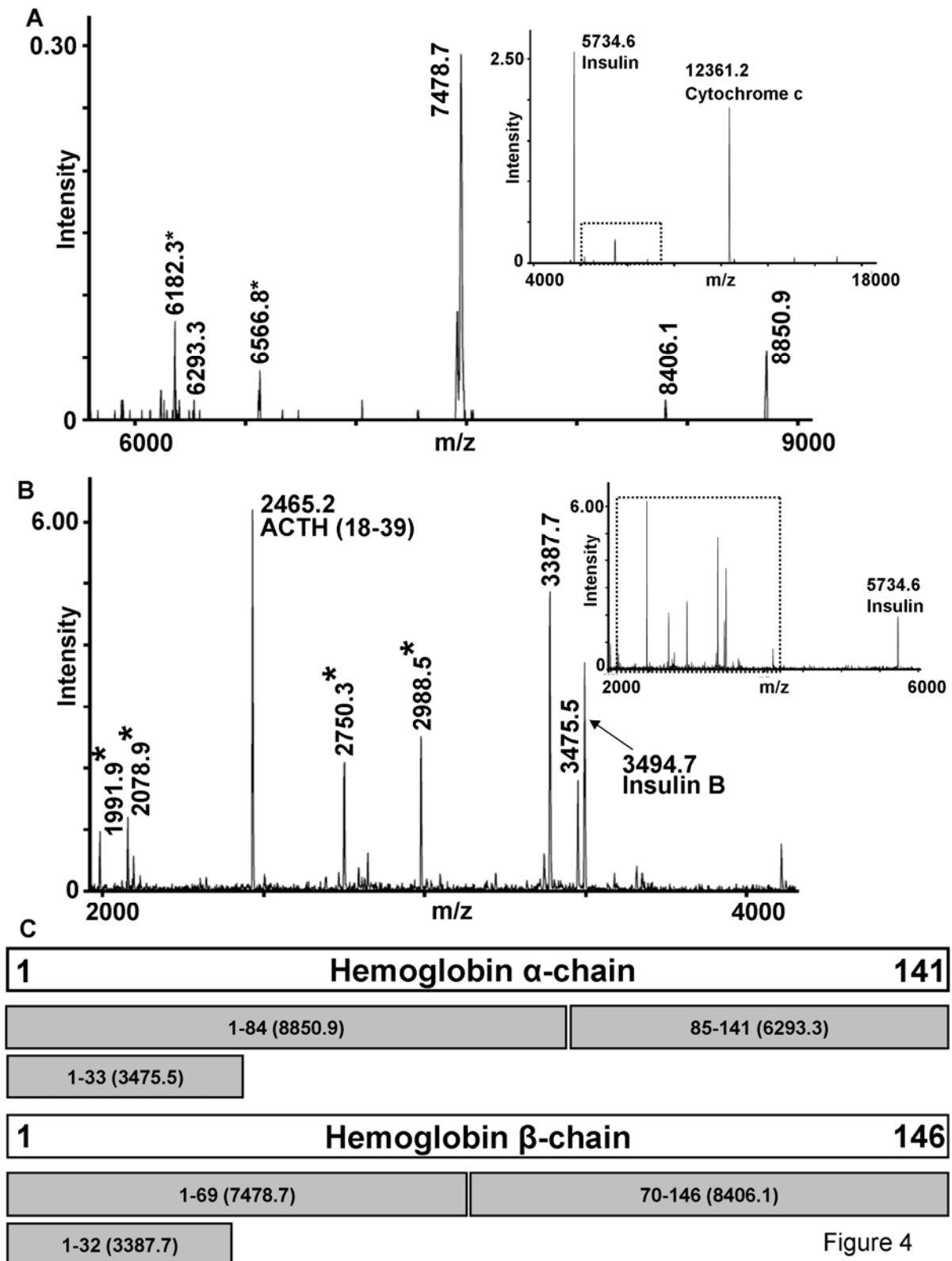


Figure 4

**Fig. 3.32:** MALDI-TOF spectrum of primary and secondary hemoglobin degradation products. The marker protein masses are also denoted. The peak mass that could not be mapped to one specific peptide fragment is denoted with an asterisk. **(A)** MALDI analysis was carried out in the molecular mass range 4000 to 18000 (inset). **(B)** MALDI analysis carried out in <4000 molecular mass range (inset). **(C)** Boxed representation of peptide fragments identified in the both chains of hemoglobin.

## RESULT AND DISCUSSION

A second MALDI-TOF spectrum was recorded for samples incubated for 10 min or less in the lower molecular mass range (<4000; Fig. 3.32B). The spectrum revealed additional peaks with masses of 1991.9, 2078.9, 2750.3, 2988.5, 3387.7 and 3475.5. Two masses could be annotated unambiguously (3387.7 and 3475.5) and once mapped to the hemoglobin sequences, these peptide fragments could be attributed to further degradation of the large primary cleavage products. The mass 3475.5 corresponds to the peptide fragment 1-33 (Fig. 3.32C) with the cleavage site between residues Phe33 and Leu34 ( $\alpha$ -chain). The other mass, 3387.7 corresponds to the peptide fragment 1-32 (Fig. 3.32C) with the cleavage site between residues Leu32 and Val33 ( $\beta$ -chain). As for the other observed masses that could not be annotated (Fig. 3.31b, marked with asterisks), too many potential peptide fragments with very little differences in their masses both from  $\alpha$ - and  $\beta$ -chain hemoglobin potentially correspond to this signal.

**3.8.8 Analysis of FP-2 substrate specificity** - The sequences of the seven primary and the six secondary cleavage sites that could be identified by MALDI-TOF are compared in Fig. 3.33A and Fig. 3.33B, respectively. The positions P1 and P1' show no readily recognizable preferences for specific residues. However, it is noticeable that none of the primary cleavage sites contains a charged residue at P1 (Fig. 3.33A), whereas 5 out of 6 of the secondary cleavage sites do prefer a hydrophobic residue at this position (Fig. 3.33B). The primary cleavage sites show a strong preference for leucine in P2 (6 out of 7 sites) (Fig. 3.33A). This preference appears to be less distinctive for the secondary cleavage sites, which can contain other hydrophobic residues and in one case even a glutamate residue at this position (Fig. 3.33B). FP-2 shows a preference for a hydrophobic residue at the P3 position, which is dominated by valine, leucine or alanine, whereas charged residues at this position are more frequent among the secondary cleavage sites (3 out of 6 sites) (Fig. 3.33B). It is noteworthy that products resulting from one of the primary cleavage sites (Phe33 and Thr34) contain an Arg at P3 (Fig. 3.33A). This cleavage site is already detectable within 10 minutes (Fig. 3.30B) but appears much more prominent after longer reaction times (Fig. 3.31A). Hence, this site may be cleaved at an intermediate rate. The positions P4 and P1' to P4' show no recognizable preferences for specific residues.



A Primary cleavage sites									
P4	P3	P2	P1	↓	P1'	P2'	P3'	P4'	
V	V	L	T <sub>70</sub>	↓	A <sub>71</sub>	L	G	G	
T	A	L	G <sub>73</sub>	↓	G <sub>74</sub>	I	L	K	
K	P	L	A <sub>90</sub>	↓	Q <sub>91</sub>	S	H	A	
P	L	A	Q <sub>91</sub>	↓	S <sub>92</sub>	H	A	T	
I	R	L	F <sub>33</sub>	↓	T <sub>34</sub>	G	H	P	
E	V	L	I <sub>30</sub>	↓	R <sub>31</sub>	L	F	T	
Q	V	L	N <sub>12</sub>	↓	V <sub>13</sub>	W	G	K	

B Secondary cleavage sites									
P4	P3	P2	P1	↓	P1'	P2'	P3'	P4'	
H	V	L	H <sub>116</sub>	↓	S <sub>117</sub>	K	H	P	
K	A	L	E <sub>136</sub>	↓	L <sub>137</sub>	F	R	N	
A	L	E	L <sub>137</sub>	↓	F <sub>138</sub>	R	N	D	
E	K	F	D <sub>44</sub>	↓	K <sub>45</sub>	F	K	H	
V	E	A	D <sub>20</sub>	↓	I <sub>21</sub>	A	G	H	
G	K	V	E <sub>18</sub>	↓	A <sub>19</sub>	D	I	A	

C									
Y	L	L	D <sub>1</sub>	↓	Q <sub>1</sub>	M	N	Y	FP-2 auto-processing of pro-enzyme
N	Y	L	G <sub>124</sub>	↓	S <sub>125</sub>	S	N	D	FP-2 processing site of pro-PM II

**Fig 3.33:** Subsite analysis of primary and secondary cleavage products. The primary and secondary cleavages sites are tabulated separately. In each case, the residues in myoglobin spanning from the subsite position P4 to P4' are denoted. **(A)** In total, seven primary cleavages could be mapped onto the myoglobin sequence. The relative positions of the cleavage sites are marked with an arrow. It is clear that FP-2 prefers leucine shaded in grey at the P2 position (6/7 cleavages). **(B)** Six secondary cleavage sites are tabulated. It can be shown that mostly hydrophobic residues occur at P2 with some exceptions (4/6 cleavages). But it is observed that the preference for leucine at P2 is greatly diminished. **(C)** FP-2 auto-processing site and the processing of Pro-PM II by FP-2 are shown. Strong preference for leucine at the P2 position is observed.

It is interesting to note that most of the cleavage sites occur within helices. Cleavage at Thr70 and Ala71, Gly73 and Gly74 occurs more or less in the middle of helix E (residues 58 to 77). Similarly, other primary cleavage sites are also located within helices (F, B, and A). Most of the secondary cleavage sites do occur within helices as well (G, H and A), with the exception of Asp44 and Lys45 (inter-helical region between helices C and D) and Glu18 and Ala19 (between helices A and B).

As reported previously (Shenai et al., 2000; Fig. 3.3), the FP-2 precursor undergoes autoproteolytic processing, thereby producing the mature active FP-2. The autoproteolytic site also contains a leucine at the P2 position (Fig. 3.33C).

Prodomain processing of PM II (Drew et al., 2008) by FP-2 also exhibits a similar preference. Similarly, one of the initial cleavage sites occurring in the  $\alpha$ -chain and two of the cleavage sites in the  $\beta$ -chain of hemoglobin show a high preference for a leucine residue at P2 as shown in Fig. 3.34 (marked with section signs §).

### **3.9 Structural significance of cleavage products -**

**3.9.1 Significance of FP-2 subsite preferences -** To extend the studies of the underlying mechanism of activity of FP-2, cleavage products from various time points were studied using MALDI-TOF MS. All the major cleavage sites could be annotated and mapped to the structure of myoglobin. Most of these cleavages do occur within helices B, E and F (Fig. 3.35). The primary peptide fragments undergo further specific cleavages producing a few more secondary cleavage products (Fig. 3.31). Comparable results were also obtained for hemoglobin from MALDI-MS experiments. Due to the more complex quaternary structure ( $\alpha_2\beta_2$ ), only primary cleavages could be unequivocally identified. Similar to myoglobin cleavage sites, the preference for leucine at P2 and a hydrophobic residue at P3 is clearly observable (Fig. 3.34).

Previous studies revealed that the interaction of the P2 residue of the substrate with its complementary S2 pocket of the enzyme is the major determinant of substrate specificity for most cysteine proteases of the papain family (Brömme et al., 1994; Bode & Huber, 2000). Earlier reports on substrate specificity of FP-2 using peptide substrates showed a strong preference for a hydrophobic residue at P2, favorably a leucine (Shenai et al., 2000; Sijwali et al., 2001b; Ramjee et al., 2006). Biochemical studies have shown that FP-2 cleaves a wider range of substrates in accordance with its role in other cellular functions and exhibits with significantly higher specific activity than FP-3 (Ramjee et al., 2006; Sijwali et al., 2001b). FP-2 cleaves the intra-erythrocyte membrane skeletal proteins ankyrin (VSAA↓PTLA)

## RESULT AND DISCUSSION

(Dhawan et al., 2003) and band 4.1 (GGIL↓LHHA) (Hanspal et al., 2002), both containing a hydrophobic residue at the P2 position. Similarly, the FP-2 autoprocessing of its precursor and precursor processing site of PM II also contain a leucine at P2 (Fig. 3.33C). All these results agree well with the cleavage specificities observed with myoglobin as substrate. Six out of seven of the primary cleavage sites contain leucine at P2, the remaining one contains alanine at this position. Studies with synthetic peptides by Ramjee and colleagues (Ramjee et al., 2006) showed that, along with Leu at P2, the presence of Glu or especially Arg at P1 greatly enhances the cleavage rates.

Unlike primary cleavage sites, the secondary FP-2 cleavage sites on myoglobin showed no clear preference for leucine at P2 but rather a general preference for hydrophobic residues (Fig. 3.33B). High preference for charged residues at the P1 site was observed. Five out of six of the secondary cleavage sites contain a charged residue (His, Glu, Asp) at the P1 position (Fig. 3.33B). The sequence around the autoproteolytic cleavage site of FP-2 contains an Asp at P1 and a Leu at P2 (Fig. 3.33B), thus exhibiting similarities to the secondary cleavage sites observed in myoglobin.

Similar preference for a charged residue at the P1 position has been previously reported for falcilysin by Eggleston et al. (1999). Drew et al. (2008) showed that a functional overlap exists among the plasmodial proteases, namely the plasmepsins and the falcipains. This study on hemoglobin and myoglobin also suggest that FP-2 may be involved in downstream degradation of hemoglobin and that some degree of functional overlap might exist between falcipains and falcilysin.

Both the primary and secondary cleavage sites show a preference for a hydrophobic residue at the P3 position, which is dominated by valine, leucine or alanine with some exceptions where lysine and arginine do occur (Fig 3.33).

**3.9.2 Prediction of hemoglobin additional cleavage sites** - Due to the similarities of the structures, the myoglobin structure (PDB ID: 1YMB; Evans & Brayer, 1990) can be superimposed onto both the  $\alpha$ - (r.m.s.d. 1.49 Å) and  $\beta$ - (r.m.s.d. 1.67 Å) chains of hemoglobin (PDB ID: 1BZ0; Kavanaugh et al., 1993). The primary and the secondary cleavage sites from the myoglobin study were mapped onto the hemoglobin structure and the residues corresponding to the cleavage sites were inspected (Fig. 3.34). This revealed that all four primary cleavage sites observed in hemoglobin using MALDI-TOF coincide with primary cleavage sites observed in myoglobin (Fig. 3.34, marked with bullets). Thus hemoglobin and myoglobin are likely to be cleaved at structurally equivalent positions by FP-2, strongly supporting the use of myoglobin as a hemoglobin substitute model. Based on these results, the additional primary and secondary cleavage sites found for myoglobin were mapped onto the hemoglobin structure. The predicted cleavage sites for both the  $\alpha$ - and  $\beta$ -chain of hemoglobin are listed in Fig. 3.34 in the same sequential order as described for myoglobin in Fig. 3.33. It is noteworthy that all predicted primary cleavage sites are located within solvent exposed regions of the intact hemoglobin tetramer and thus should be easily accessible for the protease. It is not surprising that the cleavages occur within helices of both chains. These predicted cleavage sites show high preference for hydrophobic residues at the P2 position (8/13 sites in  $\alpha$ -chain, 9/13 sites in  $\beta$ -chain), especially favoring leucine (4/13 sites in  $\alpha$ -chain, 5/13 sites in  $\beta$ -chain).

