Chapter 1

Introduction

Perhaps a third of all characterized proteins have metal ion cofactors with structural significance or functional roles in oxidation-reduction or chemical transformations such as atom transfer and hydrolysis (hydrolases).

Hydrolase enzymes, which cleave substrates with the incorporation of water, fall into the general categories of glycosidases, lipases, esterases (including nucleases, phosphatases, and kinases), polymerases, and proteases. Other hydrolases which act on more select metabolic substrates include carbonic anhydrase, arginase, urease, and β-lactamase. The relative 'simplicity' of the hydrolytic reaction in terms of its chemistry and molecularity is key to the robust applications these types of enzymes have found in biotechnology: amylases (glycosidases) in sugar processing, proteases in detergents, and more recently nucleotide polymerases in PCR.

Proteases primarily hydrolyse peptide bonds and can be divided into four general varieties: aspartic, cysteine, serine, and metallo. Aspartic or acid proteases, such as HIV protease, have two conserved catalytically essential aspartic acids in the active site, cysteine proteases such as papain from grapefruit have a similar constellation of active site residues as the serine proteases except for the eponymous catalytic nucleophile. Futher classification of proteases gives endo- vs. exo- proteases, the former cleave internal peptide bonds, and the latter peptide bonds at the terminii of (poly)peptides sequentially releasing either amino acids or di- and tri- peptides.

Metalloendoproteases are found in all forms of life, from bacteria and fungi to plants, insects, and animals and are divided into five main families by their primary sequence: thermolysin, astacin, serralysin, matrixin, and snake venom-related metalloproteinases (Rawlings and Barrett 1995; Jiang and Bond 1992; Hooper 1994)

(Table 1). Although there are distinctly conserved sequences within each family, all these proteases share the catalytic metal-binding motif, HEXXH, which contains two metal-ligating histidines, a non-ligating glutamate that is poised near the metal, and two additional amino acids of variable identity. Other metal ligands include an aquo species (water or hydroxide) and a tyrosine in families other than thermolysin. Exceptions to this classification scheme include another family of endoproteases containing an 'inverted' consensus sequence, HXXEH, a small bacterial endoprotease with a HEXXHXXGXXD motif (Harada, Kinoshita et al. 1995), and D-Ala-D-Ala carboxypeptidase which has three histidine metal ligands in a unique sequence motif. A related HEXXH hydrolase is peptide deformylase which removes formyl groups from nascent polypeptides in prokaryotes.

Table 1 -Metalloprotease Families and Subfamilies

Metalloprotease family	Other representative members	Physiological role	Consensus sequence	Conserved features
Endopeptidases				
Thermolysin (Gluzincins)[1] (clan MA)[2]	neprilysin (mammals), (M13)[2] mycolysin (*Streptomyces*), (M5) angiotensin converting enzyme (ACE) elastase (*Pseudomonas aerudinosa*) neurolysin (*Rattus norvegicus*) (M3) thermolysin is subfamily M4	hormonal regulation of blood pressure (ACE, neprilysin, neurolysin) opioid/pain responses (neprilysin, neurolysin)	HEXXH and third metal ligand E ~20 residues later	five-stranded β-sheet and two active site helices that contain the metal ligands neprilysin subfamily has two-stranded sheet, short N-terminal cytoplasmic domain, transmembrane helix, large C-terminal extracellular domain containing the active site clan M3 (including PfuCP) has a serine just below catalytic metal (Ser-turn)

Underlined residues are metal ligands and individual enzymes can fall into several families depending on the structure/sequence criteria used for similarity comparisons
[1]Nomenclature of (Rawlings and Barrett 1995) and (Hooper 1994)
[2]Clan and subfamily groupings based on additional conserved sequence features as described in (Rawlings and Barrett 1995)

Table 1 Cont'd

Astacin (Metzincin)[1] (clan MB)	meprins, bone morphogenic protein-1 astacin defines subfamily M12	vasopeptide cleavage (meprins)	HEXXHXXGXXHE...HY	clan MB has five-stranded β-sheet but metal ligands from one helix and coil region methionine residue below active site metal (Met-turn)
Serralysin (Metzincin) (clan MB)	Protease B and C (*Erwinia chrysanthemi*) subfamily M10		HEXXHXXGXXHP...SY	Met-turn
Matrixin (Metzincin) (clan MB)	membrane bound members include collagenase, gelatinase, stromelysin subfamily M10	degradation of extracellular matrix	HEXXHXXGXXHS...(I/L)Y	Met-turn Collagenase subfamily has second structural zinc and three domains (pro, catalytic, hemopexin-like)

Table 1 Cont'd

Snake-venom related proteinase (Metzincin) (clan MB)	adamalysin, reprolysin, atrolysin	degradation of extracellular matrix	HEXXHXXGXXHD ...YY	Met-turn
Inverzincin	pitrilysin, insulin degrading enzyme (rat, human, Drosophila)		HXXEH and third histidine ligand 82 residues C-terminal	
Exopeptidases				
Carboxypeptidase	digestive CPs (A, B), regulatory (H, M) membrane-bound members include M, D, P dimetal CPs (*Streptomyces griseus*) DD-CPs (*Streptomyces albus*) PfuCP (*Pyrococcus furiosus*) (M32)	digestion, peptide hormone activation and deactivation, peptide turnover	'classical' CPs: HXXE and third metal ligand 108-135 residues C-terminal DD-CPs: DHXHV and third H ligand 42 residues N-terminal in SRHMY	
Aminopeptidase	leucine AP, leukotriene A4 hydrolase dimetal APs (leucine AP, *A. proteolytica*, *S. griseus*, Methionine AP from *E. coli*)		VBXHEBXHXWFG and LWLNEGB	

Exopeptidases are similarly ubiquitous and can be further divided into amino- or carboxy- depending on which end of the poypeptide they act, and both subfamilies contain members with one or two essential metals.

The importance of proteases to human health is underscored by their involvement in common diseases (Table 2).

Table 2 –Representative Proteases in Disease

Protease	Disease
HIV (aspartic)	
Matrix metalloproteases (metallo)	Invasive cancer, angiogenesis, arthritis
ACE (metallo), renin (aspartic)	Hypertension, cardiovascular
Beta-secretase (aspartic)	Alzheimer's
Gingipain (cysteine)	Periodontal
AAA-protease (metallo)	Mitochondrial (neuronal ceroid-lipofuscinosis, spastic paraplegia)
Chymase (serine)	Chronic inflammation
Calpain (cysteine)	Neurodegenerative (stroke, muscular dystrophy, amyotrophic lateral sclerosis, Parkinson's, Huntington's)

Their uses in biotechnology are also widespread from bioremediation to research reagents in chemical syntheses and protein chemistry, but one particular application of

exopeptidases that remains largely undeveloped, and which parallels the use of DNA polymerases in PCR and DNA sequencing, is enzymatic protein sequencing. By progressively releasing amino acids from the ends of polypeptides, their primary sequence can be deduced by time-resolved amino acid analysis or in a more recent development, mass spectrometric analysis of the truncated peptides (Schar, Bornsen et al. 1991). Earlier work with mesophilic carboxypeptidases has given sporadic results with remarkably few peptides and proteins that are amenable to extensive digestion, even when the specificities of the enzyme essentially allow for release of all 20 common amino acids. One explanation is that at lower temperatures, the (poly)peptide substrate retains too much structure, secondary or otherwise, which blocks access of the enzyme to a progressively shortened C-terminus. As enzymes from hyperthermophilic microorganisms are known to be more active and stable at elevated temperatures and under otherwise denaturing conditions, a new carboxypeptidase (PfuCP) from the hyperthermophilic archaeon, *Pyrococcus furiosus*, was purified and characterized to assess its potential for high temperature C-terminal protein sequencing.

However, optimization of a carboxypeptidase as a sequencing tool requires understanding the fundamentals of how the enzyme behaves, and although the mechanics of hydrolytic proteolysis appear simple enough, PfuCP exhibits intriguing and at times contradictory behavior that hints at deeper issues:

1) Thorough steady-state velocity studies reveal non-Michaelis-Menten-like kinetics with allosteric features which may implicate either an active dimeric or monomeric form (or perhaps a mixture of both).

2) The enzyme inactivates at high temperature yet the global structure appears remarkably thermostable.
3) The dimeric-monomer disposition varies unpredictably with conditions (denaturant, temperature, chelator).
4) There is an obligatory catalytic Co^{2+} but the metal is very loosely associated.
5) There is no esterase activity.
6) The metal-binding amino acid sequence motif is characteristic of well-known zinc metalloendopeptidases, yet zinc does not reactivate the *apo*enzyme (metal specificity is also extremely strict as no divalent metal ion aside from Co^{2+} activates the enzyme). A corollary issue is why are cobalt substituted forms generally more active than native zinc holoenzymes?
7) The specificity for the C-terminal amino acid released is very broad although the size of the peptides amenable to cleavage is limited to about 30 residues or less.
8) There are unknown structural idiosyncracies that severely complicate expression and refolding.

This thesis represents the first steps into investigating this newly discovered enzyme and details exploration of the above issues in two main sections: I) the enzymology of PfuCP (Chapters 2-5) and II) its applications in protein sequencing (Chapter 6). Chapter 2 (Purification, Cloning, Expression, Refolding) and Chapter 4 (Steady-State Kinetics) describe the biochemical characteristics of PfuCP while Chapter 3 (Crystal Structure) explores structural aspects as they relate to function. Chapter 5 (Implications for Mechanism, Substrate Specificity, and Rate Theory) places

experimental observations into the larger context of metalloprotein behavior and theory. Finally, much work that has been done to optimize sequencing protocols will not be discussed here but will appear in future patent literature.

Chapter 2

Purification, Cloning, Expression, and Refolding of Native and Recombinant PfuCP

Abstract

Native PfuCP was purified from *Pyrococcus furiosus* cells by multistep chromatography under anaerobic conditions. The gene was cloned from a shotgun library of genomic DNA and subcloned into both insect (baculovirus) and bacterial (*E. coli*) expression systems. Although the insect cells produced soluble recombinant, expression in bacteria proved difficult and only one of four bacterial vectors expressed, invariably producing inclusion bodies that required refolding from 10 M urea. Three methods of refolding were compared and on-resin buffer exchange was determined to be the optimal strategy.

Introduction

Hyperthermophiles are microorganisms that thrive at elevated temperatures in such exotic locales as terrestrial hot springs, deep sea vents, and in the case of *Pyrococcus furiosus*, shallow marine sediments (solfatara) off the coast of Vulcano Island, Italy. Together with other such microorganisms which flourish under extremes of cold (psychrophiles), pressure (barophiles), ionic strength (halophiles), and pH (acidophiles and alkaliphiles), hyperthermophiles complete the domain of life known as extremophiles which count among their number both bacteria and archaebacteria (archaea). Among the most evolutionarily 'primitive' lifeforms known today, their distinct identity as the third kingdom of life arises from their S16 RNA (Woese, Kandler et al. 1990).

One of the first hyperthermophiles to be discovered and extensively studied, *P. furiosus* (Pf), is a heterotrophic, anaerobic archaeon that uses complex carbohydrates and peptides as sources of carbon and energy (Fiala and Stetter 1986; Adams and Kelly 1998). Metabolic end products include organic acids, alanine, CO_2, and H_2 (Kengen and Stams 1994; Kengen, Stams et al. 1996) and additional energy is generated when excess redox equivalents are channeled into elemental sulfur (Schicho, Ma et al. 1993). Although many proteases have been characterized from Pf (Tsunasawa, Izu et al. 1997; Voorhorst, Eggen et al. 1996; Snowden, Blumentals et al. 1992; Harwood, Denson et al. 1997; Ghosh, Grunden et al. 1998), no carboxypeptidase has previously been detected.