(a): Myoglobin	(b): Hemoglobin α-chain	(c): Hemoglobin β-chain
<b>Bond cleaved</b>	<b>Cleavage sequence</b>	<b>Cleavage sequence</b>
T <sub>70</sub> ↓ A <sub>71</sub>	KVAD <sub>64</sub> ↓ A <sub>65</sub> LTN	<u>KV</u> <u>L</u> G <sub>69</sub> ↓ A <sub>70</sub> FSD <sub>§</sub> ←
G <sub>73</sub> ↓ G <sub>74</sub>	ALTN <sub>68</sub> ↓ A <sub>69</sub> VAH	GAFS <sub>72</sub> ↓ D <sub>73</sub> GLA
A <sub>90</sub> ↓ Q <sub>91</sub>	LSAL <sub>83</sub> ↓ S <sub>84</sub> DLH	ATLS <sub>89</sub> ↓ E <sub>90</sub> LHC
Q <sub>91</sub> ↓ S <sub>92</sub>	SALS <sub>84</sub> ↓ D <sub>85</sub> LHA <sub>§</sub> ←	TLSE <sub>90</sub> ↓ L <sub>91</sub> HCD
F <sub>33</sub> ↓ T <sub>34</sub>	ERMF <sub>33</sub> ↓ L <sub>34</sub> SFP ←	<u>GR</u> <u>L</u> L <sub>32</sub> ↓ V <sub>33</sub> VYP <sub>§</sub> ←
I <sub>30</sub> ↓ R <sub>31</sub>	EAL <sub>30</sub> ↓ R <sub>31</sub> MFL	EALG <sub>29</sub> ↓ R <sub>30</sub> LLV
N <sub>12</sub> ↓ V <sub>13</sub>	NVKA <sub>12</sub> ↓ A <sub>13</sub> WGK	<u>AV</u> <u>T</u> A <sub>13</sub> ↓ L <sub>14</sub> WGK
H <sub>116</sub> ↓ S <sub>117</sub>	VTLA <sub>110</sub> ↓ A <sub>111</sub> HLP	CVLA <sub>115</sub> ↓ H <sub>116</sub> HFG
E <sub>136</sub> ↓ L <sub>137</sub>	KFLA <sub>130</sub> ↓ S <sub>131</sub> VST	KVVA <sub>135</sub> ↓ G <sub>136</sub> VAN
L <sub>137</sub> ↓ F <sub>138</sub>	FLAS <sub>131</sub> ↓ V <sub>132</sub> STV	VVAG <sub>136</sub> ↓ V <sub>137</sub> ANA
D <sub>44</sub> ↓ K <sub>45</sub>	FPHF <sub>46</sub> ↓ D <sub>47</sub> LSH	FESF <sub>45</sub> ↓ G <sub>46</sub> DLT
D <sub>20</sub> ↓ I <sub>21</sub>	KVGA <sub>19</sub> ↓ H <sub>20</sub> AGE	GKVN <sub>19</sub> ↓ V <sub>20</sub> DEV
E <sub>18</sub> ↓ A <sub>19</sub>	WGKV <sub>17</sub> ↓ G <sub>18</sub> AHA	LWGK <sub>17</sub> ↓ V <sub>18</sub> NVD

**Fig. 3.34:** Cleavage sites of myoglobin extrapolated to hemoglobin sequences. The myoglobin structure (PDB-ID: 1YMB) (Evans & Brayer, 1990) was superposed onto hemoglobin structure (PDB-ID: 1BZ0) (Kavanaugh et al., 1993). The cleavage sites on myoglobin are considered to be potential cleavages in hemoglobin. The cleavages are tabulated according to the primary and secondary cleavages observed in myoglobin. The cleavage sequence spanning from P4 to P4' is denoted. **(A)** The myoglobin cleavage sites including the primary and secondary cleavages with the respective residue numbering. **(B)** Hemoglobin α-chain sequences, whose respective bonds should be cleaved according to the superposition of the structures. It should be noted that the extrapolated cleavage sites also show high preference for hydrophobic residues at P2, especially leucine (denoted by §). The occurrence of leucine at P2 is shaded in grey (4/13 cleavages). **(C)** Cleavage sequences in hemoglobin β-chain. The preference for leucine at P2 is shaded in grey (5/13 cleavages). Grey bullets denote the cleavage sites observed using MALDI-TOF. The underlined cleavage sites in both the chains are those that match with the results from Gluzman et al., 1994 and Kolakovich et al., 1997.

Previous studies on hemoglobin degradation by plasmodial proteases have been conducted by Gluzman et al. (1994) and Kolakovich et al. (1997). Gluzman et al. (1994) isolated an active cysteine protease from the vacuoles of trophozoitic-stage *P. falciparum* and characterized some of its cleavage products. Like the falcipains 2, 2B and 3, this protease activity displayed a pH optimum in the acidic range. However, this cysteine protease was able to cleave globin but not native hemoglobin (Gluzman

## RESULT AND DISCUSSION

et al. 1994), contradictory to what was demonstrated for FP-2 (Salas et al., 1995; Shennai et al., 2000). Since the authors did not report special precautions to avoid the oxidation of the heme group, the substrate can be expected to consist at least partly of methemoglobin, which was readily degraded in our studies (Fig. 3.22). Hence, the identity of the cysteine protease isolated by Gluzman et al. (1994) seems unclear.

Kolakovich and colleagues (1997) studied and characterized the cleavage products obtained by incubation of hemoglobin with lysates from plasmodial digestive vacuoles. While both studies identified several cleavage peptides, the nature of the enzyme that generated these peptides remained ambiguous. In their paper, they refer to the cysteine protease described by Gluzman et al. (1994) as “falcipain”. It is most certainly correct that this preparation contained a falcipain. But, these results cannot be used to characterize the activity of one individual protease of the three known hemoglobin-degrading cysteine proteases known to be present within the food vacuole (Shenai et al., 2000; Sijwali et al., 2001a; Singh et al., 2006; Sim et al., 2001; Chan et al., 2005). However, the inability of this preparation to degrade native intact hemoglobin suggests that FP-2 was not the main component. Despite this uncertainty, these previous studies identified several peptides resulting from cleavages of hemoglobin at sites predicted from the myoglobin model (Sequences underlined in Fig. 3.34). This match of the predicted and observed cleavage sites provides further support to the conclusions drawn based on the current studies using myoglobin as a model substrate. The prediction suggest that some cleavages previously attributed only to plasmepsins can also occur due to the enzymatic activity of FP-2 (cleavage sites between the residues Asn68 and Ala69 in the  $\alpha$ -chain; Ala13 and Leu14, Asn19 and Val20 in the  $\beta$ -chain). A comparison with the cleavage sites reported for the metallo-protease falcilysin (Eggleson et al., 1999)

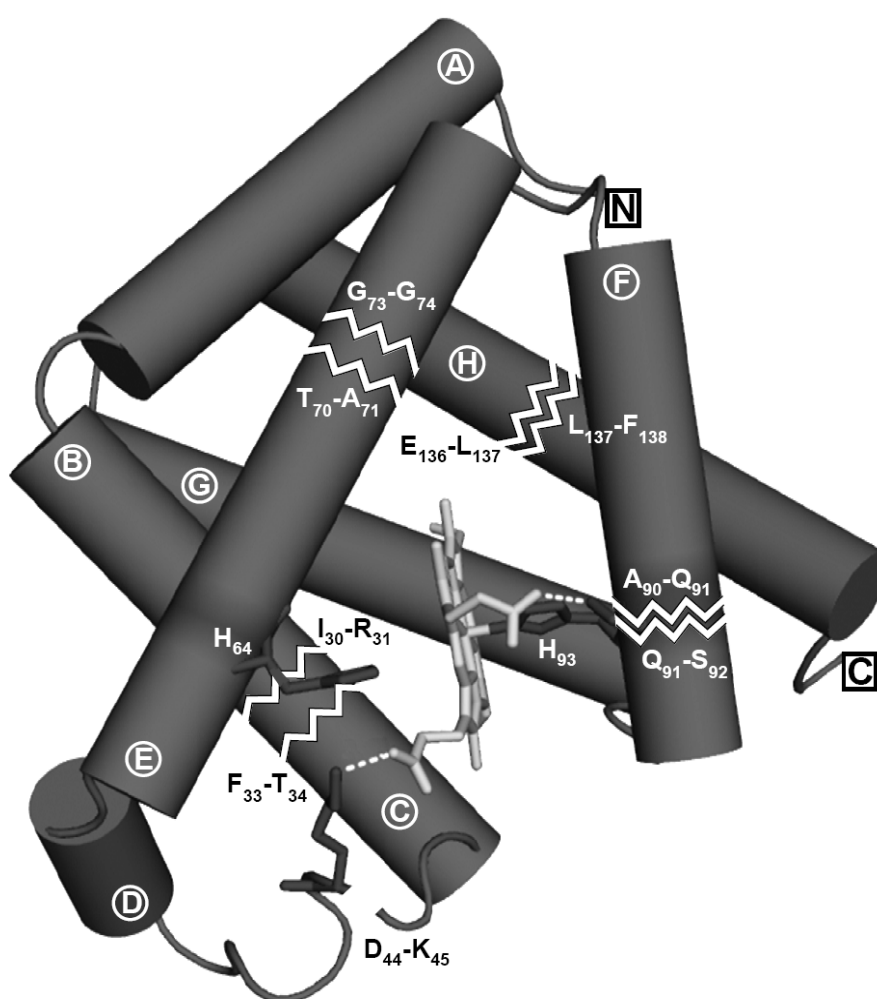
and these results revealed two matching cleavage sites, between Phe33 and Leu34, Leu32 and Val33 of the  $\alpha$ - and  $\beta$ -chain respectively.

**3.10 Role of FP-2 in heme detoxification** - Free heme is toxic to the malaria parasites (Orjih et al., 1981; Fitch et al., 1982). It is highly important for the development of the erythrocyte stages of the *Plasmodia* that heme is released from hemoglobin and detoxified as soon as possible. The subsequent globin hydrolysis provides free amino acids that are used for parasitic protein synthesis and to maintain the osmotic balance inside the erythrocytes (Rosenthal & Meshnick, 1996; Lew et al., 2003). While most of the free heme in the food vacuole is detoxified by polymerization to hemozoin, heme processing may also provide the iron required for synthesis of iron-containing plasmodial proteins such as ribonucleotide reductase (Rubin et al., 1993), superoxide dismutase (Becuwe et al., 1996), cytochromes (Fry & Beesley, 1991) and for parasitic heme biosynthesis (Surolia & Padmanaban, 1991, 1992). Even though the source of iron is uncertain (Gordeuk et al., 1994), considering the massive degradation of hemoglobin, this process has to be regarded as a principal source of iron for the parasite.

Previous reports showed that the cysteine protease inhibitor E-64, but not the aspartic protease inhibitor pepstatin, nor chloroquine or artemisinin can block the formation of hemozoin by cultured parasites (Asawamahasakda et al., 1994; Gamboa de Domínguez et al., 1996). These results suggest that cysteine proteases, most likely FP-2 along with other falcipains, play a major role in hemoglobin degradation, which facilitates the separation of heme moieties from hemoglobin.

The results from the studies with myoglobin as model substrate permit to suggest a mechanism of heme release from the globin molecule. Though many solvent-exposed leucines occur in the myoglobin structure, the selection of cleavage

sites is highly specific. The location of the primary cleavage sites suggests that the amino-acid residues at the positions P1 to P3 along with the structure of myoglobin including the heme coordination are highly selective for FP-2 degradation activity. Most of the primary cleavages occur in the vicinity of the heme group (Fig. 3.35). It is easily conceivable that even partial unfolding of the protein regions adjacent to these cleavage sites facilitates the release of heme from the globin moiety (Fig. 3.35).



**Fig 3.35:** Systematic removal of heme from globin. The myoglobin structure (PDB-ID: 1YMB) is presented as cylinders and the cleavages that occur in the vicinity of the heme group are mapped onto the structure. The cleavages are denoted by cracks (coloured white). Heme, proximal and distal histidines, and the residues Ser92, Lys45 involved in hydrogen bonding with the heme propionate groups (O2A and O2D) respectively are shown as sticks.



## RESULT AND DISCUSSION

Reports on myoglobin structure by Vallone et al. (2004) and Kachalova et al. (1999) showed that the heme coordination is stabilized by the heme-helixE-helixF system. The E and F helices are holding the heme in its designated position like a forceps (Kachalova et al., 1999). Any disturbances of these helices can displace the backbone segments near the proximal and the distal histidines (His93 and His64; myoglobin PDB ID: 1YMB; Evans et al., 1990), thereby weakening the binding of the heme group. Furthermore, Hargrove and co-workers (Hargrove et al., 1994) reported that the unfolding of helices B, C, and E will lead to complete loss of the heme binding site and entering a molten globule state, where only helices A, G, and H are intact. It is noticed that 6 out of 7 primary cleavages do occur in the helices B, E and F. Similarly, 5 out of 6 secondary cleavages do occur in the A, G, and H helices, respectively. This clearly supports the hypothesis that the specificity of FP-2 has been adapted such that it initially weakens the heme-protein interactions by strategic cleavages and thus facilitates the release of the heme group from the protein.

A comparable location of the primary cleavage sites is also observable in case of hemoglobin. MALDI-TOF analysis of hemoglobin degradation by FP-2 revealed the presence of a primary cleavage site located within the  $\beta$ -chain between Gly69 and Ala70 (Fig. 3.32C); this is one of the cleavage sites also observed in previous studies (Gluzman et al., 1994). Mapped onto the structure, it is obvious that this cleavage site occurs in helix E at equivalent positions in hemoglobin and in myoglobin (between Thr70 and Ala71). Altogether this primary cleavage site appears to be congruent in both substrates, thus validating the use of myoglobin as a model substrate for FP-2.

Similarly, a second major primary cleavage also occurs in helix E between Gly73 and Gly74. Hydrophobic and/or apolar residues (Leu72, Leu76, Ile75; Gly73, Gly74) are positioned in the vicinity of the heme group and are involved in

## RESULT AND DISCUSSION

protecting the heme pocket from the external solvent. Furthermore, Kachalova et al. (1999) showed that even slight movements of the E helix reorient the hydrophobic lining of the heme pocket. Thus, this cleavage might open the heme pocket for further cleavages as well as for water entry and subsequent heme release.