In the non-archael world, CPs are ubiquitous in animals, plants, and microorganisms being widely distributed among tissue types in soluble as well as membrane-bound forms. Sizes range from ~35 kDa (CPA and CPB) to ~180 kDa (CPC, CPD) and monomers, dimers, and tetramers have been found. Representative members of metallo, serine, and cysteine carboxypeptidases have been discovered but to date no aspartic CPs have characterized. Physiologically, CPs have been implicated in protein degradation/turnover, maturation and processing of hormonal peptides (Steiner 1998), and digestive metabolism of proteins and peptides. The majority of known CPs have temperature optima below 40 ^{o}C, however recently, thermostable CPs have been isolated from *Thermoactinomyces vulgaris* (Stepanov 1995), *Thermus aquaticus* (Lee, Minagawa et al. 1992; Lee, Taguchi et al. 1994; Sterner, Kleemann et al. 1996), *P. horikoshii* OT3 (Ishikawa, Ishida et al. 2001), and *Sulfolobus solfataricus* (Colombo, D'Auria et al. 1992).

Purification of Native PfuCP

The native enzyme was purified from *P. furiosus* cells under anaerobic conditions by multistep chromatography involving anion exchange (QFF), hydroxyapatite, hydrophobic interaction, and gel filtration resins yielding 11 mg of protein from 21 g of crude protein giving a 174-fold purification and 9% recovery (Cheng, Ramakrishnan et al. 1999). The response of the activity to inhibitors included in the purification buffers (no inhibition for 1 mM PMSF and the abolishment of activity in 1 mM EDTA) established PfuCP as a metalloprotease. SDS-PAGE and MALDI-TOF MS gave a molecular mass of 58 kDa and 59 kDa, respectively (Fig.1) while gel filtration at room temperature identified the native form as a dimer at ~128 kDa (data not shown).

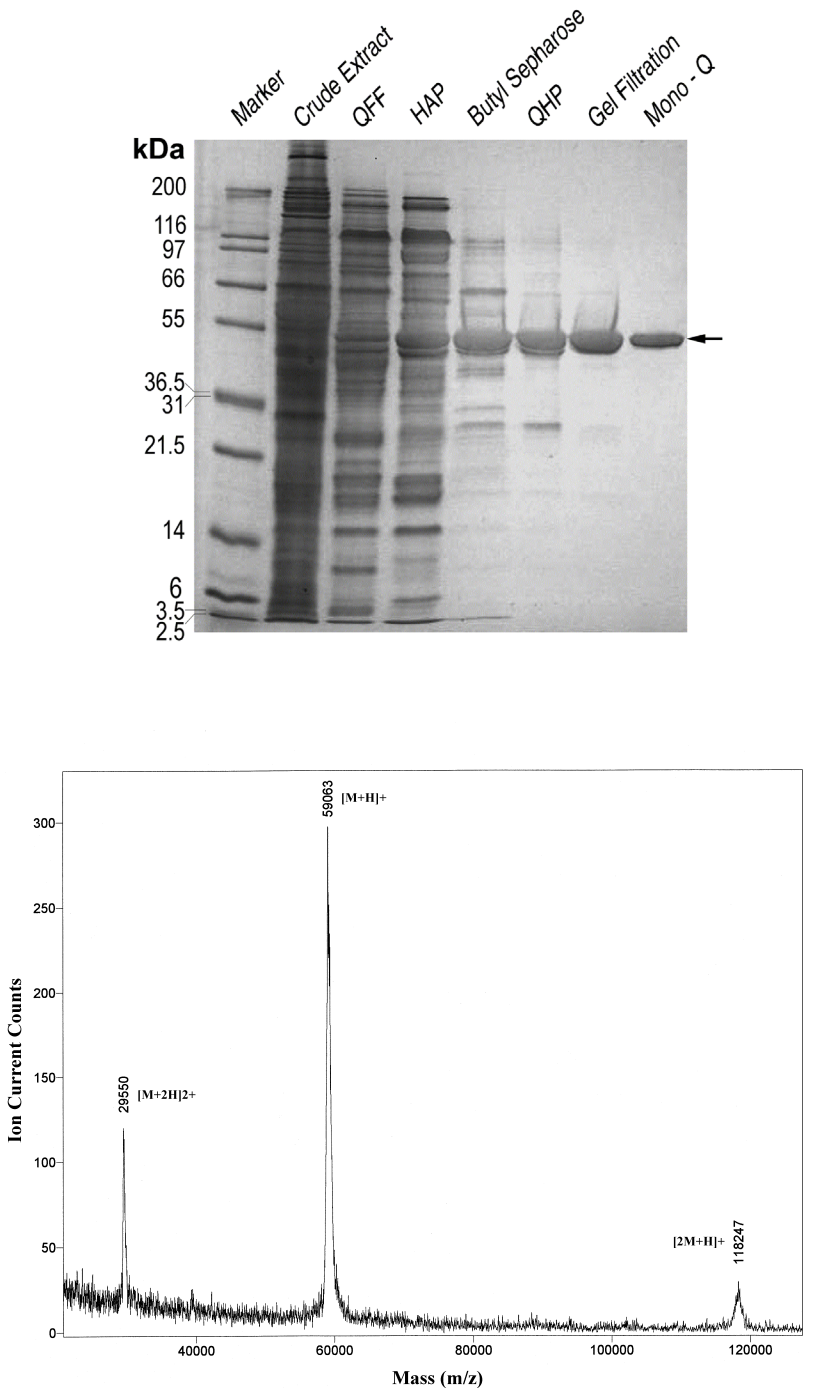


Fig. 1 Purification of native PfuCP from *P. furiosus* cells. Upper: Denaturing and reducing SDS-PAGE of the most active chromatography fractions; samples were pre-incubated at 80 ^{o}C for 1 min in 0.1 % SDS and run on a 12 % Tris-glycine gel. Lower: MALDI-TOF MS on the purified enzyme (± DTT)

Cloning, Subcloning, and Expression in Insect and Bacterial Cells

The gene was subsequently cloned into a pET24a (Bluescript) plasmid vector from genomic DNA through PCR and other standard molecular biology techniques (Maniatis and Sambrook 1989; Harwood 1996; Brown 2000) using degenerate probes designed from N-terminal protein sequence tags derived from fragments of the purified, native enzyme[1]. This generated an Xba size-selected library which enabled further subcloning into three different bacterial expression vectors, however no recombinant could be induced:

- N-terminal His-Tag (Qiagen, pQE, T5 promoter)
- C-terminal His-Tag (Qiagen, pQE, T5 promoter)
- GST-Tag (Invitrogen, pDEST, T7 promoter)

There are several possible explanations including instability of mRNA or of the nascent polypeptide, secondary structure in the mRNA transcript that hinders translation, effects of differential codon usage between bacteria and archaea, or perhaps even an obligatory chaperone that is present in native cells (Tuan 1997).

Further subcloning into an insect (baculovirus) expression system (His-Tag, pDEST, p10 promoter, Invitrogen) did yield soluble and folded recombinant but the attached His-Tag yielded spurious (and time dependent) optical absorbance during metal reconstitution

and kinetic experiments, so the gene was further subcloned into a pROTet *E. coli* expression system with an enterokinase cleavage site for removal of the His-Tag (Clontech, Ltet0-1 promoter, anhydrotetracycline (aTc) inducer). Overexpression was achieved; however, the recombinant (rPfuCP) was entirely concentrated in inclusion bodies.

Several strategies were undertaken to reduce the formation of inclusion bodies (Higgins and Hames 1999):

1) reduction of induction temperature from 37 ^{o}C to 25 ^{o}C and 21 ^{o}C;
2) change in the expression strain from BL to DH5α *E. coli* cells;
3) reduction in the concentration of inducing agent from 100 to 75, 55, and 35 ng/mL.

Generally speaking, a reduced temperature and lower aTc concentration delayed induction, however in all cases including trials with the DH5α strain, the results were the same meaning centrifugation of the cell lysate after fermentation and induction pelleted the recombinant in inclusion bodies without any trace of soluble rPfuCP in the supernatant (Fig. 2).

[1] Detailed protocol in Appendix A

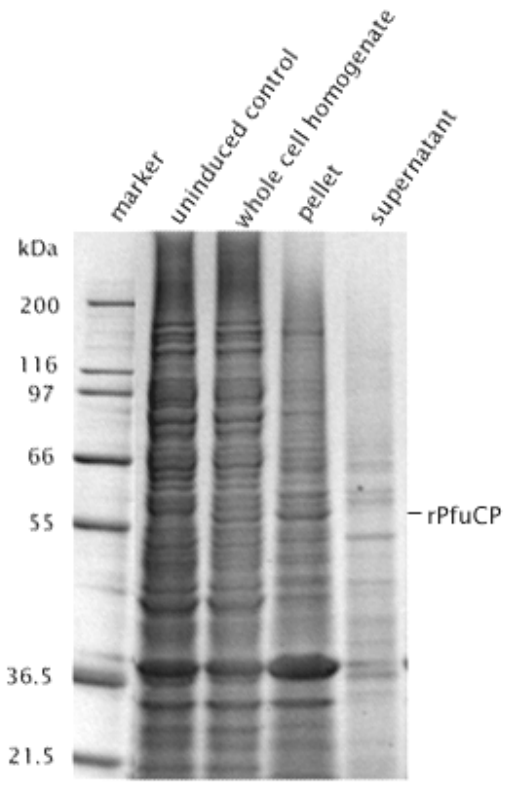


Fig. 2 Representative expression pattern from induction trials to reduce inclusion body formation: the production of recombinant entirely in inclusion bodies was invariant whether the *E. coli* strain was changed from BL to DH5α, the induction temperature lowered from 37 °C to 21, 25, 30 °C, or the concentration of the inducer anhydrotetracycline (aTc) reduced from 100 to 35, 55, and 75 ng/mL. A distinct band for the recombinant is seen in the whole cell homogenate and pelleted cell lysate but not in the either the uninduced control or the supernatant from centrifuged lysate.

As inclusion body production appeared inevitable, overexpression was maximized by increasing the concentration of anhydrotetracycline to 1 μg/mL and the induction period to 24 h (Fig. 3). The yield was not improved further with a longer induction nor with more aTc).

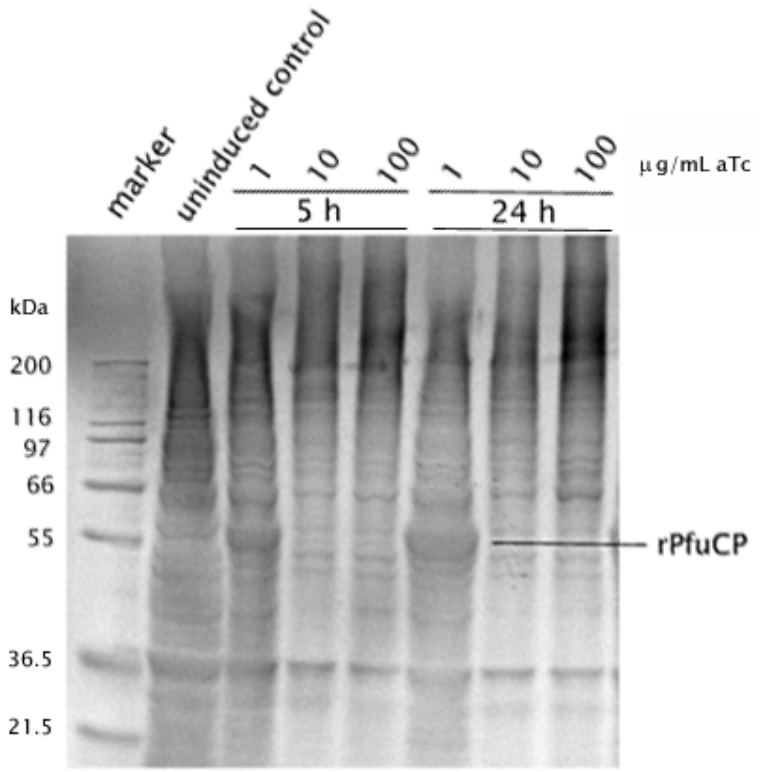


Fig. 3 Maximum overexpression of rPfuCP in BL cells (*E. coli*) at 37 °C was achieved after 24 h at 1 μg/mL anhydrotetracycline (control is 24 h uninduced cells).