Two additional primary cleavages occur in helix F between residues Ala90 and Gln91, Gln91 and Ser92. Residues Leu89 and Ser92 are involved in hydrogen bonding with the proximal His93. These bonds are responsible for holding the ring plane of His93 close to the heme pyrrole nitrogen atoms (Fig. 3.35). Furthermore, Ser92 is also involved in hydrogen bonding with the propionate groups of the heme (O2A) (Evans & Brayer, 1990). Previous mutational studies by Scott et al. (2001) showed that the Leu89Gly mutation allowed several water molecules to flood the cavity (confined by residues Leu89, Leu104, Phe138) located underneath the porphyrin ring, via the 'hole' created by this substitution. Thus, cleavage sites within the helix F seem to be strategically important and may assist in severing the hydrogen bonds and in turn disorient the His93 and weaken its interaction with the heme. This appears to be an early step in the systematic removal of heme by FP-2.

Some of the secondary cleavages may also contribute to heme release. Lys45 is located on the outer edge of the heme pocket and forms a salt bridge to the carboxyl end of the heme propionate group (O2D) (Evans et al., 1988; Evans & Brayer, 1990). The cleavage between Asp44 and Lys45 (inter-helical region between helix C and D) most likely abolishes this interaction due to the increased flexibility of the protein backbone at the cleavage site. The cleavage between His116 and Ser117 might sever the hydrogen bond between Lys118 and His24, some of the residues that appear to be involved in linking different helical units of the secondary structure together (Evans & Brayer, 1990). The other cleavage sites also seem to aid the heme removal. The cleavage site between Leu137 and Phe138 lies in the vicinity to the

## RESULT AND DISCUSSION

heme pocket. The side chains of these residues are involved in the protection of the heme cavity from external solvent. Thus, many cleavage sites of FP-2 in hemoglobin and myoglobin seem to be defined by a combination of amino-acid sequence and structure in a manner that facilitates the efficient release of the potentially toxic heme group from the protein.

The study based on myoglobin as a model lead to an in-depth understanding of the mechanism of hemoglobin degradation by FP-2. Similarly, studies with myoglobin also demonstrate that FP-2 can cleave an intact, folded protein into discrete peptide fragments (Fig. 3.28). Like in the case of hemoglobin, the degradation of myoglobin proceeds in a specific well-ordered fashion via defined intermediates. The primary cleavages produce relatively stable intermediates, which are further processed in a time-dependent manner leading to complete degradation into short peptides with masses significantly less than 1000 Da within 120 minutes of incubation (Fig. 3.28). Since the pH-dependence of this reaction and the binding affinity are comparable to the observations for hemoglobin (Fig. 3.22), these results favour a 4:1 binding mode in which, four molecules of FP-2 bind to each hemoglobin tetramer, as proposed by Pandey et al. (2005). Furthermore, this study shows that FP-2 can efficiently bind and cleave individual monomeric units. From the alternative docking model proposed by Wang et al. (2006) in which the unique loop or FP-2<sub>arm</sub> motif is placed at the interface between  $\beta$ 1 and  $\alpha$ 2 globins, a significantly lower affinity, or an altered binding mode for a monomeric substrate is expected.

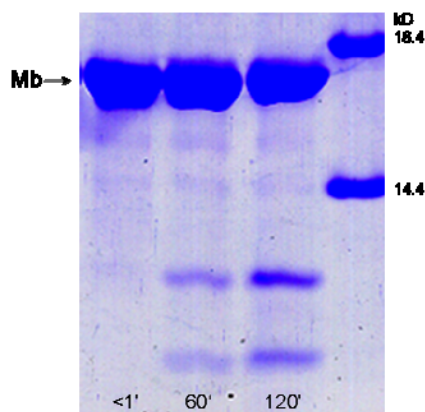
The high coincidence of the cleavage sites between myoglobin and hemoglobin makes myoglobin a perfect study model for hemoglobin degradation by FP-2. It even may permit speculations that myoglobin could be an *in-vivo* substrate for FP-2. Recent studies have shown the association of muscle cell injury as a severe complication of *falciparum* malaria. Sequestration of *P. falciparum* into muscle was

## RESULT AND DISCUSSION

observed in biopsies and the infected patients had increased level of myoglobin in plasma and urine (O'Donnell et al., 2006), indicating damage or demise of muscle cells. Further studies are necessary to ascertain whether in these cases, the muscle cells were damaged by secondary processes, or whether they were infected by *Plasmodia*.

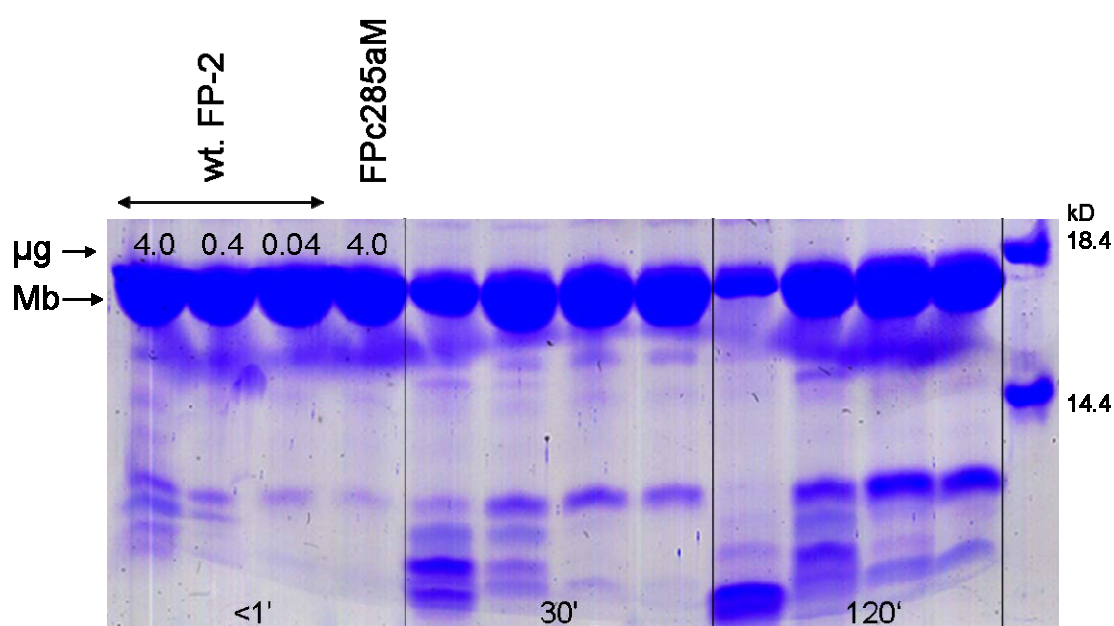
**3.11 Residual activity of the FPc285aM mutant** - The inactive FPc285aM mutant, where the active-site cysteine was mutated to an alanine was generally used as a positive control in the activity assay experiment involving the other active constructs (Fig. 3.26). From the results it was noted that hemoglobin was degraded indicating that FPc285aM also exhibited some minimal activity. The following experiments were conducted in order to verify the existence of protease activity.

**3.11.1 Active-site mutant protease activity** - The substrate myoglobin was incubated with FPc285aM mutant for different time intervals (<1, 60, and 120 min) as shown in Fig. 3.36. Though the mutant is not as efficient as wild-type FP-2, it exhibited somewhat pronounced activity and the degradation products are clearly visible in 60 min and 120 min incubated samples. The degradation bands look alike the intermediate metastable cleavages of the wild-type FP-2 (Fig 3.27), except that no third band was noticed. The accumulation of these intermediate bands from 60 min to 120 min is observable suggesting that the cleavages do occur in a well-specified manner (Fig.3.36, 60 to 120 min). However, the exact reason for this activity is not clear. It can be due to the presence of minute traces of proteases from *E.coli* (expression vector) or may even be due to a secondary active site.



**Fig. 3.36:** Degradation of myoglobin by the active site FPc285aM mutant. The reaction rate is much slower when compared to the wild-type FP-2. Accumulation of intermediate cleavages from 60 to 120 min is clearly evident.

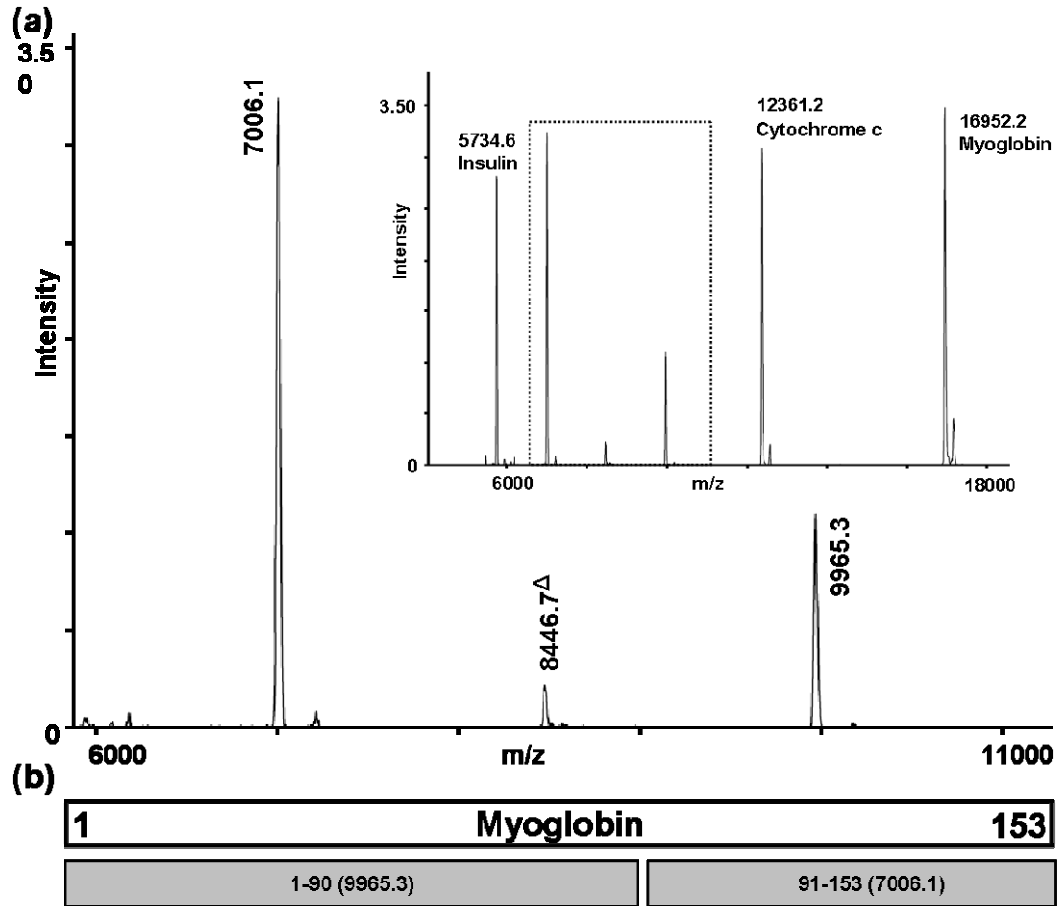
**3.11.2 FPc285aM mutant activity alongside wild-type FP-2** - To eliminate the doubt that traces of wild-type FP-2 are responsible for this activity, different concentrations of wild-type protease were used alongside the mutant. 4 µg/µl of the mutant and a series of diluted wild-type FP-2 (4 µg/µl, 0.4 µg/µl, 0.04 µg/µl) were included in the assay mixture containing 200 µg/µl of the substrate. The results analyzed by a Schagger gel showed that the wild-type FP-2 cleavage pattern and that of the mutant are very much comparable with slight differences (Fig. 3.37).



**Fig. 3.37:** Comparison of catalytic activity between wild-type FP-2 and the FP-c285aM mutant. The wild-type protein even at a very low concentration exhibited a different cleavage profile.

The first four lanes from the left (Fig. 3.37) are of <1 min incubated samples, lanes 5 - 8 are from samples incubated for 30 min and lanes 9 - 12 correspond to 90 min. The degradation profile for the wild-type protease was very much similar in all concentrations and even at the lowest concentration, three to four metastable intermediates were observed. However, in the case of FPc285aM mutant, only two metastable intermediates were produced even after 120 min of incubation. The step-wise accumulation of these intermediates was visible in the case of both the proteases. This experiment ruled out the possibility of a wild-type FP-2 impurity.

**3.11.3 MALDI-TOF analysis of cleavage products** - Similar to the experiments conducted with the wild-type protease, the myoglobin samples incubated for 120 min with the FPc285aM mutant were subjected to mass spectrometry (Fig. 3.38).

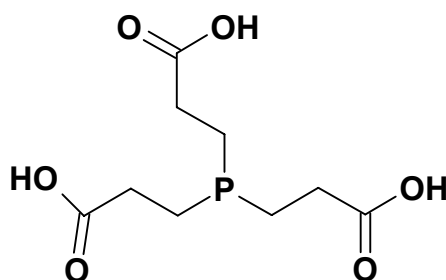


**Fig. 3.38:** MALDI-TOF spectrum of myoglobin degradation products obtained from 120 min incubated samples. **(A)** MALDI-TOF analysis was carried out in the molecular mass range 4000 to 18000 (inset). The boxed region of the inset highlights the area of the spectrum in detail. Double-ionized myoglobin peak is denoted (Δ). **(B)** Boxed representation of peptide fragments which are arranged in-line to cover the whole myoglobin sequence.

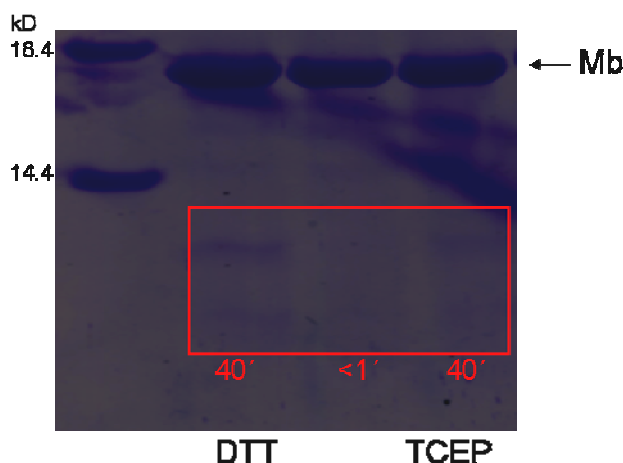
Two discrete cleavage products corresponding to 9965.3 and 7006.1 mass peaks were observed. Upon annotation, the 9965.3 peak matched the peptide fragment 1-90 and the 7006.1 peak matched the peptide fragment 91-153. These peptide fragments were produced due to the cleavage between residues Ala 90 and Gln 91 in helix F and can be placed in-line to cover the whole sequence of myoglobin. This site was one of the primary cleavage sites exploited by the wild-type protease. This

clearly indicates that the activity is not due to any traces of protease from *E.coli* but due the presence of a falcipain, in this case most probably the FPc285aM mutant.

**3.11.4 Degradation assay in the presence of reducing agent** - In all the degradation assays, 10 mM of DTT, a well-known cysteine protease activator, has been used. Previous studies have shown that DTT in the assay mixture might substitute, albeit inefficiently, for the missing thiol group from the free cysteine (Mueller et al., 2001). To rule out the possibility that DTT is acting as a substitute for the active site Cys42, TCEP (tris(2-carboxyethyl)phosphine), also a powerful reductant, was used in the assay mixture containing the FPc285aM mutant and myoglobin. The main advantage in using TCEP is that it cannot be involved in any nucleophilic substitution reactions due to the lack of thiol group (Fig. 3.39).



**Fig. 3.39:** Chemical structure of tris(2-carboxyethyl)phosphine (TCEP).



**Fig 3.40:** FPc285aM mutant exhibiting activity in the presence of DTT and TCEP. The sample incubation time are shown. The boxed region demonstrated the intermediate meta-stable bands.

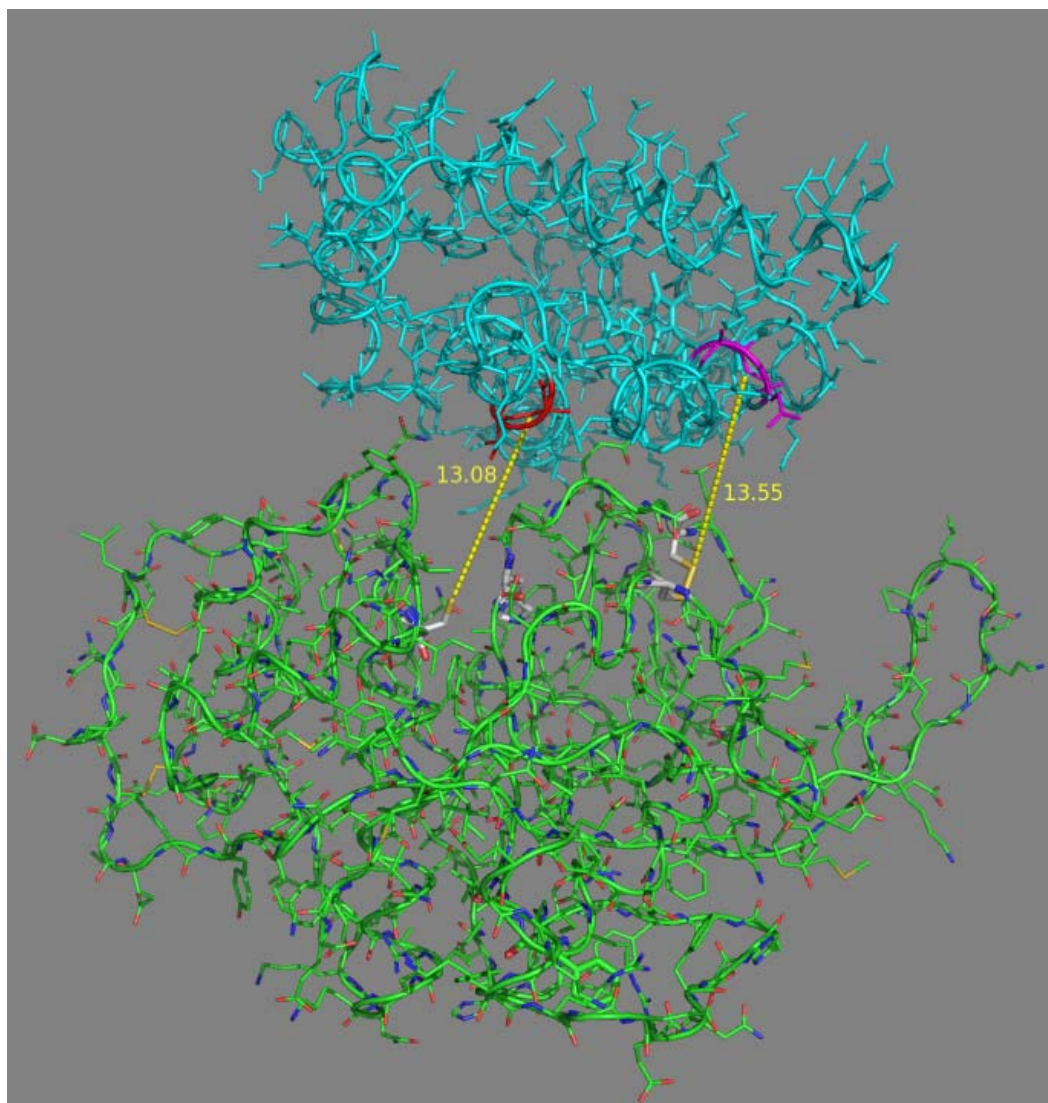


Sample mixture containing 10 mM DTT or 10 mM TCEP were incubated for various time intervals. Fig.3.40 showed no alteration in the cleavage pattern due to different reducing agents. The two metastable intermediate bands can be observed in both the cases (Fig. 3.40, 40 min). This result strongly suggests that the catalytic activity exhibited by the FPc285aM mutant is not due to thiol substitution from an external agent.

**3.11.5 Protein-protein docking studies** - From the results (Sections 3.11.2, 3.11.3, and 3.11.4), it can be concluded that the catalytic activity exhibited by the mutant is not due to any external protease activity or due to reducing agents. This leaves the option that the cysteine residues present in FP-2 could be responsible for this residual activity. Other than the active-site cysteine (Cys42), there are eight more cysteine residues present in FP-2 (Section 3.5.3). But all of them are involved in disulphide bridges (Cys39-Cys80, Cys73-Cys114, Cys99-Cys119 and Cys168-Cys229). It can be hypothesized that a free cysteine might be generated by reduction of any of these disulphide bonds.

Due to the unavailability of FP-2 : myoglobin or FP-2 : hemoglobin complex structural data, protein-protein docking and molecular dynamics simulation experiments were used to study the binding mode of FP-2 to myoglobin. Based on the cleavage information available from the earlier results (Section 3.8.8, Fig.3.33), two different constraint files were set between the enzyme and the substrate. The first constraint was between FP-2 active-site Cys42 and one of the myoglobin primary cleavage sites involving the residue Thr70. The second constraint was set between residues Ile85, Gln36 of FP-2 to Leu69, Leu72 of myoglobin, respectively. The docking studies were carried out at pH 5.0, where the protease activity of FP-2 is at its highest. The rigid-body docking results revealed multiple binding modes and

molecular dynamic simulations were obtained for those binding modes that best satisfied the constraints (Fig. 3.41).

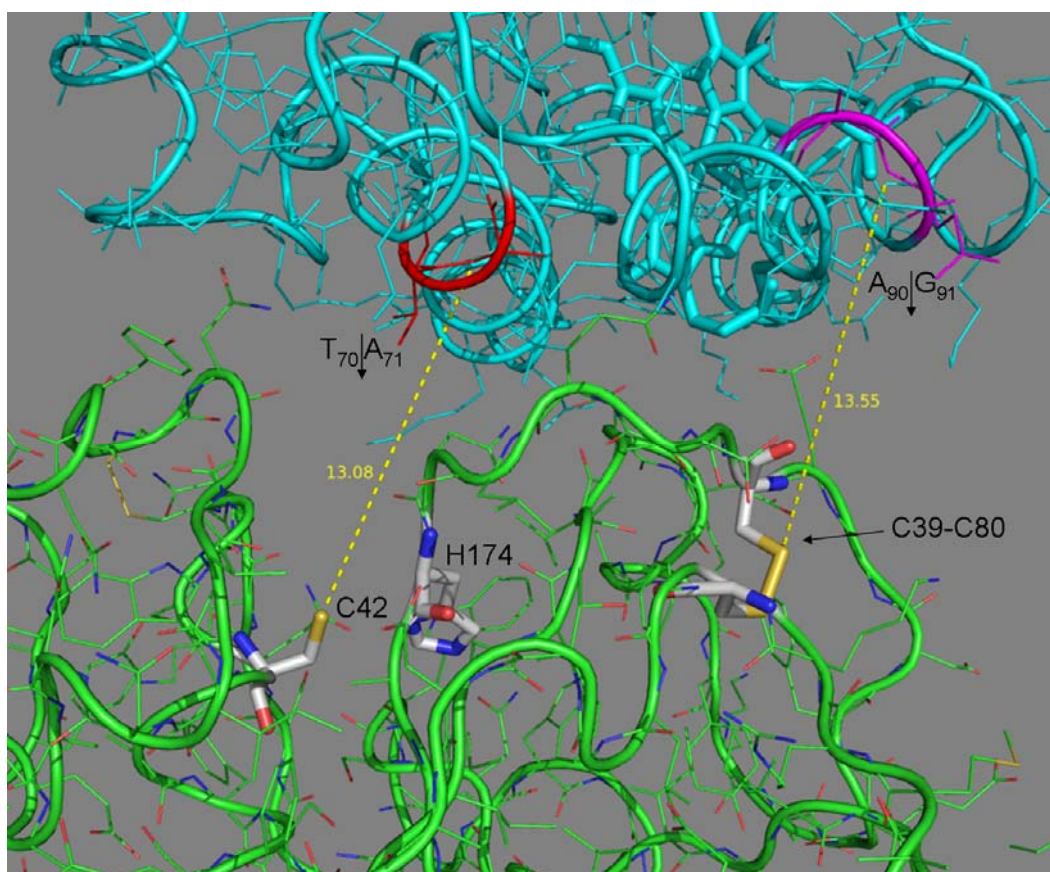


**Fig. 3.41:** Docking of myoglobin on FP-2 structure. The over all structure of myoglobin represented as ribbons (cyan) with side chain residues as thin lines. The cleavage sites between Thr70 and Ala71 (coloured red) and between residues Ala90 and Gln91 (coloured magenta) are also shown.

The docking models typically placed helix E nearer to the active site Cys42. This result complemented the results obtained from MALDI-TOF experiments, where it was shown that two of the major primary cleavages do occur in this helix (Fig 3.30). Though the unique motif or the FP-2<sub>arm</sub> exhibited many conformation states due to highly flexibility, none of the docking models placed the substrate molecule near to the unique motif.

## RESULT AND DISCUSSION

At a closer look at the docking model, it can be observed that a second helix (helix F) agreeably aligns on top of the Cys39-Cys80 disulphide bridge (Fig. 3.42). The MALDI-TOF experiments with the FPc285aM mutant (Fig. 3.37), already showed that the only cleavage site in myoglobin actually occurs between residues Ala90 and Gln91, which is situated in helix F (for helix information, see Fig. 3.35).



**Fig. 3.41:** Magnified view of Fig 3.40. The cleavage-site residues in helices E and F are denoted respectively. The active-site C42, H174 are shown as sticks. The disulphide bridge C39-C80 present in the vicinity of the helix F cleavage sites (Ala90 and Gln91) is shown as sticks. The distances between the respective cysteines and the cleavage sites are marked.

The docking models suggested that myoglobin in its native form cannot be brought closer than ~13 Å to the active site of the FP-2. This distance is too far for any protease activity if the myoglobin structure is intact. It could be hypothesized that the helices must unfold at least locally upon binding to FP-2 to initiate the proteolysis activity. This speculation is supported by previous studies on myoglobin (Hargrove et al., 1994). They reported that at room temperature (pH 7.0), myoglobin unfolds

## RESULT AND DISCUSSION

slowly. The process is initiated by unfolding of helices B, C, and E along with F to form a molten globular intermediate structure which is then completely unfolded. The authors further mentioned that the process is at least 50 times faster in the case of oxidized myoglobin. Considering the facts that the pH of the food vacuole as well as the experimental assays in this study is (pH 5.0 to 5.5), the myoglobin will be obviously denatured.

Due to the lack of complex crystal structure and supportive literature for a cysteine protease with a second active site, further insight is needed to establish the novelty in protease activity demonstrated by the FPc285aM mutant.

### 3.12 FP-2 Inhibitor studies -

**3.12.1 Falstatin, the natural inhibitor of FP-2** - Endogenous cysteine protease inhibitors have been described in a number of eukaryotic systems. In mammalian and plant cells, the cystatin superfamily of inhibitors and calpains regulate the functions of cysteine proteases (Dubin, 2005). Recently, a number of protease inhibitors have been reported for protozoan parasites such as *Trypanosoma cruzi* (Monteiro et al., 2001), *Leishmania mexicana* (Besteiro et al., 2004), and *Entamoeba histolytica* (Riekenberg et al., 2005). Similarly, falstatin a potent inhibitor of falcipains has been identified (Pandey et al., 2006). Pandey and colleagues reported that falstatin is present even during the early blood stages to control the activity of the protease and binds very tightly and totally inactivates the FP-2. It has been speculated that the inhibitor somehow facilitates the invasion of the parasite during blood stages. Similarly, a potential inhibitor of rodent malarial parasite *P. berghei* (PbICP- *Plasmodium berghei* inhibitor cysteine protease) was also identified by sequence search using chagasin (Inhibitor of *T. cruzi*) as a template (Pandey et al., 2006).

Homology modeling of PbICP has been performed for the C-terminal part of the protein and inhibitory studies with FP-2 were carried out with this PbICP C construct.

**3.12.2 Sequence alignment and homology modeling of PbICP** - Sequence alignment was performed for PbICP with chagasin using ClustalX. From the alignment it was clear that the chagasin-like domain is present in the C-term part (AA 203 to AA 354) of PbICP (full-length protein- 354 AA). Fig. 3.43 demonstrates the alignment between chagasin (Salmon et al., 2006) and PbICP C.

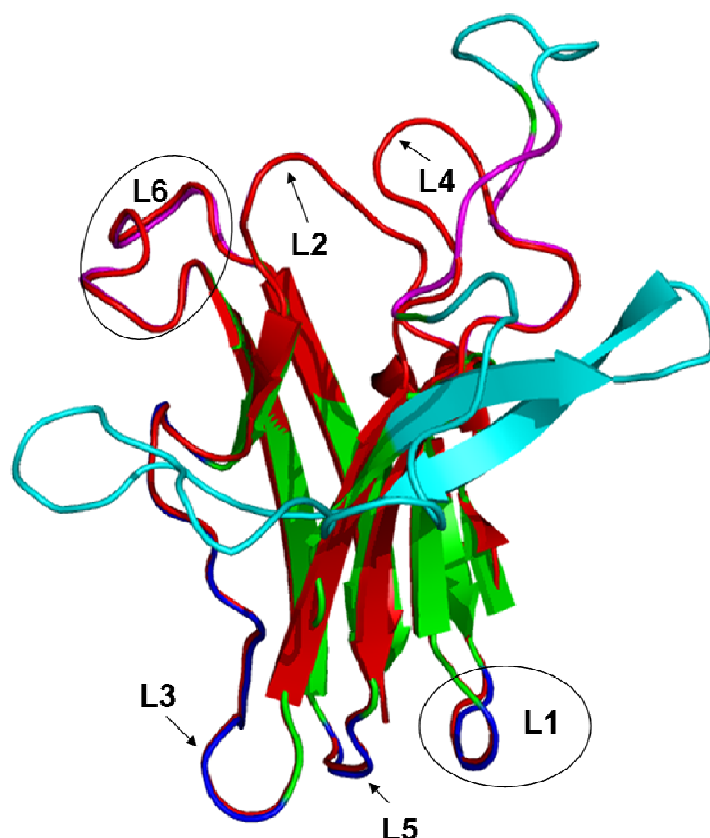
## RESULT AND DISCUSSION

				L1		L2
PbICP	1		ISNQT	NQETITQSLS	VGEILCIDLE	GNAGTGYLWV
Chagasin	3		HKVTK	AHNGATLTVA	VGELVEIQLP	SNPTTGFAWY
				. . * . .	*** . . *. *	*. . **.
		L3				
PbICP	36	LLGIHKDEPI	INPENFPTKL	TKKSFFSEEI	SVTQPKKYKI	DEHDSSKNVN
Chagasin	38	FEGGTKESPN	ESMFTVENKY	-----	-----	-----
		. * *. *	.	*		
			L4		L5	
PbICP	86	REIESPEQKE	SDSKPKKPQM	QLLGPD RMR	SVIKGHKPGK	YYIVYSYYRP
Chagasin	58	-----FP	PDSK-----	-LLGAGGTEH	FHVT VKAAGT	HAVNLTYMRP
			. ***	***.	.	. *
		L6				
PbICP	136	FSPTSGANTK	IIYVTVQ			
Chagasin	93	WTGPSHDSER	FTVYLKAN			
		*				

**Fig. 3.43:** Sequence alignment between chagasin and PblCP C (C-terminal). Conserved loop regions are highlighted in yellow and identical residues in these sequences are coloured in red.