Surprisingly, extreme denaturing conditions were required to solubilize the inclusion bodies (Fig. 4). Only 10 M urea at the solubility limit of the denaturant was effective, not lesser concentrations nor the detergents Triton X-100 (0.1-1 %) and sodium deoxycholate (0.1-1 %).[2] Many other proteins were extracted by both detergents, suggesting that pre-washing the inclusion body pellets with these agents would achieve some measure of purification, however, this was found to complicate subsequent refolding and chromatography due to reduced binding to the resin and foaming effects.

[2] 7 M guanidium was also effective but not used due to its degradative effects on the IMAC (immobilized metal affinity chromatography) resin used in the subsequent refolding (see below). Isopropanol (40-60%) was not effective in solubilizing the inclusion bodies.

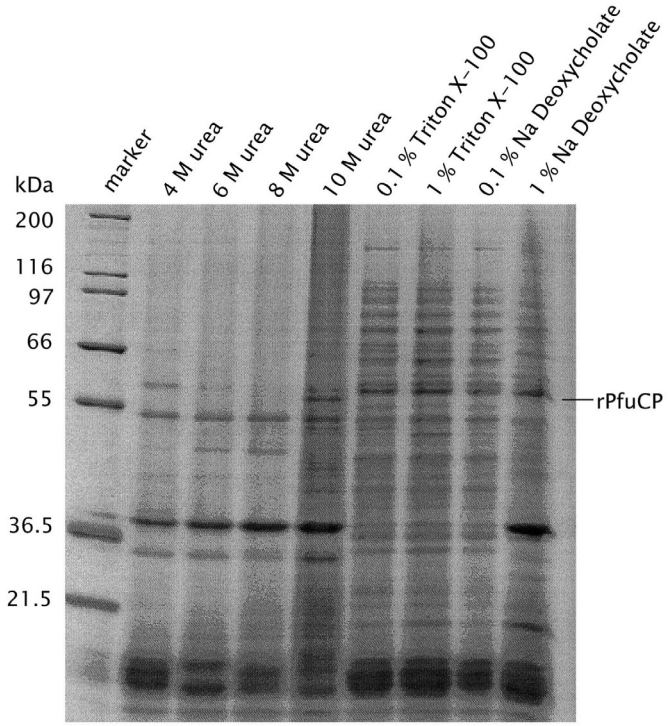


Fig. 4 Solubilization of the inclusion bodies: a band for rPfuCP is only visible following resuspension of the inclusion bodies in 10 M urea.

Refolding of Recombinant from Inclusion Bodies

Refolding[3] of PfuCP from inclusion bodies in bacterial cells has several advantages: the yield of expressed protein is greater than for insect cells (and the cost lower), there is an immediate and efficient separation of the recombinant from soluble proteins and other cellular components, and the bulk of the recombinant is accumulated in a protease-resistant state that is not susceptible to turnover either during fermentation or following cell lysis. Centrifugation of crude lysate readily isolates the recombinant which can then be refolded by the following methods:

[3] Detailed protocols in Appendix B

1) rapid dilution of a concentrated solution of inclusion body + denaturant;
2) buffer exchange, with simultaneous dilution of both denaturant and recombinant as the sample passes through a gel filtration column;
3) buffer exchange, with simultaneous dilution of denaturant where the refolding protein is attached to a resin column.

Refolding by method 1 was accomplished by rapid 100-fold dilution of a 10 M urea solution of inclusion body (80 mg/mL) into refolding buffer (20 mM MOPS pH 7 + 100 mM NaCl) inside the mixing apparatus of a Pharmacia FPLC system (LC-500) at room termperature, followed by passing the diluted solution through a Co^{2+} IMAC column equilibrated at 4 ^{o}C. Extensive concentration of the column eluant after imidazole elution (to ~33 μL) allowed detection of activity at the threshold of the standard assay (~0.006 mM/min) but the amount of refolded protein remained below the detection threshold of SDS-PAGE (5 ng by Coomassie stain) which suggested that less than 5 ppb of the originally loaded sample refolded. The low yield was not improved by the addition of either 1 mM $CoCl_2$, 1 mM EDTA, 1 mM PMSF, or various 'stabilizers' that have been reported to improve refolding (0.5% PEG 600, 0.5% PEG 8000, 1 M triethanolamine) nor was the yield affected by lowering the refolding temperature to 4 ^{o}C or changing the pH of the refolding buffer to 6 or 8, so further optimization was discontinued. The appeal of this method would be in its technical simplicity and high throughput (i.e. potential for large loadings per run with minimum handling of sample).

Refolding by method 2 was accomplished by 250 µL injections of a 10 M urea solution of concentrated inclusion body solution (40 mg/mL) onto a 24 mL prepacked Superdex G-200 gel filtration column at 0.4 mL/min with an elution buffer of 20 mM MOPS pH 7 + 150 mM NaCl + 1 mM PMSF. Further optimization of gel filtration methods was discontinued due to the limited sample volume that could be loaded and the low net yield of refolded protein which is estimated from SDS-PAGE at <0.2 ppt of the original loaded sample. Similar trials with two Pharmacia HiTrap™ desalting columns connected in tandem did not resolve separate peaks for urea and protein such that denatured protein eluted in the pass-through.

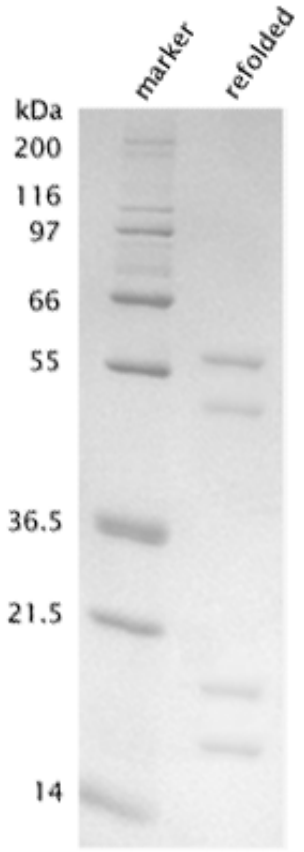


Fig. 5 Protein refolded by gel filtration (method 2). Aliquots of 250 µL of a 10 M urea solution of inclusion bodies (40 mg/mL) were loaded onto a Superdex G-200 at 0.4 mL/min with an elution buffer of 20 mM MOPS pH 7 + 150 mM NaCl + 1 mM PMSF. Refolded protein was < 0.2 ppt.

Refolding by method 3 (Fig. 6) allowed for the highest yield since the loaded protein sample could be recovered as pass-through and then reloaded again in the next refolding cycle (load, wash, refold, elute; see Appendix B for details). Furthermore, the selective attachment of denatured PfuCP to the Co^{2+} IMAC column allowed for a simultaneous purification. A single step-gradient from 10 to 0 M urea by buffer exchange with 20 mM MOPS pH 7 + 100 mM NaCl was then used to refold the enzyme on-resin followed by elution with 200 mM imidazole, removal of the His-Tag by enterokinase digestion, and ion-exchange chromatography on a MonoQ column. Another advantage of the on-resin method is that costly protease inhibitors do not need to be added until after the final elution from the IMAC column as proteolysis was reduced by both the high initial concentration of urea and the spatial separation of proteins on the resin.

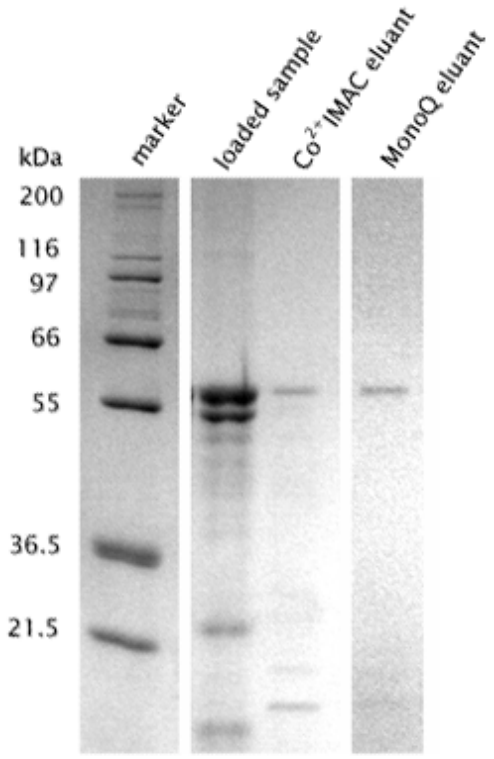


Fig. 6 On-resin refolding (method 3): each cycle saturated the Co^{2+} IMAC column with 50 mL of an 80 mg/mL solution of inclusion body in 10 M urea (+ 20 mM MOPS pH 8 + 100 mM NaCl) followed by refolding in 20 mM MOPS pH 8 + 100 mM NaCl and elution in the same buffer with 200 mM imidazole which yields 2-4 mg/cycle. Six runs in total were pooled yielding ~13 mg of refolded rPfuCP which was then cut by enterokinase (42 ng mL^{-1} at 20 ^{o}C for 16 h) and re-run through the Co^{2+} IMAC to remove uncut recombinant prior to chromatography on an anion exchange MonoQ column.

Guanidium chloride (7 M) was initially used but the IMAC resin gradually degraded reducing the overall yield and preventing multiple cycles of refolding. Pre-washing the inclusion body pellet with 1% Triton X-100, 1% deoxycholate, or 6 M urea did not improve the refolding yield. SDS-PAGE (Fig. 6) confirmed a MW of ~59 257.04 kDa for rPfuCP which is the MW predicted from the cloned gene with the two residual N-terminal amino acids (V and D) from the enterokinase cleavage site.

Biochemical Characterization and Metal Dependence

Substrate Specificity

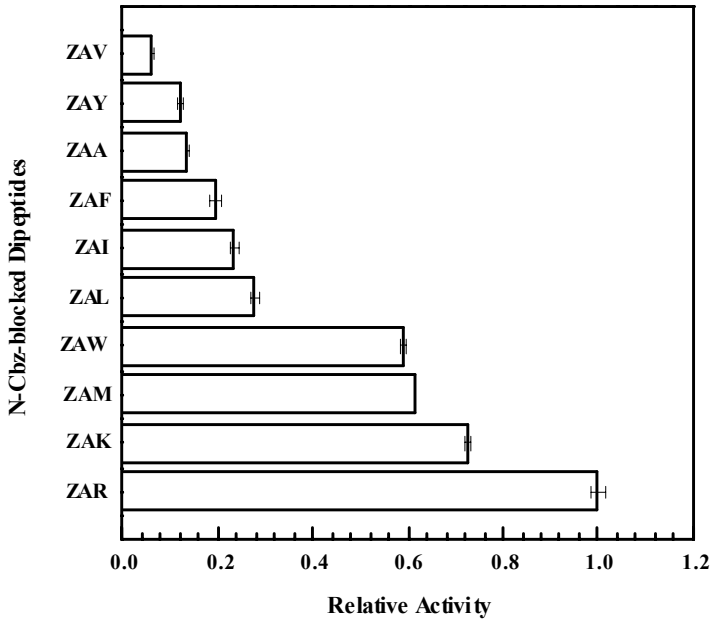


Fig. 7 Representative C-terminal specificities of PfuCP: Data represent duplicate points under standard conditions (10', 2 mM ZAX, 0.1 M KMes pH 6.5, 0.4 mM $CoCl_2$) with 30 nM enzyme. The activity against ZAS, ZAH, and ZAN was determined at longer digest times (30') to have a relative activity < 0.01.