In previous reports on chagasin interaction studies with cruzain, it has been shown that the loops (highlighted in Fig. 3.43) are conserved and loops L2, L4 and L6 are involved in substrate recognition. This study was also supported by the chagasin crystal structure in complex with FP-2 (Wang et al., 2007). Based on these studies, the loop sequences of PbICP C were manually adjusted to align with chagasin loops. From the sequence alignment, it was clear that many residues in the loops were identical, suggesting evolutionary conservation among ICP's (dos Reis et al., 2008). Surprisingly, the presence of 40 AA between loops L3 and L4 was observed in PbICP in comparison to chagasin. Similarly, an extra 7 AA was observed within the loop L4 whereas the loop was continuous in case of chagasin. The reason for this extra region is not yet clear.

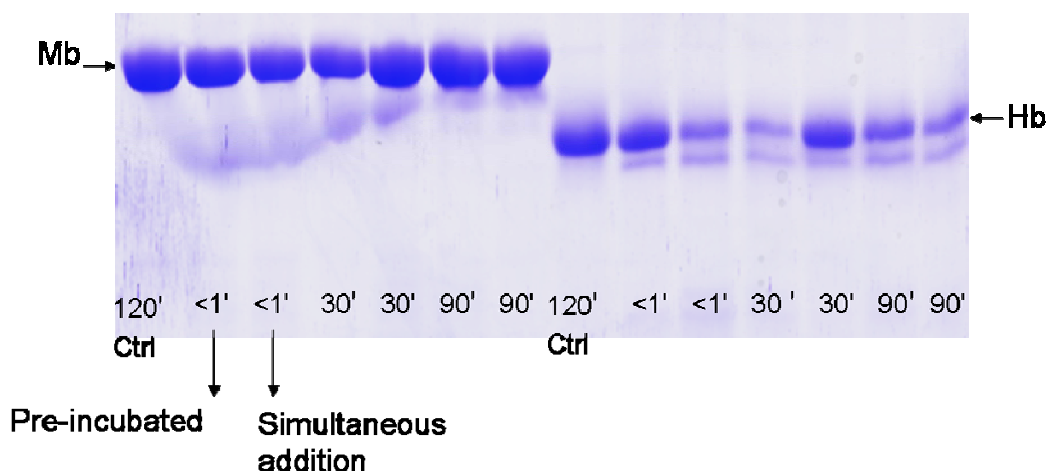
With the sequence alignment as the basis, homology modeling was performed using the chagasin structure (PDB-ID: 2FO8, Salmon et al., 2006). The loops L2 and L6 in the substrate recognition side and the loops L1, L3 and L5 are well defined. The L4 loop is slightly longer and the extra 40 AA forms a long antiparallel  $\beta$ -sheet and further protrudes out of the structure (Fig 3.44).



**Fig. 3.44:** Superposition of PBICP C model (green) on top of chagasin (red) shown as cartoon. The conserved loops are marked. Overall the model fits well except the extra 40 AA and 7 AA regions (coloured in cyan).

**3.12.3 Activity assay in the presence of *PbICP C*** - Affinity column purified *PbICP C* was kindly provided by our collaborators from Bernhard-Nocht-Institute for Tropical Medicine, Hamburg (Dr. Volker Heussler). The protein was stored in column elution buffer (20 mM  $\text{NaH}_2\text{PO}_4$ , 300 mM NaCl, and 500 mM imidazole at pH 8.0). Upon receiving, the storage buffer was exchanged to 20 mM Tris, 150 mM NaCl pH 7.5. Hemoglobin and myoglobin cleavage assays were carried out with FP-2 in the presence of *PbICP C*. 0.2 mg/ml of *PbICP C* was pre-incubated with 0.02 mg/ml of wild-type FP-2 at room temperature for 15 min before the addition of the substrates. Similarly, simultaneous addition of the inhibitor and substrate to the reaction mixture was also performed (Fig. 3.45).





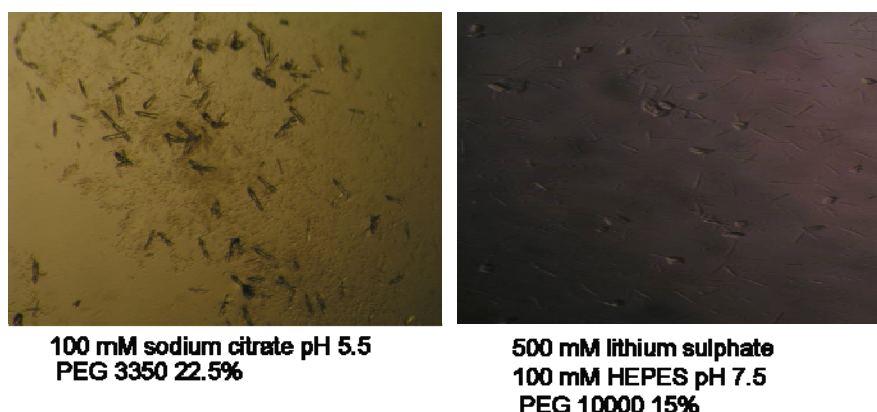
**Fig 3.45:** Activity assay of the wild-type FP-2 with myoglobin and hemoglobin in the presence of PbICP C. Cleavage is blocked almost totally. No metastable intermediate cleavage bands were observed. Reaction mixture with PbICP C without the enzyme was used as a control.

PbICP C efficiently inhibited the hydrolysis of FP-2. In case of myoglobin, no cleavage products were observed even after 90 min of incubation. The results were similar even when the substrate and the inhibitor were added simultaneously, indicating the high affinity of PbICP for FP-2. Strikingly, hemoglobin exhibited very minimal activity. Two distinct cleavage bands were visible even in <1 min samples, which might be due to a single cleavage in the chains (Fig. 3.45). The exact reason behind different inhibition efficiency of PbICP is not yet clear and further studies are underway.

**3.12.4 Crystallization of PbICP C** - Initial crystallization attempts have been performed with and without the protease. Crystallization using commercially available screens did not produce any PbICP C crystals. Due to high structural similarity, the crystallization conditions for chagasin (Figueiredo et al., 2007) were followed. By further optimization, tiny crystals of PbICP C were obtained in two different conditions (100 mM sodium citrate pH 5.5, 22.5% PEG 3350 and 500 mM lithium sulfate, 100 mM HEPES pH 7.5, 15% PEG 10000). Needle-like crystals started to appear within 2 to 3 days (Fig. 3.46). Crystal characterization procedures such as cracking and



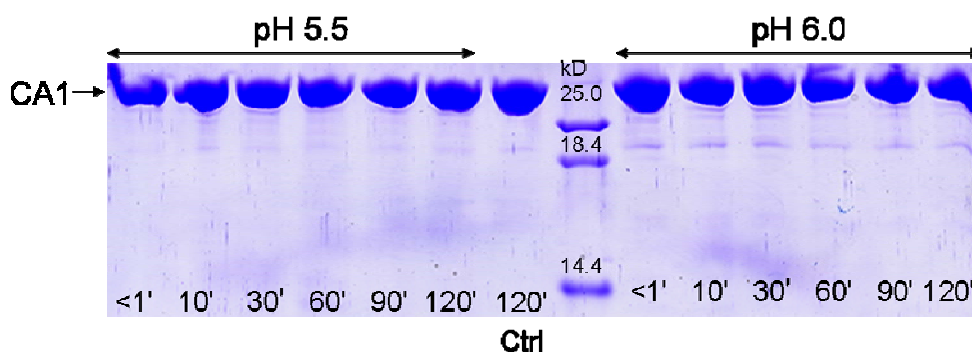
addition of IZIT dye (Hampton Research) confirmed that the crystals were of the protein. Further optimization of these crystals is required.



**Fig. 3.46:** Initial crystals of PbICP C in two different crystallization conditions as shown. Tiny needle-like crystals started to appear within 3 days.

**3.13 Search for other potential *in-vivo* substrates** - Carbonic anhydrase isoform 1 (CA1) is the second most abundant protein in the erythrocytes. The primary function of the enzyme in animals is to interconvert carbon dioxide and bicarbonate to maintain the acid-base balance in blood and other tissues and to help transport carbon dioxide out of tissues (Cramer et al., 1979). Experiments were conducted using CA1 as a substrate, to find whether FP-2, the multi function enzyme exhibit any cleavage activity.

**3.13.1 FP-2 degradation assays in presence of Carbonic anhydrase 1** - Similar to hemoglobin and myoglobin assays, 0.2 mg/ml CA1 was incubated with 0.02 mg/ml of active wild-type FP-2 for various time intervals at pH 5.5 and 6.0, where the activity of the protease is at its maximum (Fig. 3.47). The sample incubated at all time intervals (<1 to 120 min) were stable even after 120 min and showed no detectable degradation products. From this experiment, it can be concluded that CA1 is not an *in-vivo* substrate for FP-2.



**Fig 3.47:** FP-2 activity assay with CA1. No detectable cleavages could be observed at pH 5.0 and 6.0.

### 3.14 Search for novel cysteine protease Inhibitors -

**3.14.1 Virtual screen and binding studies of FP-2 inhibitors** – Virtual screening of potential compounds with inhibitory effect against active FP-2, synthesis of compounds and binding affinity studies using Surface Plasmon resonance (Biacore) were performed by our collaborators (Shanghai Institute of Materia Medica).

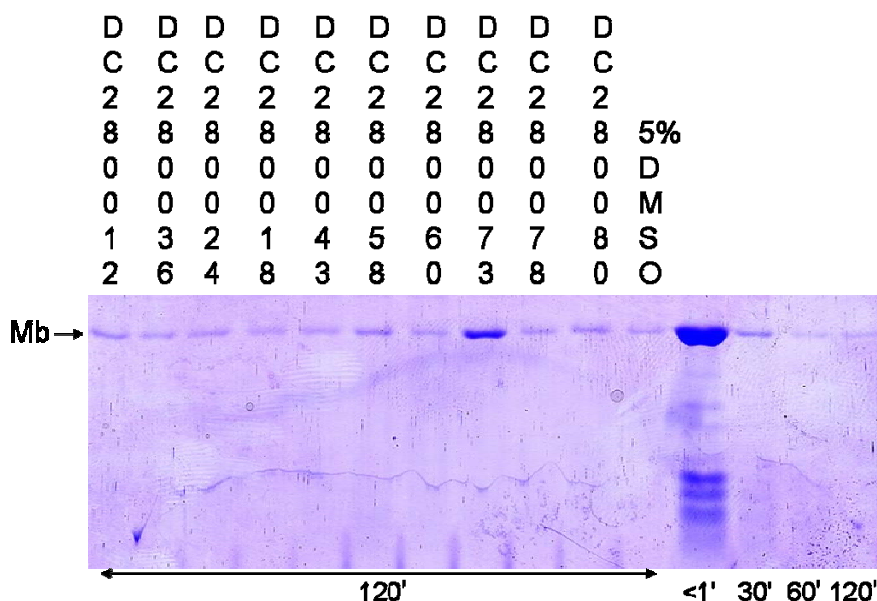
**Table 3.4:** The compounds and their  $K_d$  values obtained from Surface Plasmon resonance (SPR) studies are tabulated below.

No. of Compounds	Mol.wt (Da)	$K_d$ ( $\mu$ M)
Z-Phe-Arg-pNA.HCl (substrate)	612.09	~32.10
DC280012	413.48	5.88
DC280036	466.56	6.03
DC280024	489.39	2.79
DC280018	374.85	2.80
DC280043	528.98	49.0
DC280058	526.91	8.50
DC280060	413.48	31.90
DC280073	491.53	19.70
DC280078	500.65	12.90
DC280080	484.62	52.50

**3.14.2 Myoglobin degradation assay in the presence of inhibitors** - From the binding studies, it was clear that some of the compounds exhibited strong binding affinity towards FP-2. The myoglobin degradation assay was performed in the presence of these inhibitors. Five times  $K_d$  concentration (obtained from SPR binding studies) of each inhibitor was used in the assays. Inhibitor solutions were prepared

## RESULT AND DISCUSSION

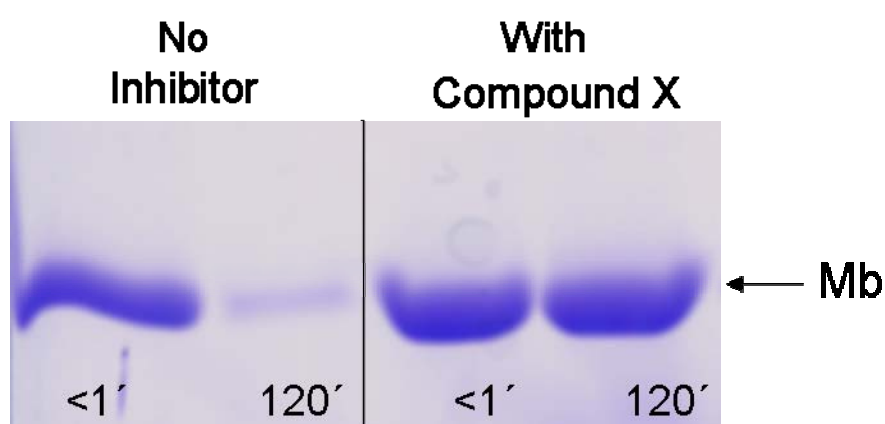
from stocks in DMSO (maximum concentration of DMSO in the assay was 1%). All the inhibitors were pre-incubated for 15 min with 0.02 mg/ml of active wild-type FP-2 at room temperature before addition of the substrate. Samples were collected after 120 min of incubation at 37°C. A control experiment without the inhibitors was done in parallel.



**Fig. 3.48:** Analysis of myoglobin degradation in the presence of inhibitors. FP-2 along with the inhibitors was pre-incubated for 15 min at room temperature. The samples without inhibitors (<1, 30, 60, and 120 min) were used in control experiments.