PfuCP cleaves most N-blocked carbobenzoxy-alanyl-amino acid (Z-Ala-Xaa or ZAX) dipeptides with significant activity under standard assay conditions (Fig. 7) although ZAS, ZAH, and ZAN required longer digest times (30') and dipeptides with C-terminal glutamate, aspartate, glycine, or proline were not cut[4]. Too little or too much rigidity may cause poor binding of Gly or Pro to an S1' subsite or unfavorable conformations along the reaction coordinate, while the sidechains of Glu or Asp may cause nonproductive chelation of the catalytically essential metal by the side-chain carboxyl perhaps in tandem with the C-terminus; incidentally, no activity is observed in

0.05 M sodium acetate buffer (pH 5.8) with otherwise standard assay conditions. Unlike mesophilic metalloCPs, which primarily have specificity for hydrophobic (CPA) or basic (CPB) terminal residues, PfuCP and other characterized thermophilic CPs such as CPT from *T. actinomyces* (Stepanov 1995), exhibit a broad specificity for neutral, aliphatic, basic, polar, and aromatic C-terminal residues that is suggestive a more general functionality in earlier evolutionary history.

Although some CPs are active against substrates where the C-terminus is blocked by an alcohol or amine (ester or amide), no esterolytic or aminolytic behavior was seen for PfuCP against the ester ZAR-OMe nor the synthetic peptide ZLY-NH_2 under standard conditions. The absence of esterolytic activity agrees with one report for CPA which found that unlike peptides which can bind to the *apo*enzyme (Christianson 1989), esters only bind when there is a metal ion already attached to the enzyme and the ester K_m varies with substituted metal.

[4] ZAC, ZAQ, and ZAT were not commercially available but the observed activities of ZAS and ZAN were taken to be representative of ZAC/ZAT and ZAQ to a first approximation.

pH Dependence of Activity for Native PfuCP

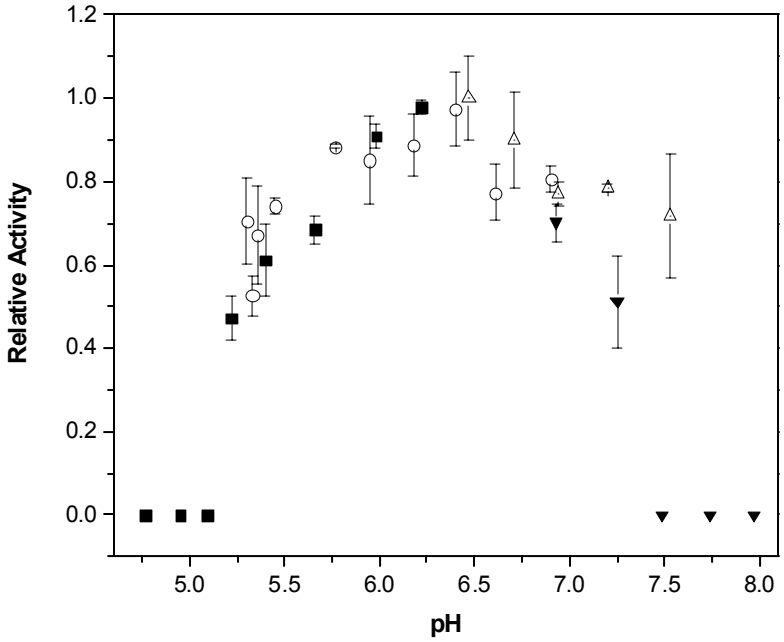


Fig. 8 The pH dependence of PfuCP under standard assay conditions: piperazine (■), pH 4.8-6.2; KMes (○), pH 5.3-6.9; MOPS (Δ), pH 6.5-7.5; HEPES (▼), pH 6.9-8.0. Data shown represent duplicate measurements with standard error normalized to the maximum activity observed in 0.1 M MOPS (pH 6.5).

Temperature Dependence of Activity

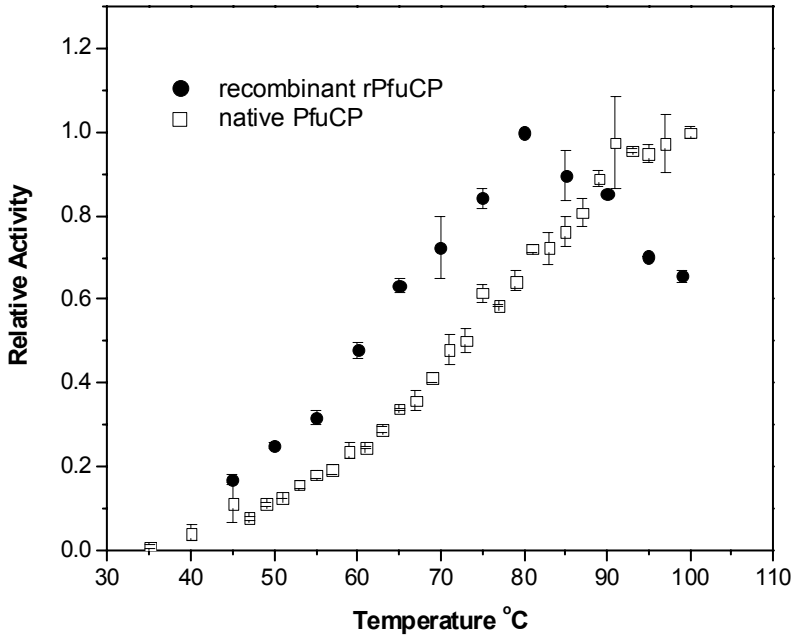


Fig. 9 The temperature dependence of both native and recombinant PfuCP activity: relative activity is shown for convenience as rPfuCP currently remains only partially purified with a sepcific activity of approximately 0.07 units/mg (measured from its peak at 80 °C). Duplicate data is shown with standard error.

The optimal pH range of 6.2-6.6 (Fig. 8) suggests that *P. furiosus* may be able to maintain a near neutral intracellular pH despite potentially extreme variations in acidity from surrounding volcanic sediments. But the temperature optima (Fig. 9) which is greater than 90 °C is perhaps more a reflection of the need to accommodate the native environmental conditions in which *P. furiosus* grows optimally at 100 °C. The difference in the temperature maxima between the native and recombinant may be instability effects from the presence of the His-Tag or structural artifacts from the refolding process[5].

[5] Enterokinase removal of the His-Tag is currently unoptimized and insufficient amounts of stable and refolded protein were recovered for full biophysical characterization.

An apparent activation energy can be calculated for the near linear portion of the temperature dependence data for native enzyme between 59 and 93 ^{o}C yielding E_a = 47.8 kJ mol^{-1} and Q_{10} = 1.19 (no measurements were taken above 100 ^{o}C). Although, no activity was observed in the standard assay below 40 ^{o}C, room temperature activity was observed in C-terminal ladder sequencing experiments (Ch. 6) reflecting the higher sensitivity of mass spectrometry.

Pre-incubation of the enzyme for varying amounts of time at 80 ^{o}C prior to a standard assay reveals that PfuCP inactivates irreversibly with a apparent second order rate constant of 6.25 x $10^{-1} \pm 0.036$ mM^{-1} (Fig. 10), despite an apparently thermostable physical structure; changes in both the circular dichroism (CD) and intrinsic fluorescence spectra are essentially reversible when the temperature is ramped from 20 ^{o}C to 80 ^{o}C (over a similar time period) with only a slight decrease in the negative ellipticity from 205-230 nm and a partial quenching of fluorescence intensity without a noticeable shift in the emission wavelength at 344 nm (Fig. 11). MALDI-TOF MS analysis of control digestion reactions (without substrate) showed no apparent autodigestion or fragmentation of the enzyme.

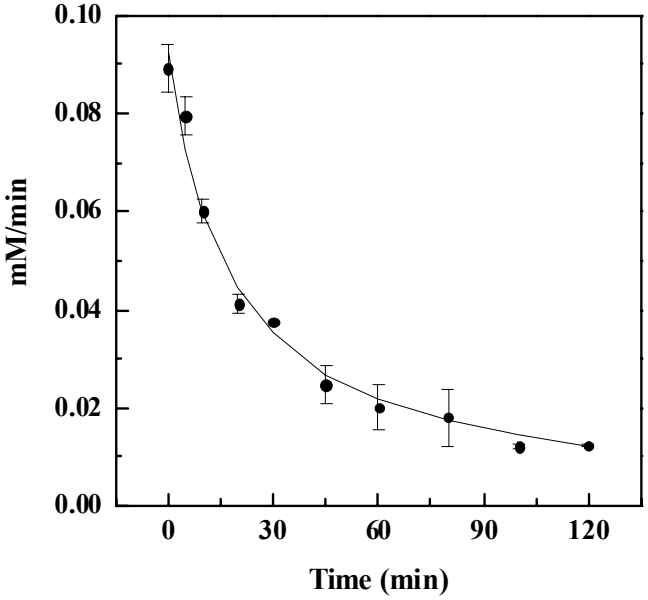


Fig. 10 Apparent second-order thermal inactivation of PfuCP ($6.25 \times 10^{-1} \pm 0.036$ mM^{-1}) in the presence or absence of Co^{2+}: enzyme (1.4 nM) was incubated for varying amounts of time at 80 ^{o}C under otherwise standard conditions prior to addition of substrate to initiate reaction. Duplicate data are shown with standard error.

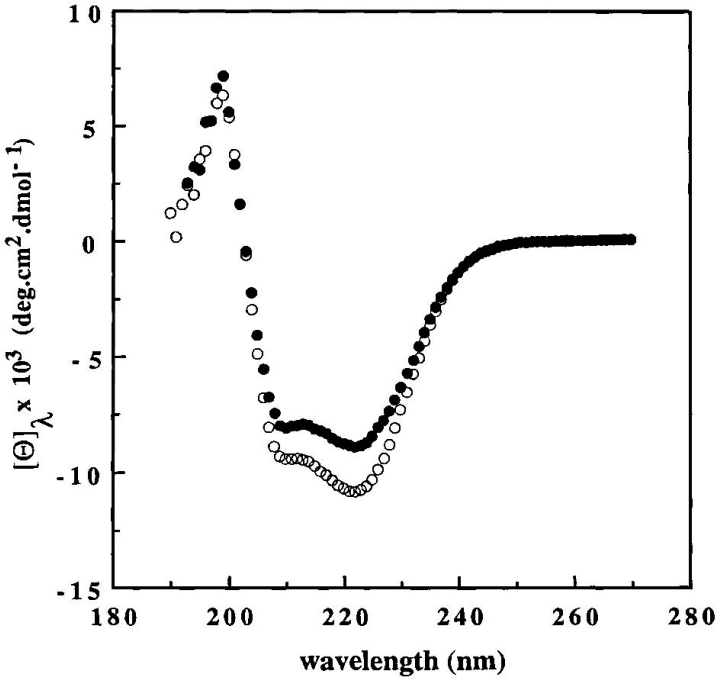


Fig. 11 Far-UV CD showing *apo*PfuCP (15 μM) at room temperature (○) and 80 ^{o}C (●) in 10 mM potassium phosphate (pH 7.0). Minima are at 210 and 222 nm and a maximum at 199 nm.

The thermal inactivation profiles, CD spectra, and fluorescence behavior are invariant in the presence or absence of Co^{2+}, implicating no role for the catalytic metal in either structural thermostability or thermoinactivation, so the mechanism for activity loss remains unclear. Further, this suggests that the presence of Co^{2+} does not contribute additional oxidative effects to the intrinsic inactivation process. One could speculate that contributions to inactivation might include aggregation, local denaturation, and chemical modifications such as hydrolysis of critical asparagines or glutamines, but the second order dependence is suggestive of interactions between two active subunits or monomers.

Oligomeric State

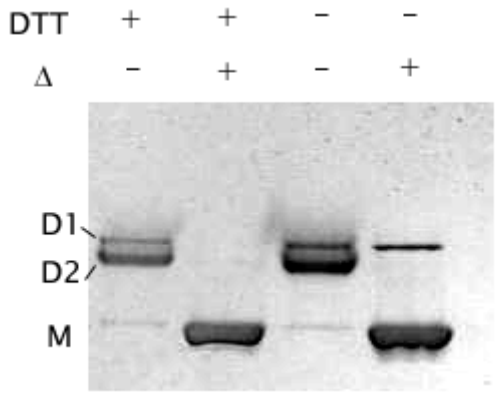


Fig. 12 Oligomeric behavior of PfuCP in 0.1% SDS-PAGE: *apo*PfuCP exists in two dimer forms in the absence of heat but only one dimer form is thermostable in the absence of 2 mM DTT (D1 and D_2 have an apparent MW ~85 kDa while the monomer M appears at 55 kDa)

Both gel filtration and crystallography clearly identify a dimeric PfuCP at room temperature; however, evidence from SDS-PAGE is somewhat more ambiguous (MALDI-TOF MS on the other hand shows a predominant monomer and a trace of dimer

but this distribution likely results from laser desorption efficiencies and the highly acidic sample loading conditions which include organic solvents). In 0.1% SDS (Fig. 12), there appears to be one dimer form (D1) that is stable against heat denaturation in the absence of DTT while a second and perhaps more compact molecular form (D2) is consistently monomerized upon pre-incubation at 80 ^{o}C. However, in the presence of 2 mM DTT, the higher MW form appears to monomerize as well. Since there are no cysteines in the primary sequence, the stability in the absence of DTT suggests the involvement of stabilizing metals (bridging or otherwise) that are chelated away by DTT in the purified protein but which were not lost during purification, however, the crystal structure (Ch. 4) does not reveal additional metals bound beyond the catalytic metal center.

Here a serendipitous piece of evidence that was discarded earlier but recently revisited can be invoked. After the initial large-scale purification, purified native enzyme was inadvertently divided into two batches prior to storage, one with DTT added and another without. The DTT-free samples were completely consumed during optimization of the initial characterization experiments including i) the SDS-PAGE oligomeric studies (Fig. 12) and ii) an initial thermostability study which demonstrated remarkable retention of activity in the presence of Co^{2+} (Fig. 13). It was later realized in retrospect that a few of the stored enzyme aliquots lacked DTT, however, the significance was uncertain. It now seems that the stable dimer form observed in early SDS-PAGE (Fig. 12) may be the stable active species in the initial thermal stability studies (Fig.13). All subsequent studies presented in this thesis (including more recent thermostability experiments (Fig. 10) and crystallographic structural studies) involve the demetallated enzyme form which arises after the stabilizing metal(s) have been removed by DTT. That the removal of the

stabilizing metal(s) is irreversible (see lack of reconstitution below) suggests a kinetic trap, an issue that will revisited in steady-state kinetic studies (Ch. 4). That these stabilizing metal(s) appear to be important to dimeric integrity and activity suggests a role in protein folding and another possible reason for the low yield in refolding studies.

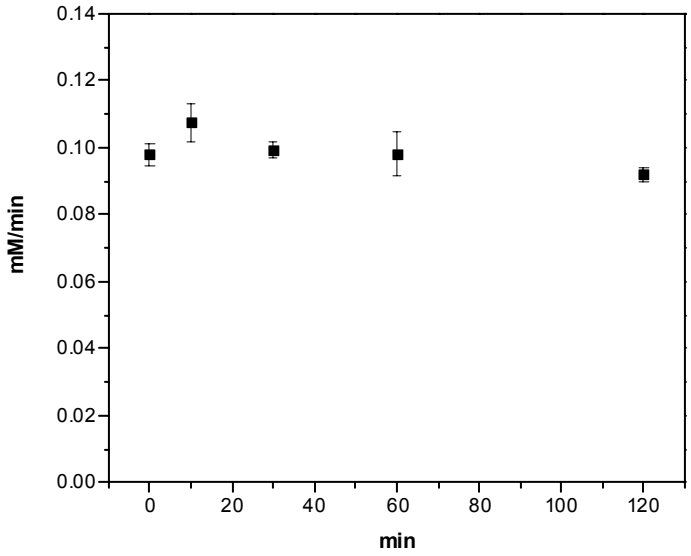


Fig. 13 Early thermostability experiment demonstrating remarkable retention of activity for 'DTT-free' PfuCP (1.4 nM samples of PfuCP were incubated at 80 ^{o}C for varying amounts of time in the presence of 400 µM $CoCl_2$ and 0.1 M KMes pH 6.5 before quenching on ice and subsequent addition of 2 mM ZAR to assay activity under standard conditions). Duplicate data with standard error shown.

The early thermostable activity data (Fig. 13), the dimeric stability in 0.1 % SDS (Fig. 12), together with the second-order dependence of recent inactivation data on the demetallated form (Fig. 10), suggest the dimer may be the active form at higher temperatures although it is possible that *in vivo* conditions (thiols, chelators, metal ion distribution) may favor the monomer. Oligomeric interactions are thought to contribute to thermostability in hyperthermophilic proteins (Sterner and Liebl 2001; Hess, Kruger et

al. 1995; Kohlhoff, Dahm et al. 1996; Legrain, Villeret et al. 1997), but the small contact surface between subunits (helix-helix contacts at the N-terminii) makes this a dubious assumption for PfuCP (Ch. 3). Steady-state experiments (Ch. 4) suggest instead that the dimeric interactions might have an allosteric role. Interestingly, the recombinant does not seem to have a stable dimer form (Fig. 6), perhaps reflecting the effects of the extra residues at the N-terminus retained from the enterokinase cleavage site.

Metal ion Specificity and Binding

Enzymes that use Co^{2+} as the native catalytic metal are few in number (Table 3), and the suggestion has been made that they are evolutionary holdovers from an era when Co^{2+} was more prevalent in the biosphere (Frausto da Silva and Williams 2001)[6]. Although replacing a native Zn^{2+} by Co^{2+} in many metalloproteases can increase the activity, the modern bioavailability of Zn^{2+}, its redox stability, and its strong ability to polarize and activate substrates have been taken as compelling evidence for the obligatory, native role of Zn^{2+} in metalloproteases, even when the metal is lost during purification and hence unknown.

Table 3 –Proteins and Enzymes with Presumed Native Cobalt

[6] It has also been suggested that the native metal for some presumed Co^{2+} enzymes is actually Fe^{2+} due to the bioavailability of the latter and other metal uptake and anaerobic experimental evidence (D'souza 1999).

Protein	Function	Comment
methionine aminopeptidase	removal of N-terminal dipeptide	dicobalt (bacteria to animals)
prolidase	cleavage of C-terminal to proline	dicobalt
glucose isomerase		Actinomycetes
methylomalonyl-CoA carboxy transferase		Co(II) + Zn(II) dimer (bacteria)
aldehyde decarbonylase		algae
lysine 2,3 amino-mutase		Zn(II), Fe_n/S_n as well (bacteria)
adenylate kinase		Zn(II) too (sulfur bacteria)
bromoperoxidase	bromination	V or Fe in other organisms (bacteria)
nitrile hydratase	-CN bond hydration	Co(III) (bacteria)
cobalt transporter COT1, GRR1, HOXN	transfer of Co(II)	also for Zn(II) and Ni(II) (bacteria and yeast)
cobalt porphyrin protein	possibly electron transfer	(bacteria)

Adapted from (Frausto da Silva and Williams 2001) p. 446

The purification of native PfuCP from *P. furiosus* cells included 1 mM EDTA as an inhibitor since significant loss of activity was observed in its absence, however the *apo*enzyme so prepared was only reactivated by Co^{2+} (Table 4). There is slight activity at the limit of detection when Mn^{2+} is added to the *apo*enzyme, but no other metal dichloride tested yields activity, nor is activity induced when glutathione or dithiothreitol is included to mimic the reducing environment of the cell interior. Interestingly enough, Mn^{2+} does restore full activity to crude cell lysate that has been pre-treated then washed free of 1 mM EDTA, which suggests either the presence of another Mn^{2+} activated protease which can cleave the same assay substrate (a trace contaminant of which is

causing the near negligible activity in purified PfuCP preparations) or Mn^{2+} has a different effect on the *in vivo* form of PfuCP versus its demetallated conformation. If Co^{2+} is in fact the native enzyme, mechanisms must exist to keep it from becoming trapped in inaccessible sulphide complexes.

No esterase activity was observed for any metal except for Co^{2+}, which showed a very slight activity at the limit of detection (<0.003 mM/min). Activity was not increased by raising either the ester concentration to 4 mM or the Co^{2+} metal concentration to 0.5, 1, and 15 mM.

Table 4 -Relative Peptidase and Esterase Activities for Different Metal-Substituted Isoforms

Metal[a]	PfuCP[b]	CPA[c] (peptidase)	CPA[d] (esterase)
Zn^{2+}	0	1, 1	1
Co^{2+}	1	2.23, 1.6	0.95
Mn^{2+}	~ 0 [b]	0.42, 0.08	0.35
Ni^{2+}	0	0, 1.06	0.87
Ca^{2+}	0		
Mg^{2+}	0		
Fe^{2+}	0		
Cu^{2+}	0	0, 0	0
Cd^{2+}	0	0, 0	1.5
Na^{+}	0		
Li^{+}	0		
K^{+}	0		
Pd^{2+}	0		
Sn^{2+}	0		
Hg^{2+}	0	0, 0	1.16
Pb^{2+}	0	-, 0	0.52

[a] The metal concentrations used in assays were 200 µM of the metal chloride (or nitrate and sulfate for Pb^{2+} and Mg^{2+}, respectively). The [ZAR] was fixed at 10 mM.
[b] The activity seen for PfuCP with Mn^{2+} is attributed to a trace amount of contaminating Mn^{2+} activated protease or different effects on the *in vivo* vs demetallated forms. The k_{cat} for CoPfuCP is 2 x 10^5 min^{-1} (Ch. 4)
[c] Values for CPA taken from (Davies, Riordan et al. 1968) and (Feinberg, Greenblatt et al. 1993) for Carbobenzoxy-glycyl-phenylalanine (ZGF). The k_{cat} for CoCPA is 12.3 x 10^3 min^{-1} from (Davies, Riordan et al. 1968)
[d] Values for CPA taken from (Feinberg, Greenblatt et al. 1993)
PfuCP shows no esterase activity against 10 mM Z-Ala-Arg(OMe) ie. ZAR(OMe) where [Me^{2+}] = 200 µM in 0.1 M KMes pH 6.5 where Me^{2+} is a divalent metal other than cobalt.

The inhibition properties of various metal ions was estimated by adding 200 µM of the divalent cation to standard assay reactions and the inhibition strength generally follows a decrease in ionic radius although there is one noticeable deviation from the

Irving-Williams series: Ni^{2+} inhibition < Fe^{2+} (perhaps due to facile oxidation of Fe(II) to the more exchange inert Fe(III)).

Table 5 -Residual Activity of PfuCP + 200 μM Co^{2+} in Presence of 200 μM Me^{2+}

Me^{2+}	Relative Activity
Co^{2+}	1
K^{+}	0.98
Li^{+}	0.79
Mn^{2+}	0.78
Ca^{2+}	0.77
Mg^{2+}	0.74
Na^{+}	0.74
Hg^{2+}	0.50
Pd^{2+}	0.30
Sn^{2+}	0.30
Ni^{2+}	0.24
Fe^{2+}	0.08
Pb^{2+}	0
Cd^{2+}	0
Cu^{2+}	0

The optical spectrum for CoPfuCP (Fig. 14) shows a slight increase in peak absorbance at 510 nm without a significant shift in the band envelope (the discrepancy in the lower lefthand shoulder is an artifact resulting from baseline correction during subtraction of the protein background). The K_d for Co^{2+} binding to enzyme derived from steady-state kinetic simulations (Ch. 4) is on the order of 10 μM, so it can be assumed

that PfuCP (1.5 mM) is saturated by Co^{2+} (10 mM)[7]. The difference in peak absorbance can then be used to calculate an absorption coefficient at 510 nm for CoPfuCP of ~ 7.9 M^{-1} cm^{-1} which is consistent with an octahedral Co^{2+} coordination (Rosenberg 1975), in keeping with the absence of a wavelength shift relative to free hexaaquo Co^{2+} (Fig. 14).[8] Tetrahedral Co^{2+} coordination would be characterized by an absorption with $\varepsilon > 300$ M^{-1} but at a higher wavelength (625 ± 50 nm) due to a smaller ligand field stabilization energy whereas pentacoordinate Co^{2+} exhibits an intermediate ε ~50-250 M^{-1} cm^{-1} and absorbances between 525 and 625 nm.