Myoglobin was completely degraded within 120 min even in the presence of inhibitors. The rate of the reaction was very much comparable to the control experiments which contained no inhibitors. The sample containing 5% DMSO showed no inhibition of FP-2 activity. The sample incubated with the inhibitor DSC280073 displayed slightly undegraded myoglobin band, suggesting some inhibitory effect (Fig. 3.48). This inhibitor can be used for further design of inhibitors; though this compound was not the one with the highest affinity from the binding assays (Table 3.4).

**3.14.3 Degradation assay in the presence of compound X** - The inhibitor compound X was kindly provided by Prof. Dr. François Diederich, ETH- Zurich. Upon receiving, the compound X (molecular weight: ~300 Da) was dissolved in 100% methanol. 1 mM of compound X was pre-incubated with 0.02 mg/ml of wild-type FP-2 in the assay mixture for 15 min at room temperature. 0.2 mg/ml of myoglobin was used as substrate. Samples were obtained after 120 min of incubation (Fig. 3.49). Samples without compound X was used in control experiments.



**Fig 3.49:** Degradation assay performed in the presence of compound X. The samples were incubated for 120 min. Samples without the inhibitor were used as control experiments.

The myoglobin cleavage activity by FP-2 was efficiently inhibited by compound X. The control experiments showed efficient proteolysis of myoglobin whereas no detectable degradation was visible even after 120 min of incubation in the presence of the inhibitor. From this result, it is clear that compound X exhibits tight binding and this compound can be exploited for further engineering of FP-2 specific inhibitors.

### 4. SUMMARY AND OUTLOOK

Malaria is one of the major infectious diseases known to man for several thousands of years. Each year, there are 300 - 500 million cases of malaria and is responsible for 1 - 3 million deaths each year. The disease is caused by protozoan parasites of the genus *Plasmodium*. The most serious form of the disease is caused by *Plasmodium falciparum*. There is an urgent call for new antimalarials targeting new sites. This present work is focused on the malarial cysteine protease FP-2. It is directed to answer several unknown questions such as structural aspects of FP-2, binding and cleavage specificity towards hemoglobin. Finally, this work presents studies for designing potential inhibitors against FP-2.

This dissertation provides a clear picture of hemoglobin binding and cleavage by FP-2. The results also indicated that FP-2 can autoproteolytically cleave its precursor *in trans*. The activity was at its maximum at neutral to weakly alkaline pH, whereas the binding of hemoglobin to FP-2 displayed a pH optimum in acidic range mimicking the food vacuole pH. The SPR binding studies showed that methemoglobin is the preferred substrate for FP-2 within the acidic food vacuole hinting at the survival strategy of the parasites.

The three-dimensional crystal structure of FP-2 provided further insight. The protein was crystallized in the presence of the inhibitor iodoacetamide at acidic pH. The structure was solved at 3.1 Å resolution in space group C222<sub>1</sub> using the Molecular Replacement method. Though the overall fold of the protein is similar to that of the papain family proteases, some structural aspects unique to malarial proteases were observed. The presence of a dynamic unique motif and a flexible amino-terminal helical extension were noticed. This N-terminal extension has been speculated to be involved in proper folding of the protein (Pandey et al., 2004) and

from the result from the mutant activity assays it may play a critical role in hemoglobin binding.

It has been proposed that FP-2 captures hemoglobin via this unique loop or FP-2<sub>arm</sub> hairpin prior to the subsequent cleavage of the substrate (Pandey et al., 2005). Two different binding models (4:1, 2:1), both signifying the importance of the loop were proposed (Pandey et al., 2005; Wang et al., 2006). The FP-2 mutants (FP-2 $\Delta$ 185-194, FP-2 $\Delta$ KK/G) efficiently degraded the hemoglobin leaving an uncertainty in the role played by this motif.

There have been many studies conducted for rational drug design for FP-2, using synthetic inhibitors as a model to study the substrate catalysis and specificity of FP-2 cleavage sites (Ramjee et al., 2006; Choe et al., 2006; Gelhaus et al., 2004; Rosenthal, 2004; Desai et al., 2006; Rosenthal, 2001). Ramjee et al., 2006; reported the cleavage specificity of FP-2 with various synthetic peptides as substrates. Peptides and irreversible peptidic inhibitors have always been used to study the substrate specificity of cysteine proteases. But for the first time, another globin protein was used as a potential substrate for the study the binding models and to invoke further insight in understanding of FP-2 substrate specificity.

Myoglobin, a monomeric protein structurally and functionally similar to hetero-tetrameric hemoglobin, was used as model substrate. Active mature FP-2 and also mutants readily bind and hydrolyze myoglobin. The pH-optimum is in the acidic range, as it has been reported for hemoglobin (Yayon et al., 1984; Krogstad et al., 1985). The cleavage reaction proceeds in a sequential manner and produces defined cleavage intermediates. The presence of only one unique protein chain in myoglobin compared to two different chains in the hemoglobin  $\alpha_2\beta_2$ -tetramer facilitated the analysis of the cleavage products by MALDI-TOF mass spectroscopy. Multiple cleavage sites were unequivocally identified. The sequential cleavage of peptide

bonds and the localization of these cleavage sites suggest that the specificity of FP-2 has been optimized to facilitate an early removal of the parasite-toxic heme group from the protein moiety. This result provided a mechanistic explanation for the previously observed inhibition of hemozoin formation by cysteine protease inhibitors (Gamboa de Domínguez et al., 1996). Furthermore, the primary and the secondary cleavage sites of myoglobin illustrated the substrate specificity of FP-2 in detail.

Degradation assays supported by MALDI-TOF mass spectrometry analysis of hemoglobin showed that the peptide fragments previously attributed only to aspartic proteases (plasmepsins) were also observed in the cleavages by FP-2. This result provides a strong support that FP-2 can degrade intact hemoglobin without the intervention of other proteases (Salas et al., 1995; Shenai et al., 2000).

FP-2 inhibition studies by PblCP results further enhanced the knowledge of natural inhibitors of cysteine proteases. The synthetic inhibitor (compound X) completely blocked the FP-2 activity, opening a new field for the design of potent inhibitors. A comparative structural analysis of the FP-2 active site with that of cruzain (cruzipain) from *Trypanosoma cruzi* suggests that existing vinyl sulfone inhibitors designed against cruzipain may also be useful starting compounds.

Recently, it has been proposed that berghepain-2, FP-2 equivalent in *P. berghei*, does not fulfil the roles of FP-2 (Chan et al., 2005) and the animal model studies (Fidock et al., 2004) may not represent the real scenario for the *in-vivo* analysis of *P. falciparum* inhibitors and for designing anti-malarial therapeutics. Berghepain-2 demonstrated optimal activity at narrower pH optima of 5.5 - 6.0 and exhibited a lack of preference for substrates with leucine at P2 (Chan et al., 2005). Taking this into account it can be concluded the results provided in this dissertation will be instrumental for the design of FP-2 specific inhibitors and will open a new era in rational design of novel antimalarials.

## 5. REFERENCES

- Akompong, T., Ghorri, N. & Haldar, K. (2000): *In-vitro* activity of riboflavin against the human malaria parasite *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* **44**, 88-96
- Alkadi, H.O. (2007): Antimalarial drug toxicity: a review. *Chemotherapy* **53**, 385-391
- Asawamahasakda, W., Ittarat, I., Chang, C.C., McElroy, P. & Meshnick, S.R. (1994): Effects of antimalarials and protease inhibitors on plasmodial hemozoin production. *Mol. Biochem. Parasitol.* **67**, 183-191
- Aslan, K., Lakowicz, J.R. & Geddes, C.D. (2005): Plasmon light scattering in biology and medicine: new sensing approaches, visions and perspectives. *Curr. Opin. Chem. Biol.* **9**, 538-544
- Baird, J.K., Wiady, I., Sutanihardja, A., Suradi, Purnomo, Basri, H., Sekartuti, Ayomi, E., Fryauff, D.J. & Hoffman, S.L. (2002): Short report: therapeutic efficacy of chloroquine combined with primaquine against *Plasmodium falciparum* in northeastern Papua, Indonesia. *Am. J. Trop. Med. Hyg.* **66**, 659-660
- Baird, J.K. (2005): Effectiveness of antimalarial drugs. *N. Engl. J. Med.* **352**, 1565-1577
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J. & McCammon, J.A. (2001): Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037-10041
- Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klembe, M. & Goldberg, D.E. (2002): Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proc. Natl. Acad. Sci. USA* **99**, 990-995
- Banerjee, R., Francis, S.E. & Goldberg, D.E. (2003): Food vacuole plasmepsins are processed at a conserved site by an acidic convertase activity in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **129**, 157-165
- Barrett, A.J. & Rawlings, N.D. (2001): Evolutionary lines of cysteine peptidases. *Biol. Chem.* **382**, 727-733
- Bécuwe, P., Gratepanche, S., Fourmaux, M.N., Van Beeumen, J., Samyn, B., Mercereau-Puijalon, O., Touzel, J.P., Slomianny, C., Camus, D. & Dive, D. (1996): Characterization of iron-dependent endogenous superoxide dismutase of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **76**, 125-134
- Bennett-Lovsey, R.M., Herbert, A.D., Sternberg, M.J. & Kelley, L.A. (2008): Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **70**, 611-625
- Bergfors T: Protein Crystallization Techniques, Strategies, and Tips - A Laboratory Manual. International University Line, La Jolla, USA, 1999.



## REFERENCES

- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000): The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-242
- Besteiro, S., Coombs, G.H. & Mottram, J.C. (2004): A potential role for ICP, a Leishmanial inhibitor of cysteine peptidases, in the interaction between host and parasite. *Mol. Microbiol.* **54**, 1224-1236
- Bode, W. & Huber, R. (2000): Structural basis of the endoproteinase-protein inhibitor interaction. *Biochim. Biophys. Acta* **1477**, 241-252
- Bonilla, J.A., Bonilla, T.D., Yowell, C.A., Fujioka, H. & Dame, J.B. (2007): Critical roles for the digestive vacuole plasmepsins of *Plasmodium falciparum* in vacuolar function. *Mol. Microbiol.* **65**, 64-75
- Bozdech, Z., Llinas, M., Pulliam, B.L., Wong, E.D., Zhu, J. & DeRisi, J.L. (2003): The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* **1**, E5
- Bozdech, Z., Zhu, J., Joachimiak, M.P., Cohen, F.E., Pulliam, B. & DeRisi, J.L. (2003): Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol.* **4**, R9
- Brinen, L.S., Hansell, E., Cheng, J., Roush, W.R., McKerrow, J.H. & Fletterick, R.J. (2000): A target within the target: probing cruzain's P1' site to define structural determinants for the Chagas' disease protease. *Structure* **8**, 831-840
- Bromme, D., Bonneau, P.R., Lachance, P. & Storer, A.C. (1994): Engineering the S2 subsite specificity of human cathepsin S to a cathepsin L- and cathepsin B-like specificity. *J. Biol. Chem.* **269**, 30238-30242
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. & Warren, G.L. (1998): Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst.* **D54**, 905-921
- Burmester, T., Weich, B., Reinhardt, S. & Hankeln, T. (2000): A vertebrate globin expressed in the brain. *Nature* **407**, 520-523
- Chan, C., Goh, L.L. & Sim, T.S. (2005): Differences in biochemical properties of the Plasmodial falcipain-2 and berghepain-2 orthologues: implications for *in-vivo* screens of inhibitors. *FEMS Microbiol. Lett.* **249**, 315-321
- Choe, Y., Brinen, L.S., Price, M.S., Engel, J.C., Lange, M., Grisostomi, C., Weston, S.G., Pallai, P.V., Cheng, H., Hardy, L.W., Hartsough, D.S., McMakin, M., Tilton, R.F., Baldino, C.M. & Craik, C.S. (2005): Development of alpha-keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease. *Bioorg. Med. Chem.* **13**, 2141-2156
- Choe, Y., Leonetti, F., Greenbaum, D.C., Lecaille, F., Bogyo, M., Bromme, D., Ellman, J.A. & Craik, C.S. (2006): Substrate profiling of cysteine proteases using a

## REFERENCES

- combinatorial peptide library identifies functionally unique specificities. *J. Biol. Chem.* **281**, 12824-12832
- CCP4 (1994): The CCP4 suite: programs for protein crystallography. *Acta Cryst.* **D50**, 760-763
- Corpet, F. (1988): Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881-10890
- Cowman, A.F. & Crabb, B.S. (2006): Invasion of red blood cells by malaria parasites. *Cell* **124**, 755-766
- Cramer, F., Gould, H., Barlow, S. & Carter, N. (1979): Synthesis of carbonic anhydrase in rabbit and chicken reticulocyte lysates. *Eur. J. Biochem.* **95**, 99-105
- Curley, G.P., O'Donovan, S.M., McNally, J., Mullally, M., O'Hara, H., Troy, A., O'Callaghan, S.A. & Dalton, J.P. (1994): Aminopeptidases from *Plasmodium falciparum*, *Plasmodium chabaudi* and *Plasmodium berghei*. *J. Eukaryot. Microbiol.* **41**, 119-123
- Dahl, E.L. & Rosenthal, P.J. (2005): Biosynthesis, localization, and processing of falcipain cysteine proteases of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **139**, 205-212
- DeLano WL. The PyMOL Molecular Graphics System, DeLano Scientific. 2002; San Carlos, CA, USA.
- Desai, P.V., Patny, A., Gut, J., Rosenthal, P.J., Tekwani, B., Srivastava, A. & Avery, M. (2006): Identification of novel parasitic cysteine protease inhibitors by use of virtual screening. 2. The available chemical directory. *J. Med. Chem.* **49**, 1576-1584
- Dhawan, S., Dua, M., Chishti, A.H. & Hanspal, M. (2003): Ankyrin peptide blocks falcipain-2-mediated malaria parasite release from red blood cells. *J. Biol. Chem.* **278**, 30180-30186
- Dolinsky, T.J., Nielsen, J.E., McCammon, J.A. & Baker, N.A. (2004): PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665-667
- dos Reis, F.C., Smith, B.O., Santos, C.C., Costa, T.F., Scharfstein, J., Coombs, G.H., Mottram, J.C. & Lima, A.P. (2008): The role of conserved residues of chagasin in the inhibition of cysteine peptidases. *FEBS Lett.* **582**, 485-490
- Drew, M.E., Banerjee, R., Uffman, E.W., Gilbertson, S., Rosenthal, P.J. & Goldberg, D.E. (2008): *Plasmodium* food vacuole plasmepsins are activated by falcipains. *J. Biol. Chem.* **283**, 12870-12876
- Dua, M., Raphael, P., Sijwali, P.S., Rosenthal, P.J. & Hanspal, M. (2001): Recombinant falcipain-2 cleaves erythrocyte membrane ankyrin and protein 4.1. *Mol. Biochem. Parasitol.* **116**, 95-99