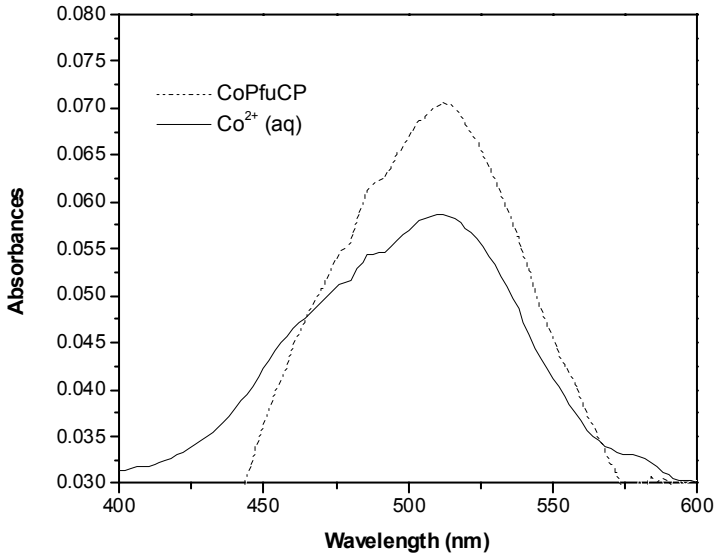


Fig. 14 Optical spectrum of 10 mM $CoCl_2$ and 1.5 mM CoPfuCP in 0.1 M KMes pH 6.5 (in the presence of 10 mM $CoCl_2$) showing a peak and difference absorbance at 510 nm typical of octahedral coordination (ε ~ 7.9 M^{-1} cm^{-1}). Deviations in the shoulders are artifacts from baseline correction.

[7] See Appendix C for details of calculation.

[8] Analogous difference EPR spectra were collected but the use of recombinant PfuCP due to a limited amount of purified native enzyme gave artifacts from progressive binding of Co^{2+} to the histidine tag. Optimization of digestions to remove the His-tag are currently underway.

Octahedral coordination for the catalytic metal in a Zn^{2+}-type metalloprotease is atypical and likely a direct result of the extreme lability of Co^{2+} in PfuCP. Interestingly enough, the crystal structure of the *apo*enzyme has a Mg^{2+} from crystallization buffers bound to the catalytic metal site in a distorted octahedral coordination sphere which includes 3 waters. Anaerobic and aerobic incubation of *apo*PfuCP (1.5 mM enzyme in 0.1 M KMes, pH 6.5, at room temperature) for up to 24 h in the presence of 1 mM $CoCl_2$ followed by removal of excess metal by dialysis (1 h ultrafiltration at 4 ^{o}C) does not incorporate Co^{2+} as evidenced by the lack of signal in absorption and EPR[9] spectra as well as ICP-MS. Addition of Co^{2+} to the *apo*enzyme at room temperature does not alter the far-UV CD or the intrinsic fluorescence (Fig. 11). Incubation of the *apo*enzyme (150 μM) with Co^{2+} (1 mM) and a substrate analog β-phenylpropionate (130 μM) across a wide pH range (from pH 4.9-8.8 in 0.3 increments) does not yield a stable complex as evidenced by optical spectroscopy. No optical absorbance was observed after heating of the enzyme aerobically at 80 ^{o}C (15 μM PfuCP) in the presence of 1 mM Co^{2+} suggesting that Co^{2+} in the catalytic site does not become oxidized. Exposure of the enzyme (1.5 mM + 10 mM Co^{2+}) to 1 mM H_2O_2 generated an unquantifiable amount of Co(III)PfuCP which appeared as a slight modulation of the noise from 450-560 nm. Structural origins for this loose binding are hinted at in the crystal structure (Ch. 3) and implications for regulation and catalysis will be discussed in chapters 4 and 5.

[9] See Appendix E for detailed methods.

Chapter 3

Crystal Structure and Sequence Analysis

Abstract

The structure of *apo*PfuCP was solved to 2.2 Å resolution through multiwavelength anomalous diffraction (MAD) methods and revealed a homodimer with predominantly helical structure and a deep active site groove in each subunit (Arndt, Hao et al. 2001). The structure is highly homologous with neurolysin (rat) and multiple alignment of a BLAST search revealed conserved residues that define a new family of carboxypeptidases, M32.

Introduction

Similarities between the crystal structures of various metalloproteases have suggested commonalities in the mechanism; however, biochemical experiments continue to reveal additional complexity due in part to subtle differences in the identity and disposition of critical active site residues (see Ch. 5). In the broadest sense, metalloproteases appear to deprotonate water which then attacks the substrate via its scissile peptide bond coordinated to the catalytic metal (Fig. 15). Coordination of the catalytic metal is nominally four or five depending on the denticity of a glutamate ligand (carboxyl) or alternately the presence of a conserved tyrosine ligand (Table 1) and the resulting geometries are most commonly described as distorted tetrahedral or trigonal bipyramid. The catalytic metal activates the substrate by polarizing the scissile carbonyl,

which is subsequently attacked by a water/hydroxide (aquo species) generating a tetrahedral transition state that collapses with protonation of the scissile amide nitrogen, releasing the C-terminal peptide or amino acid (Fig. 15). To date, the exact identity of the general base (B), the order and nature (sequential or concerted) of the deprotonation steps, the location of the attacking water (metal bound or not), and exact role of active site residues in stabilizing intermediates along the reaction coordinate remain incompletely determined for most metalloproteases (Ch. 5).

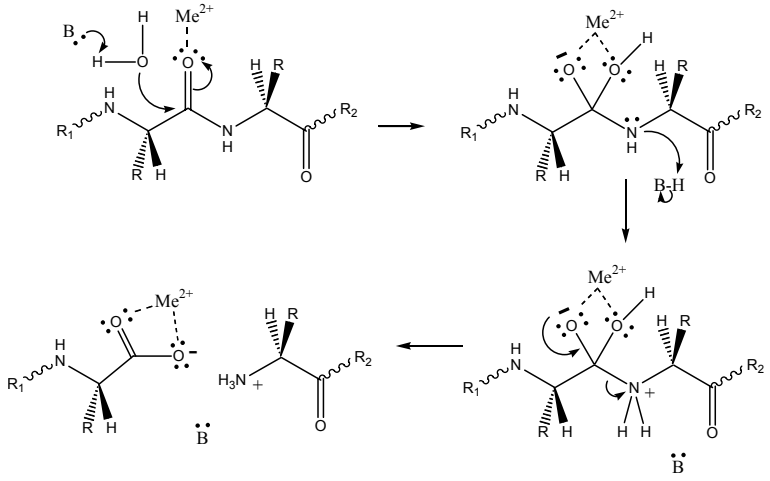


Fig. 15 Generalized hydrolysis mechanism for metalloproteases showing a general base or proton shuttle (B) and peptide bond where R_1 = H for aminopeptidases, R_2 = O for carboxypeptidases or R_1 and R_2 = additional amino acid residues for endopeptidases.

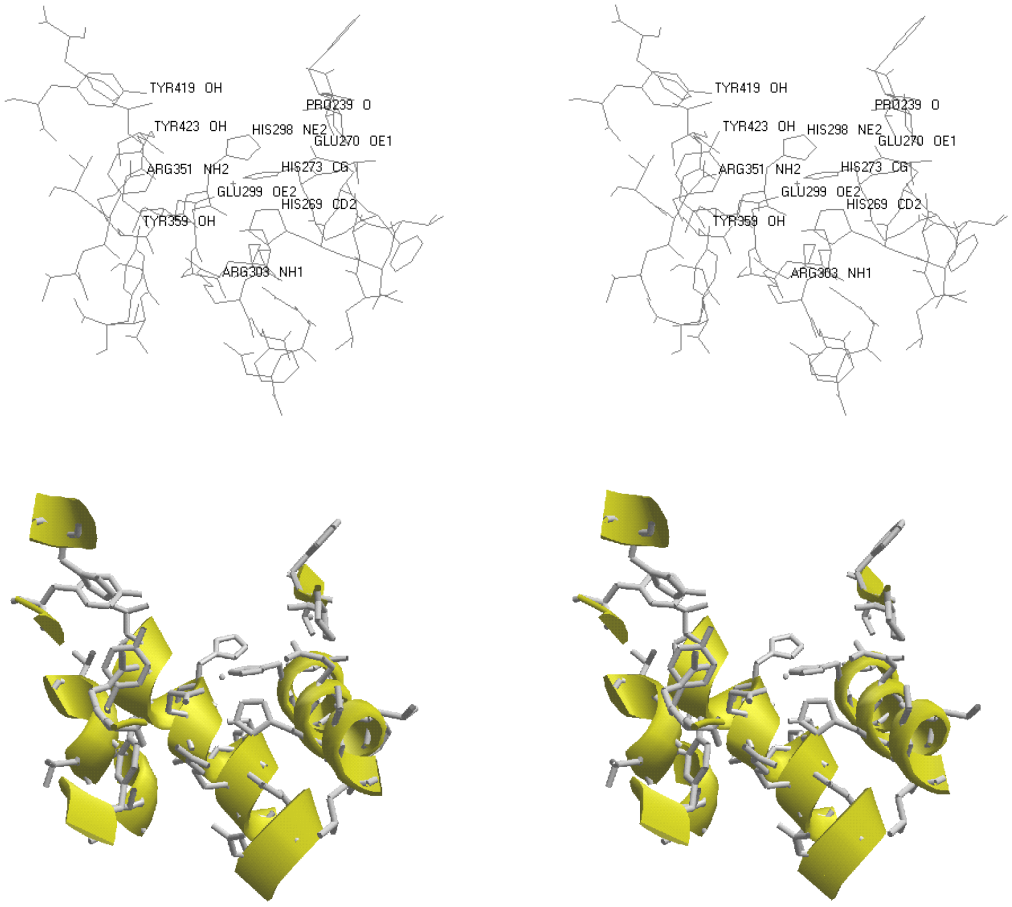


Fig. 16 Stereoview of the active site of PfuCP within 8 Å of the catalytic metal. The metal ligands arise from two consecutive α-helices packed against β-sheet to the top right (graphics generated with Swiss-PDB Viewer (Guex and Peitsch 1997)).

The active site of PfuCP is a close match to that of the archetypical endopeptidase thermolysin, and in PfuCP the metal ligands His 269 and 273 and Glu 299 arise from consecutive helices (in carboxypeptidase A (CPA), the corresponding metal ligands arise from β-sheet and loop segments (Rees, Lewis et al. 1983))[10]. Arg 351 is in a position to bind the C-terminal carboxyl and Glu 270 (Glu 143 in thermolysin) is poised to act as a critical counterion or as the general base in deprotonation of a nucleophilic water and perhaps reprotonation of the scissile amide prior to release of the terminal amino acid.

Other residues, conserved with respect to homologs of PfuCP (Fig. 20), are positioned to stabilize various intermediates along the reaction coordinate. Pro 239 positions the scissile amide to be protonated prior to amino acid release, His 238 can be modeled to hydrogen-bond to a putative tetrahedral intermediate, and Y423 can stabilize the C-terminal carboxyl and transition state (Arndt, Hao et al. 2001). In thermolysin and the related endopeptidase astacin respectively, H231and Y149 are in the same position as Y423 from PfuCP and have been suggested to stabilize the tetrahedral intermediate by binding to one of the coordinating oxygens; prolines in these two enzymes also mirror Pro 239 from PfuCP, and Asn 112 in thermolysin is in a similar location to His 238 from PfuCP (Holden, Tronrud et al. 1987; Grams, Dive et al. 1996). Residues in a position to form the S_1' specificity pocket for the C-terminal amino acid include Tyr 423 and Tyr 359 which would explain the specificity for basic, aromatic, and aliphatic C-terminii. Another interesting conserved feature akin to the Met-turn in endopeptidases appears to be a serine just below the metal site (Ser 302 PfuCP, Ser 506 in neurolysin (Brown, Madauss et al. 2001), and Ser 169 in thermolysin) that forms hydrogen bonds to the glutamate metal ligand in PfuCP and neurolysin.