## REFERENCES

- Dubin, G. (2005): Proteinaceous cysteine protease inhibitors. *Cell. Mol. Life Sci.* **62**, 653-669
- Eggleson, K.K., Duffin, K.L. & Goldberg, D.E. (1999): Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* **274**, 32411-32417
- Eksi, S., Czesny, B., Greenbaum, D.C., Bogyo, M. & Williamson, K.C. (2004): Targeted disruption of *Plasmodium falciparum* cysteine protease, falcipain 1, reduces oocyst production, not erythrocytic stage growth. *Mol. Microbiol.* **53**, 243-250
- Emsley, P. & Cowtan, K. (2004): Coot: model-building tools for molecular graphics. *Acta Cryst.* **D60**, 2126-2132
- Evans, S.V. & Brayer, G.D. (1988): Horse heart metmyoglobin. A 2.8 Å resolution three-dimensional structure determination. *J. Biol. Chem.* **263**, 4263-4268
- Evans, S.V. & Brayer, G.D. (1990): High-resolution study of the three-dimensional structure of horse heart metmyoglobin. *J. Mol. Biol.* **213**, 885-897
- Fidock, D.A., Rosenthal, P.J., Croft, S.L., Brun, R. & Nwaka, S. (2004): Antimalarial drug discovery: efficacy models for compound screening. *Nature Rev. Drug Discov.* **3**, 509-520
- Figueiredo da Silva, A.A., de Carvalho Vieira, L., Krieger, M.A., Goldenberg, S., Zanchin, N.I. & Guimaraes, B.G. (2007): Crystal structure of chagasin, the endogenous cysteine-protease inhibitor from *Trypanosoma cruzi*. *J. Struct. Biol.* **157**, 416-423
- Fitch, C.D., Chevli, R., Banyal, H.S., Phillips, G., Pfaller, M.A. & Krogstad, D.J. (1982): Lysis of *Plasmodium falciparum* by ferriprotoporphyrin IX and a chloroquine-ferriprotoporphyrin IX complex. *Antimicrob. Agents Chemother.* **21**, 819-822
- Florent, I., Derhy, Z., Allary, M., Monsigny, M., Mayer, R. & Schrevel, J. (1998): A *Plasmodium falciparum* aminopeptidase gene belonging to the M1 family of zinc-metallopeptidases is expressed in erythrocytic stages. *Mol. Biochem. Parasitol.* **97**, 149-160
- Francis, S.E., Gluzman, I.Y., Oksman, A., Knickerbocker, A., Mueller, R., Bryant, M.L., Sherman, D.R., Russell, D.G. & Goldberg, D.E. (1994): Molecular characterization and inhibition of a *Plasmodium falciparum* aspartic hemoglobinase. *EMBO J.* **13**, 306-317
- Francis, S.E., Banerjee, R. & Goldberg, D.E. (1997): Biosynthesis and maturation of the malaria aspartic hemoglobinases plasmepsins I and II. *J. Biol. Chem.* **272**, 14961-14968
- Fry, M. & Beesley, J.E. (1991): Mitochondria of mammalian *Plasmodium* spp. *Parasitology* **102**, 17-26

## REFERENCES

- Gamboa de Dominguez, N.D. & Rosenthal, P.J. (1996): Cysteine proteinase inhibitors block early steps in hemoglobin degradation by cultured malaria parasites. *Blood* **87**, 4448-4454
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. & Bairoch, A. (2003): ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784-3788
- Gattiker, A., Bienvenut, W.V., Bairoch, A. & Gasteiger, E. (2002): FindPept, a tool to identify unmatched masses in peptide mass fingerprinting protein identification. *Proteomics* **2**, 1435-1444
- Gelhaus, C., Vicik, R., Hilgenfeld, R., Schmidt, C.L., Leippe, M. & Schirmeister, T. (2004): Synthesis and antiparasmodial activity of a cysteine protease-inhibiting biotinylated aziridine-2,3-dicarboxylate. *Biol. Chem.* **385**, 435-438
- Gibrat, J.F., Madej, T. & Bryant, S.H. (1996): Surprising similarities in structure comparison. *Curr. Opin. Struct. Biol.* **6**, 377-385
- Gillmor, S.A., Craik, C.S. & Fletterick, R.J. (1997): Structural determinants of specificity in the cysteine protease cruzain. *Protein Sci.* **6**, 1603-1611
- Gluzman, I.Y., Francis, S.E., Oksman, A., Smith, C.E., Duffin, K.L. & Goldberg, D.E. (1994): Order and specificity of the *Plasmodium falciparum* hemoglobin degradation pathway. *J. Clin. Invest.* **93**, 1602-1608
- Goh, L.L. & Sim, T.S. (2004): Homology modeling and mutagenesis analyses of *Plasmodium falciparum* falcipain 2A: implications for rational drug design. *Biochem. Biophys. Res. Commun.* **323**, 565-572
- Goh, L.L. & Sim, T.S. (2005): Characterization of amino acid variation at strategic positions in parasite and human proteases for selective inhibition of falcipains in *Plasmodium falciparum*. *Biochem. Biophys. Res. Commun.* **335**, 762-770
- Goldberg, D.E., Slater, A.F., Beavis, R., Chait, B., Cerami, A. & Henderson, G.B. (1991): Hemoglobin degradation in the human malaria pathogen *Plasmodium falciparum*: a catabolic pathway initiated by a specific aspartic protease. *J. Exp. Med.* **173**, 961-969
- Goldberg, D.E. (1993): Hemoglobin degradation in *Plasmodium*-infected red blood cells. *Semin. Cell. Biol.* **4**, 355-361
- Gordeuk, V.R., Thuma, P.E. & Brittenham, G.M. (1994): Iron chelation therapy for malaria. *Adv. Exp. Med. Biol.* **356**, 371-383
- Greenbaum, D.C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradsky, K.F., Engel, J., DeRisi, J., Holder, A.A. & Bogoy, M. (2002): A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* **298**, 2002-2006

## REFERENCES

- Greenwood, B.M., Bojang, K., Whitty, C.J. & Targett, G.A. (2005): Malaria. *Lancet* **365**, 1487-1498
- Guex, N. & Peitsch, M.C. (1997): SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723
- Guncar, G., Klemencic, I., Turk, B., Turk, V., Karaoglanovic-Carmona, A., Juliano, L. & Turk, D. (2000): Crystal structure of cathepsin X: a flip-flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease. *Structure* **8**, 305-313
- Hanspal, M., Dua, M., Takakuwa, Y., Chishti, A.H. & Mizuno, A. (2002): *Plasmodium falciparum* cysteine protease falcipain-2 cleaves erythrocyte membrane skeletal proteins at late stages of parasite development. *Blood* **100**, 1048-1054
- Hargrove, M.S., Krzywda, S., Wilkinson, A.J., Dou, Y., Ikeda-Saito, M. & Olson, J.S. (1994): Stability of myoglobin: a model for the folding of heme proteins. *Biochemistry* **33**, 11767-11775
- Hogg, T., Nagarajan, K., Herzberg, S., Chen, L., Shen, X., Jiang, H., Wecke, M., Blohmke, C., Hilgenfeld, R. & Schmidt, C.L. (2006): Structural and functional characterization of Falcipain-2, a hemoglobinase from the malarial parasite *Plasmodium falciparum*. *J. Biol. Chem.* **281**, 25425-25437
- Huang, L., Brinen, L.S. & Ellman, J.A. (2003): Crystal structures of reversible ketone-Based inhibitors of the cysteine protease cruzain. *Bioorg. Med. Chem.* **11**, 21-29
- Jambou, R., Legrand, E., Niang, M., Khim, N., Lim, P., Volney, B., Ekala, M.T., Bouchier, C., Esterre, P., Fandeur, T. & Mercereau-Puijalon, O. (2005): Resistance of *Plasmodium falciparum* field isolates to in-vitro artemether and point mutations of the SERCA-type PfATPase6. *Lancet* **366**, 1960-1963
- Jones, T.R. & Hoffman, S.L. (1994): Malaria vaccine development. *Clin. Microbiol. Rev.* **7**, 303-310
- Kachalova, G.S., Popov, A.N. & Bartunik, H.D. (1999): A steric mechanism for inhibition of CO binding to heme proteins. *Science* **284**, 473-476
- Kamphuis, I.G., Kalk, K.H., Swarte, M.B. & Drenth, J. (1984): Structure of papain refined at 1.65 Å resolution. *J. Mol. Biol.* **179**, 233-256
- Kavanaugh, J.S., Moo-Penn, W.F. & Arnone, A. (1993): Accommodation of insertions in helices: the mutation in hemoglobin Catonsville (Pro 37  $\alpha$ -Glu-Thr 38  $\alpha$ ) generates a 3<sub>10</sub>→ $\alpha$  bulge. *Biochemistry* **32**, 2509-2513
- Kolakovich, K.A., Gluzman, I.Y., Duffin, K.L. & Goldberg, D.E. (1997): Generation of hemoglobin peptides in the acidic digestive vacuole of *Plasmodium falciparum* implicates peptide transport in amino acid production. *Mol. Biochem. Parasitol.* **87**, 123-135

## REFERENCES

- Krogstad, D.J., Schlesinger, P.H. & Gluzman, I.Y. (1985): Antimalarials increase vesicle pH in *Plasmodium falciparum*. *J. Cell Biol.* **101**, 2302-2309
- Lämmli, U.K. (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
- LaLonde, J.M., Zhao, B., Janson, C.A., D'Alessio, K.J., McQueney, M.S., Orsini, M.J., Debouck, C.M. & Smith, W.W. (1999): The crystal structure of human procathepsin K. *Biochemistry* **38**, 862-869
- Laufer, M.K., Thesing, P.C., Eddington, N.D., Masonga, R., Dzinjalamala, F.K., Takala, S.L., Taylor, T.E. & Plowe, C.V. (2006): Return of chloroquine antimalarial efficacy in Malawi. *N. Engl. J. Med.* **355**, 1959-1966
- Lecaille, F., Kaleta, J. & Bromme, D. (2002): Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. *Chem. Rev.* **102**, 4459-4488
- Lee, K., Bae, D. & Lim, D. (2002): Evaluation of parameters in peptide mass fingerprinting for protein identification by MALDI-TOF mass spectrometry. *Mol. Cells* **13**, 175-184
- Leslie AGW. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography*, 1992; No. 26.
- Leung-Toung, R., Zhao, Y., Li, W., Tam, T.F., Karimian, K. & Spino, M. (2006): Thiol proteases: inhibitors and potential therapeutic targets. *Curr. Med. Chem.* **13**, 547-581
- Lew, V.L., Tiffert, T. & Ginsburg, H. (2003): Excess hemoglobin digestion and the osmotic stability of *Plasmodium falciparum*-infected red blood cells. *Blood* **101**, 4189-4194
- Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, Ilin A, Antosiewicz J, Gilson MK, Bagheri B, Scott LR and McCammon JA. (1995): Electrostatics and Diffusion of Molecules in Solution: Simulations with the University of Houston Brownian Dynamics Program. *Comp. Phys. Commun.* **91**, 57-95.
- Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F. & Sali, A. (2000): Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291-325
- Matthews, B.W. (1968): Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497
- McGrath, M.E., Eakin, A.E., Engel, J.C., McKerrow, J.H., Craik, C.S. & Fletterick, R.J. (1995): The crystal structure of cruzain: a therapeutic target for Chagas' disease. *J. Mol. Biol.* **247**, 251-259
- McPherson, A. (2004): Protein crystallization in the structural genomics era. *J. Struct. Funct. Genomics* **5**, 3-12

## REFERENCES

- Mehlin, C. (2005): Structure-based drug discovery for *Plasmodium falciparum*. *Comb. Chem. High Throughput Screen* **8**, 5-14
- Monteiro, A.C., Abrahamson, M., Lima, A.P., Vannier-Santos, M.A. & Scharfstein, J. (2001): Identification, characterization and localization of chagasin, a tight-binding cysteine protease inhibitor in *Trypanosoma cruzi*. *J. Cell Sci.* **114**, 3933-3942
- Morris, A.L., MacArthur, M.W., Hutchinson, E.G. & Thornton, J.M. (1992): Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364
- Motiejunas, D., Gabdoulline, R., Wang, T., Feldman-Salit, A., Johann, T., Winn, P.J. & Wade, R.C. (2008): Protein-protein docking by simulating the process of association subject to biochemical constraints. *Proteins* **71**, 1955-1969
- Mueller, E.G., Palenchar, P.M. & Buck, C.J. (2001): The role of the cysteine residues of ThiI in the generation of 4-thiouridine in tRNA. *J. Biol. Chem.* **276**, 33588-33595
- Müller, S. (2004): Redox and antioxidant systems of the malaria parasite *Plasmodium falciparum*. *Mol. Microbiol.* **53**, 1291-1305
- Murata, C.E. & Goldberg, D.E. (2003a): *Plasmodium falciparum* falcilysin: a metalloprotease with dual specificity. *J. Biol. Chem.* **278**, 38022-38028
- Murata, C.E. & Goldberg, D.E. (2003b): *Plasmodium falciparum* falcilysin: an unprocessed food vacuole enzyme. *Mol. Biochem. Parasitol.* **129**, 123-126
- Murphy, G.S., Basri, H., Purnomo, Andersen, E.M., Bangs, M.J., Mount, D.L., Gorden, J., Lal, A.A., Purwokusumo, A.R., Harjosuwarno, S. & et al. (1993): Vivax malaria resistant to treatment and prophylaxis with chloroquine. *Lancet* **341**, 96-100
- Murshudov GN, Vagin AA and Dodson EJ. (1997): Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst.* **D53**, 240-255.
- Musil, D., Zucic, D., Turk, D., Engh, R.A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N. & et al. (1991): The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J.* **10**, 2321-2330
- O'Donnell, A., Weatherall, D.J., Taylor, A.M., Reeder, J.C. & Allen, S.J. (2006): Muscle cell injury, haemolysis and dark urine in children with *falciparum* malaria in Papua New Guinea. *Trans. R Soc. Trop. Med. Hyg.* **100**, 817-825
- O'Hara, B.P., Hemmings, A.M., Buttle, D.J. & Pearl, L.H. (1995): Crystal structure of glycyl endopeptidase from *Carica papaya*: a cysteine endopeptidase of unusual substrate specificity. *Biochemistry* **34**, 13190-13195
- Olliaro, P., Nevill, C., LeBras, J., Ringwald, P., Mussano, P., Garner, P. & Brasseur, P. (1996): Systematic review of amodiaquine treatment in uncomplicated malaria. *Lancet* **348**, 1196-1201