PfuCP does not appear to have drastically fewer direct hydrogen bonds to its metal ligands (Glu 299 OE2-Ser 302 and His 269 ND1-Glu 306 OE1) compared with either thermolysin (His 142 ND2-Asp 170 OD1 and Tyr 157-Glu 166 OE2) or CPA (Glu 72 OE2-Ser 197, Asp 142 OD1-His 69 NE2, Arg 127 NH2-His 69 NE2)[11] nor does the

---

[10] Although the apoenzyme was crystallized, a Mg^{2+} appeared to bind in an octahedral coordination to the metal site with three additional water ligands.

[11] Coordinates for substrate free thermolysin and CPA are from the Research Collaboratory for Structural Biology (RCSB) files 1L3F and 1CPA5 and hydrogen bonds were calculated with Swiss-PDB Viewer (Guex 1997)

extended hydrogen bonding network appear to be less compared with either thermolysin or CPA. The number of residues within an 8 Å sphere of the catalytic metal is also very similar (PfuCP, 20; thermolysin, 20; CPA, 21) so the reasons for loose binding of the catalytic metal is not immediately obvious; however, it is possible that the specific low energy hydrogen bonds and hydrophobic contacts contribute to protein flexibility and hence metal lability in addition to the effects of a predominantly helical framework (which allows for more segmental mobility than a β-sheet core) and a solvent rich internal cavity (which increases the relative dielectric).

In overall topology, PfuCP is predominantly α-helical (63%) with a three-stranded β-sheet (3.2%) near the active site (Fig. 18). Two crystal forms confirm that the enzyme is dimeric with dimensions of 100 Å x 65 Å x 55 Å and a surface area of 34,800 $Å^2$; the distance of ~50 Å between the active sites (which face opposite directions) argues against concerted action of both active sites on a single substrate (Fig. 17). The interface between pairs of N-terminal helices (α1 and α2) on each subunit amounts to 4400 $Å^2$ of buried surface, with contacts being primarily hydrophobic (A22, I23, A26, V29, L30, G42, V48, A49, L53, F286) and in part electrostatic (salt bridges R19-E59 and D33-R250). The subunits are essentially identical (α-C RMSD of 0.473) with dimensions of 75 Å x 55 Å x 55Å and a deep groove 40 Å long and 30 Å deep that runs the length of the protein. One end is wide and exposed (Fig. 17) while the other is blocked by a helix 'cap' (α4), and modeling of a decapeptide into a subunit places the N-terminus of the substrate near the 'cap' and the C-terminus close to the active site which is about midway down the groove, suggesting that peptide substrates may feed into the enzyme N-terminally through the wide end. Similar to astacin and thermolysin (Holden, Tronrud et

al. 1987) (Grams, Dive et al. 1996), the 10-mer would bind antiparallel to the β-strand nearest to the metal (β2) via four hydrogen bonds.

Fig. 17 PfuCP dimer: each subunit faces in an opposite direction about a vertical two-fold rotation axis with a small contact area between pairs of helices from each subunit (α1 and α2)

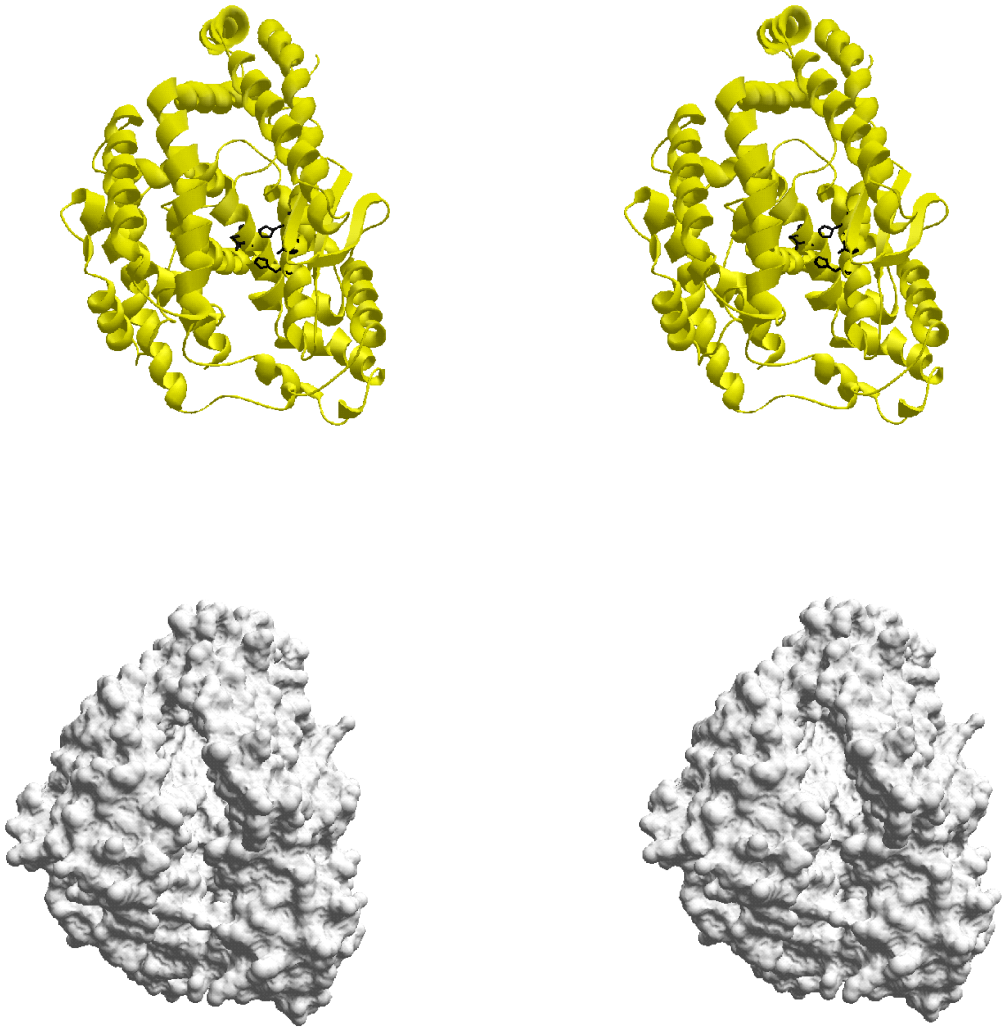


Fig. 18 PfuCP monomer: 'front' view into the active site groove showing the helical 'cap' blocking the upper end and the open bottom. The active site is located beside the β-sheet towards the center of the cleft (metal and ligands shown). Ribbons and molecular surface generated with Swiss PDB Viewer (Guex and Peitsch 1997).

The deep groove appears to allow size selectivity for substrates (Ch. 6) while simultaneously accomodating a varied range of peptide sequences, possibly in a manner similar to MHC I and other proteins that bind polypeptides extensively but non-specifically; such a manner of substrate binding has be implicated in processivity (Breyer and Matthews 2001). The large active site cleft also appears to have a functional role in

allostery (Ch. 4) and may also act as a ‘sticky’ trap to collect substrates which are low abundance in order to mitigate any limiting effects of mass transfer.

Sequence Analysis

Structural homology searches with DALI and FUGUE (Shi, Blundell et al. 2001) (Holm and Sander 1998) found one match with a zinc endopeptidase from *Rattus norvegicus*, neurolysin (Protein Data bank code 1lll) (Brown, Madauss et al. 2001). This enzyme has an established role in cleaving a 13-residue neuropeptide, neurotensin, between residues 10 and 11 and it can be superimposed on PfuCP with a rmsd of 2.1 Å for the 251 C_α atoms used in the superposition (Fig. 19) (Arndt, Hao et al. 2001).