## REFERENCES

- Orjih, A.U., Banyal, H.S., Chevli, R. & Fitch, C.D. (1981): Hemin lyses malaria parasites. *Science* **214**, 667-669
- Pandey, K.C., Sijwali, P.S., Singh, A., Na, B.K. & Rosenthal, P.J. (2004): Independent intramolecular mediators of folding, activity, and inhibition for the *Plasmodium falciparum* cysteine protease falcipain-2. *J. Biol. Chem.* **279**, 3484-3491
- Pandey, K.C., Wang, S.X., Sijwali, P.S., Lau, A.L., McKerrow, J.H. & Rosenthal, P.J. (2005): The *Plasmodium falciparum* cysteine protease falcipain-2 captures its substrate, hemoglobin, via a unique motif. *Proc. Natl. Acad. Sci. USA* **102**, 9138-9143
- Pandey, K.C., Singh, N., Arastu-Kapur, S., Bogyo, M. & Rosenthal, P.J. (2006): Falstatin, a cysteine protease inhibitor of *Plasmodium falciparum*, facilitates erythrocyte invasion. *PLoS Pathog.* **2**, e117
- Panisko, D.M. & Keystone, J.S. (1990): Treatment of malaria-1990. *Drugs* **39**, 160-189
- Pesce, A., Dewilde, S., Nardini, M., Moens, L., Ascenzi, P., Hankeln, T., Burmester, T. & Bolognesi, M. (2003): Human brain neuroglobin structure reveals a distinct mode of controlling oxygen affinity. *Structure* **11**, 1087-1095
- Pickersgill RW, Harris GW and Garman E. (1992): Structure of monoclinic papain at 1.60 Å resolution. *Acta Cryst.* **B48**, 59-67
- Ponder, J.W. & Case, D.A. (2003): Force fields for protein simulations. *Adv. Protein. Chem.* **66**, 27-85
- Ramjee, M.K., Flinn, N.S., Pemberton, T.P., Quibell, M., Wang, Y. & Watts, J.P. (2006): Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepain-2: implications for peptidase anti-malarial drug discovery. *Biochem. J.* **399**, 47-57
- Read RJ. (1986): Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst.* **A42**, 140-149.
- Riekenberg, S., Witjes, B., Saric, M., Bruchhaus, I. & Scholze, H. (2005): Identification of EhICP1, a chagasin-like cysteine protease inhibitor of *Entamoeba histolytica*. *FEBS Lett.* **579**, 1573-1578
- Rosenthal, P.J., McKerrow, J.H., Aikawa, M., Nagasawa, H. & Leech, J.H. (1988): A malarial cysteine proteinase is necessary for hemoglobin degradation by *Plasmodium falciparum*. *J. Clin. Invest.* **82**, 1560-1566
- Rosenthal, P.J. & Nelson, R.G. (1992): Isolation and characterization of a cysteine proteinase gene of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **51**, 143-152
- Rosenthal, P.J. & Meshnick, S.R. (1996): Hemoglobin catabolism and iron utilization by malaria parasites. *Mol. Biochem. Parasitol.* **83**, 131-139



## REFERENCES

- Rosenthal, P.J. (1998): Proteases of malaria parasites: new targets for chemotherapy. *Emerg. Infect. Dis.* **4**, 49-57
- Rosenthal, P.J. (2002): Hydrolysis of erythrocyte proteins by proteases of malaria parasites. *Curr. Opin. Hematol.* **9**, 140-145
- Rosenthal, P.J., Sijwali, P.S., Singh, A. & Shenai, B.R. (2002): Cysteine proteases of malaria parasites: targets for chemotherapy. *Curr. Pharm. Des.* **8**, 1659-1672
- Rosenthal, P.J. (2004): Cysteine proteases of malaria parasites. *Int. J. Parasitol.* **34**, 1489-1499
- Rubin, H., Salem, J.S., Li, L.S., Yang, F.D., Mama, S., Wang, Z.M., Fisher, A., Hamann, C.S. & Cooperman, B.S. (1993): Cloning, sequence determination, and regulation of the ribonucleotide reductase subunits from *Plasmodium falciparum*: a target for antimalarial therapy. *Proc. Natl. Acad. Sci. USA* **90**, 9280-9284
- Sabnis, Y., Rosenthal, P.J., Desai, P. & Avery, M.A. (2002): Homology modeling of falcipain-2: validation, de novo ligand design and synthesis of novel inhibitors. *J. Biomol. Struct. Dyn.* **19**, 765-774
- Sachs, J. & Malaney, P. (2002): The economic and social burden of malaria. *Nature* **415**, 680-685
- Salas, F., Fichmann, J., Lee, G.K., Scott, M.D. & Rosenthal, P.J. (1995): Functional expression of falcipain, a *Plasmodium falciparum* cysteine proteinase, supports its role as a malarial hemoglobinase. *Infect. Immun.* **63**, 2120-2125
- Salmon, D., do Aido-Machado, R., Diehl, A., Leidert, M., Schmetzer, O., de, A.L.A.P., Scharfstein, J., Oschkinat, H. & Pires, J.R. (2006): Solution structure and backbone dynamics of the *Trypanosoma cruzi* cysteine protease inhibitor chagasin. *J. Mol. Biol.* **357**, 1511-1521
- Santos, C.C., Scharfstein, J. & Lima, A.P. (2006): Role of chagasin-like inhibitors as endogenous regulators of cysteine proteases in parasitic protozoa. *Parasitol. Res.* **99**, 323-324
- Schägger, H. & von Jagow, G. (1987): Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem.* **166**, 368-379
- Schwarzenbacher, R., Godzik, A., Grzechnik, S.K. & Jaroszewski, L. (2004): The importance of alignment accuracy for molecular replacement. *Acta Cryst.* **D60**, 1229-1236
- Scott, E.E., Gibson, Q.H. & Olson, J.S. (2001): Mapping the pathways for O<sub>2</sub> entry into and exit from myoglobin. *J. Biol. Chem.* **276**, 5177-5188
- Shenai, B.R., Sijwali, P.S., Singh, A. & Rosenthal, P.J. (2000): Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and

## REFERENCES

essential hemoglobinase of *Plasmodium falciparum*. *J. Biol. Chem.* **275**, 29000-29010

Sijwali, P.S., Brinen, L.S. & Rosenthal, P.J. (2001a): Systematic optimization of expression and refolding of the *Plasmodium falciparum* cysteine protease falcipain-2. *Protein Expr. Purif.* **22**, 128-134

Sijwali, P.S., Shenai, B.R., Gut, J., Singh, A. & Rosenthal, P.J. (2001b): Expression and characterization of the *Plasmodium falciparum* haemoglobinase falcipain-3. *Biochem. J.* **360**, 481-489

Sijwali, P.S., Shenai, B.R. & Rosenthal, P.J. (2002): Folding of the *Plasmodium falciparum* cysteine protease falcipain-2 is mediated by a chaperone-like peptide and not the prodomain. *J. Biol. Chem.* **277**, 14910-14915

Sijwali, P.S., Kato, K., Seydel, K.B., Gut, J., Lehman, J., Klemba, M., Goldberg, D.E., Miller, L.H. & Rosenthal, P.J. (2004): *Plasmodium falciparum* cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites. *Proc. Natl. Acad. Sci. USA* **101**, 8721-8726

Sijwali, P.S. & Rosenthal, P.J. (2004): Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **101**, 4384-4389

Sijwali, P.S., Koo, J., Singh, N. & Rosenthal, P.J. (2006): Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **150**, 96-106

Sim, T.S., Loke, P., Lee, M.A., Singh, M. & Flotow, H. (2001): Cloning and sequence characterisation of falcipain-2 from *Plasmodium falciparum* Gombak A strain (Malaysia). *Parasitol. Res.* **87**, 683-686

Singh, A. & Rosenthal, P.J. (2001): Comparison of efficacies of cysteine protease inhibitors against five strains of *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* **45**, 949-951

Singh, A., Shenai, B.R., Choe, Y., Gut, J., Sijwali, P.S., Craik, C.S. & Rosenthal, P.J. (2002): Critical role of amino acid 23 in mediating activity and specificity of vinckepain-2, a papain-family cysteine protease of rodent malaria parasites. *Biochem. J.* **368**, 273-281

Singh, N., Sijwali, P.S., Pandey, K.C. & Rosenthal, P.J. (2006): *Plasmodium falciparum*: biochemical characterization of the cysteine protease falcipain-2'. *Exp. Parasitol.* **112**, 187-192

Singh, A., Walker, K.J., Sijwali, P.S., Lau, A.L. & Rosenthal, P.J. (2007): A chimeric cysteine protease of *Plasmodium berghei* engineered to resemble the *Plasmodium falciparum* protease falcipain-2. *Protein Eng. Des. Sel.* **20**, 171-177

## REFERENCES

- Snow, R.W., Guerra, C.A., Noor, A.M., Myint, H.Y. & Hay, S.I. (2005): The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**, 214-217
- Sobolewski, P., Gramaglia, I., Frangos, J.A., Intaglietta, M. & van der Heyde, H. (2005): *Plasmodium berghei* resists killing by reactive oxygen species. *Infect. Immun.* **73**, 6704-6710
- Steller I, Bolotovskiy R and Rossmann MG. (1997): An Algorithm for Automatic Indexing of Oscillation Images using Fourier Analysis. *Appl. Cryst.* **30**, 1036-1040.
- Stocker, R., Hunt, N.H., Buffinton, G.D., Weidemann, M.J., Lewis-Hughes, P.H. & Clark, I.A. (1985): Oxidative stress and protective mechanisms in erythrocytes in relation to *Plasmodium vinckei* load. *Proc. Natl. Acad. Sci. USA* **82**, 548-551
- Storoni, L.C., McCoy, A.J. & Read, R.J. (2004): Likelihood-enhanced fast rotation functions. *Acta Cryst.* **D60**, 432-438
- Surolia, N. & Padmanaban, G. (1991): Chloroquine inhibits heme-dependent protein synthesis in *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **88**, 4786-4790
- Surolia, N. & Padmanaban, G. (1992): *de novo* biosynthesis of heme offers a new chemotherapeutic target in the human malarial parasite. *Biochem. Biophys. Res. Commun.* **187**, 744-750
- Taylor, W.R. & White, N.J. (2004): Antimalarial drug toxicity: a review. *Drug Saf.* **27**, 25-61
- Terwilliger, T.C. (2004): Using prime-and-switch phasing to reduce model bias in molecular replacement. *Acta Cryst.* **D60**, 2144-2149
- Than, M.E., Helm, M., Simpson, D.J., Lottspeich, F., Huber, R. & Gietl, C. (2004): The 2.0 Å crystal structure and substrate specificity of the KDEL-tailed cysteine endopeptidase functioning in programmed cell death of *Ricinus communis* endosperm. *J. Mol. Biol.* **336**, 1103-1116
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. (1997): The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882
- Vallone, B., Nienhaus, K., Brunori, M. & Nienhaus, G.U. (2004): The structure of murine neuroglobin: Novel pathways for ligand migration and binding. *Proteins* **56**, 85-92
- Vander Jagt, D.L., Baack, B.R. & Hunsaker, L.A. (1984): Purification and characterization of an aminopeptidase from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **10**, 45-54
- Vriend, G. (1990): WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph* **8**, 52-56

## REFERENCES

- Wang, S.X., Pandey, K.C., Somoza, J.R., Sijwali, P.S., Kortemme, T., Brinen, L.S., Fletterick, R.J., Rosenthal, P.J. & McKerrow, J.H. (2006): Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease. *Proc. Natl. Acad. Sci. USA* **103**, 11503-11508
- Wang, S.X., Pandey, K.C., Scharfstein, J., Whisstock, J., Huang, R.K., Jacobelli, J., Fletterick, R.J., Rosenthal, P.J., Abrahamson, M., Brinen, L.S., Rossi, A., Sali, A. & McKerrow, J.H. (2007): The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. *Structure* **15**, 535-543
- Yayon, A., Cabantchik, Z.I. & Ginsburg, H. (1984): Identification of the acidic compartment of *Plasmodium falciparum*-infected human erythrocytes as the target of the antimalarial drug chloroquine. *EMBO J.* **3**, 2695-2700

## PUBLICATIONS AND PRESENTATIONS

Hogg, T., Nagarajan, K., Herzberg, S., Chen, L., Shen, X., Jiang, H., Wecke, M., Blohmke, C., Hilgenfeld, R. and Schmidt, C.L. (2006): Structural and functional characterization of Falcipain-2, a hemoglobinase from the malarial parasite *Plasmodium falciparum*. *J. Biol. Chem.* **281**, 25425-25437.

Tan, J., Kusov, Y., Mutschall, D., Tech, S., Nagarajan, K., Hilgenfeld, R., Schmidt, C.L. (2007): The "SARS-unique domain" (SUD) of SARS coronavirus is an oligo (G)-binding protein. *Biochem. Biophys. Res. Commun.* **364**, 877-882.

*International Symposium on Recent Trends in macromolecular Structure & Function*, Chennai, India, 18<sup>th</sup> to 20<sup>th</sup>, January 2006.

**Oral Presentation:** Structural analysis of the *Plasmodium falciparum* cysteine protease falcipain-2.

*Mini-Symposium „Biochemistry of the pathogen microbe interaction“* in collaboration with the Collaborative Research Center 34 and Network of Excellence, Kloster Banz near Staffelstein, Germany, 13<sup>th</sup> to 15<sup>th</sup>, November, 2006.

**Oral Presentation:** *Plasmodium* Proteases: New targets for Malaria therapy.

*Schwerpunktprogramm „Intrazelluläre Lebensformen“ (SPP 1131/3)* Kolloquium in Marburg, Germany, 13<sup>th</sup> to 14<sup>th</sup> November, 2007.

**Oral Presentation:** The *Plasmodium* Protease Falcipain-2: Substrate Recognition and Inhibition

*23<sup>rd</sup> European Crystallographic Meeting (ECM-23)*, Leuven, Belgium, 6<sup>th</sup> to 11<sup>th</sup>, August 2006.

**Poster Presentation:** Structural and biochemical analysis of the *Plasmodium falciparum* cysteine protease falcipain-2

*Herbsttagung der Gesellschaft für Biochemie und Molekularbiologie (GBM Fall meeting)*, Hamburg, Germany, 16<sup>th</sup> to 19<sup>th</sup> September 2007.

**Poster Presentation:** Structural and biochemical analysis of the *Plasmodium falciparum* cysteine protease falcipain-2.

*59. Jahrestagung der Deutschen Gesellschaft für Hygiene und Molekularbiologie (DGHM 2007)*, Göttingen, Germany, 30<sup>th</sup> to 4<sup>th</sup> October, 2007.

**Poster Presentation:** Structural and biochemical analysis of the *Plasmodium falciparum* cysteine protease falcipain-2

## PERSÖNLICHE DATEN

Name: Krishna Nagarajan

Geburtsdatum und -ort: 11.11.1981 in Mayiladuthurai, Indien

Familienstand: ledig



### Schulische Ausbildung:

Juni 1987 bis Mai 1988 Sai Vidhyalaya Matriculation School, Chennai, Indien

Juni 1988 bis Mai 1992 P.S. Senior Higher Secondary School, Chennai, Indien

Juni 1992 bis Mai 1997 Rotary Club Matriculation School, Mayiladuthurai, Indien

Juni 1997 bis Mai 1999 D.B.T.R National Higher Secondary School, Mayiladuthurai, Indien

### Studium:

August 1999 bis Mai 2002 A.V.C. College of Arts and Science (Bharathidasan University), Mayiladuthurai, Indien  
Abschluß: Bachelor of Science in Physics

August 2002 bis Mai 2004 University of Madras, Chennai, Indien  
Abschluß: Master of Science in Biophysics

Seit August 2004 Anfertigung der vorliegenden Dissertation unter der Anleitung von Prof. Dr. Dr. h.c. R. Hilgenfeld am Institut für Biochemie, Universität zu Lübeck, Germany.