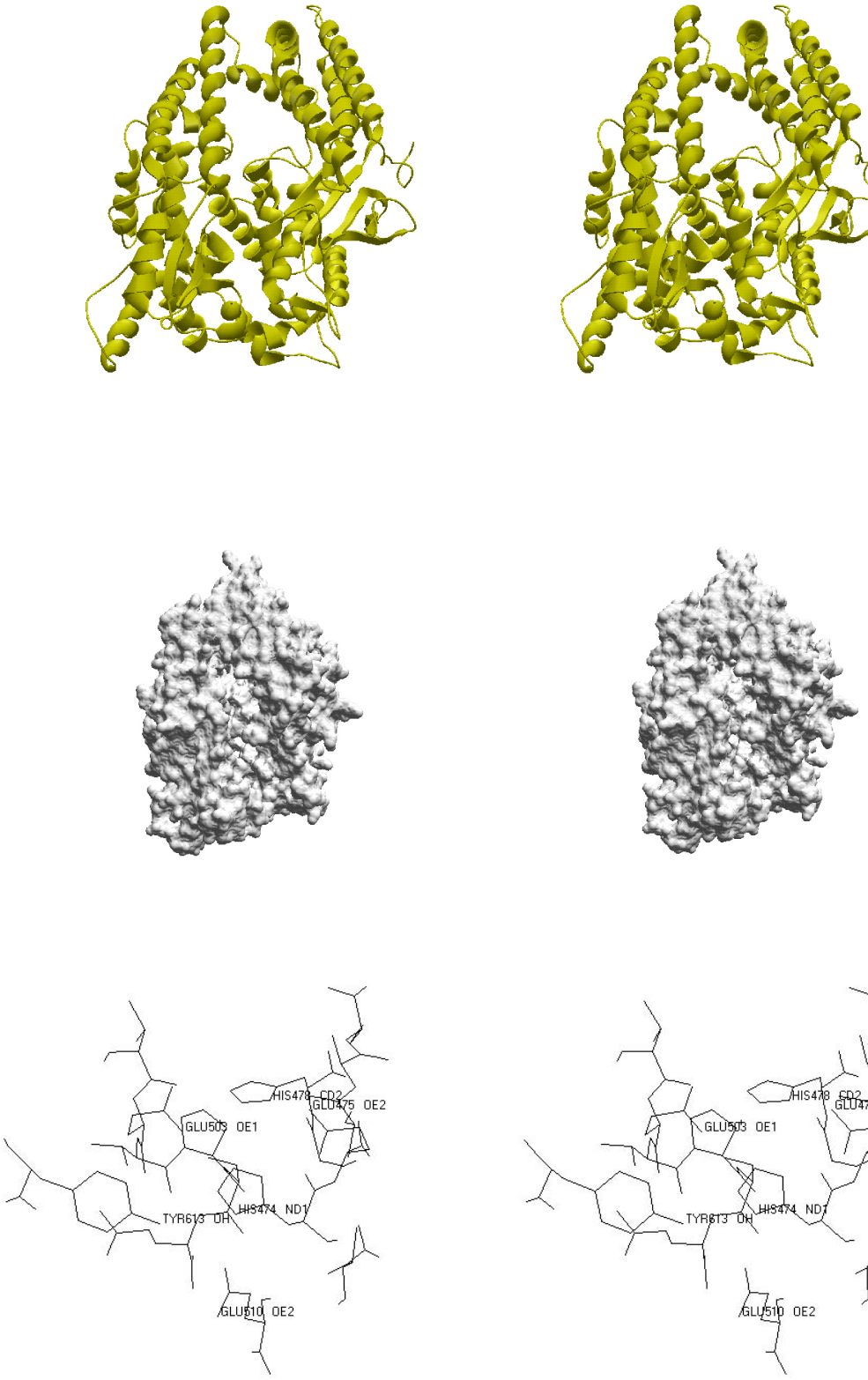
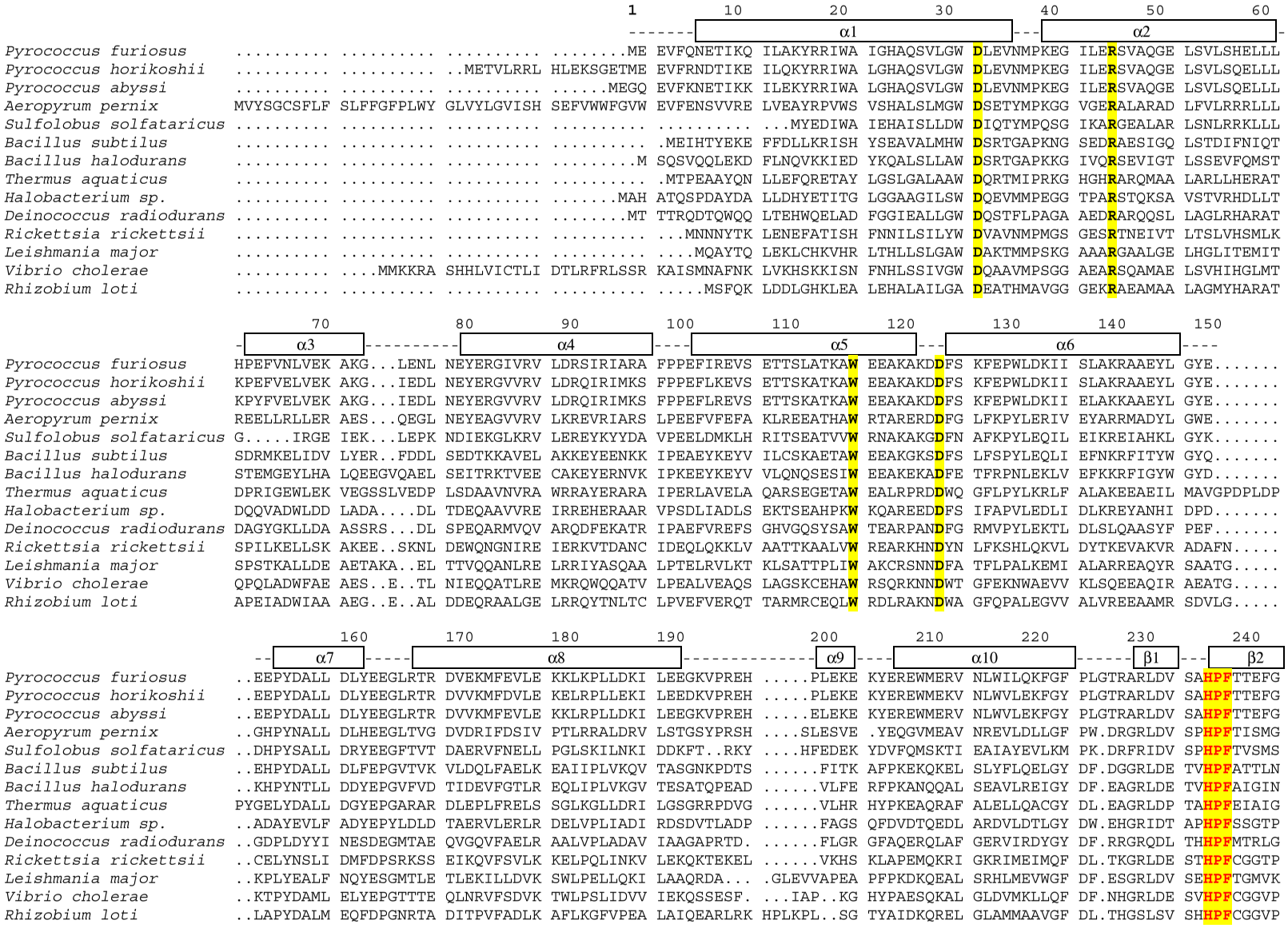
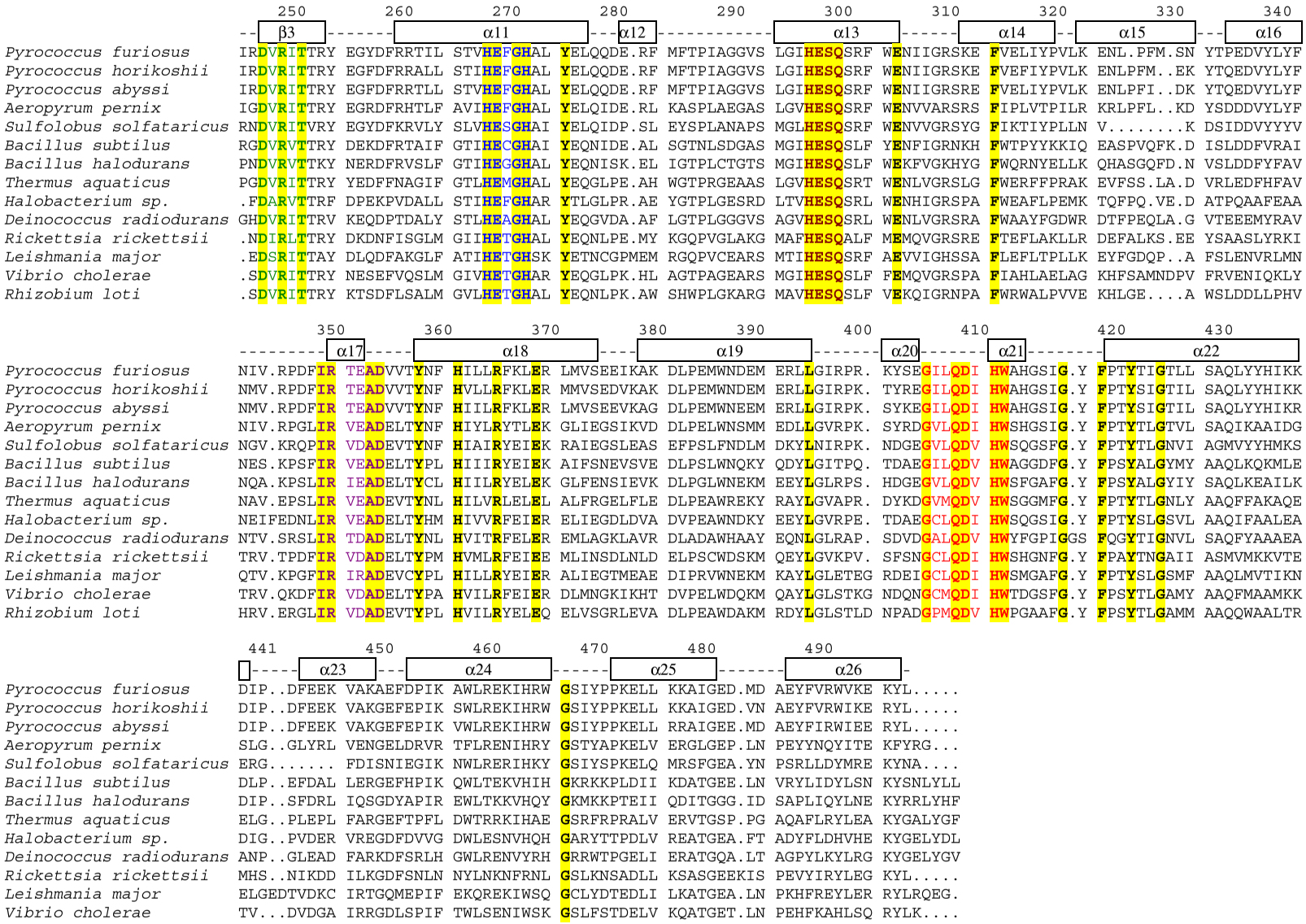


Fig. 19 Neurolysin in the same orientation as Figure 16 for PfuCP; more occluding structure is evident on the lower left of the substrate groove.

The larger size of neurolysin (681 residues vs 499 for PfuCP) accomodates additional structural elements, a third α-helix at the N-terminus, another sheet-rich subdomain that narrows the bottom end of the substrate groove, and two more β-strands near the active site. The additional loop structure in neurolysin is thought to contribute to non-specific substrate interactions whereas the reduced amount of coil in PfuCP may be related to maintaining thermal integrity. The overall similarities between the two proteins, however, suggests divergent evolution from a common ancestor despite the low amino acid identity between the two (<16%).

PfuCP lacks an HXXE$(X)_{123\text{-}132}$H motif characteristic of ‘classical’ metalloCPs but its homology with 8 other proteases defines a new M32 family of proteases which contain a conserved HEXXH motif more typical of endoproteases (Arndt, Hao et al. 2001).


1 10 20 30 40 50 60
α1
α2
Pyrococcus furiosus ....ME EVFQNETIKQ ILAKYRRIWA IGHAQSVLGW DLEVNMPKEG ILERSVAQGE LSVLSHELLL
Pyrococcus horikoshii ....METVLRRL HLEKSGETME EVFRNDTIKE ILQKYRRIWA LGHAQSVLGW DLEVNMPKEG ILERSVAQGE LSVLSQELLL
Pyrococcus abyssi ....MEGQ EVFKNETIKK ILEKYRRIWA LGHAQSVLGW DLEVNMPKEG ILERSVAQGE LSVLSQELLL
Aeropyrum pernix MVYSGCSFLF SLFFGFPLWY GLVYLGVISH SEFVWWFGVW EVFENSVVRE LVEAYRPVWS VSHALSLMGW DSETYMPKGG VGERALARAD LFVLRRRLLL
Sulfolobus solfataricus ....MYEDIWA IEHAISLLDW DIQTYMPQSG IKARGEALAR LSNLRRKLLL
Bacillus subtilus .... .MEIHTYEKE FFDLLKRISH YSEAVALMHW DSRTGAPKNG SEDRAESIGQ LSTDIFNIQT
Bacillus halodurans ....M SQSVQQLEKD FLNQVKKIED YKQALSLLAW DSRTGAPKKG IVQRSEVIGT LSSEVFQMST
Thermus aquaticus .... .MTPEAAYQN LLEFQRETAY LGSLGALAAW DQRTMIPRKG HGHRARQMAA LARLLHERAT
Halobacterium sp. ....MAH ATQSPDAYDA LLDHYETITG LGGAAGILSW DQEVMMPEGG TPARSTQKSA VSTVRHDLLT
Deinococcus radiodurans ....MT TTRQDTQWQQ LTEHWQELAD FGGIEALLGW DQSTFLPAGA AEDRARQQSL LAGLRHARAT
Rickettsia rickettsii ....MNNNYTK LENEFATISH FNNILSILYW DVAVNMPMGS GESRTNEIVT LTSLVHSMLK
Leishmania major ....MQAYTQ LEKLCHKVHR LTHLLSLGAW DAKTMMPSKG AAARGAALGE LHGLITEMIT
Vibrio cholerae ....MMKKRA SHHLVICTLI DTLRFRLSSR KAISMNAFNK LVKHSKKISN FNHLSSIVGW DQAAVMPSGG AEARSQAMAE LSVHIHGLMT
Rhizobium loti ....MSFQK LDDLGHKLEA LEHALAILGA DEATHMAVGG GEKRAEAMAA LAGMYHARAT
70 80 90 100 110 120 130 140 150
α3
α4
α5
α6
Pyrococcus furiosus HPEFVNLVEK AKG...LENL NEYERGIVRV LDRSIRIARA FPPEFIREVS ETTSLATKAW EEAKAKDDFS KFEPWLDKII SLAKRAAEYL GYE.......
Pyrococcus horikoshii KPEFVELVEK AKG...IEDL NEYERGVVRV LDRQIRIMKS FPPEFLKEVS ETTSKATKAW EEAKAKDDFS KFEPWLDKII SLAKRAAEYL GYE.......
Pyrococcus abyssi KPYFVELVEK AKG...IEDL NEYERGVVRV LDRQIRIMKS FPPEFLREVS ETTSKATKAW EEAKAKDDFS KFEPWLDKII ELAKKAAEYL GYE.......
Aeropyrum pernix REELLRLLER AES...QEGL NEYEAGVVRV LKREVRIARS LPEEFVFEFA KLREEATHAW RTARERDDFG LFKPYLERIV EYARRMADYL GWE.......
Sulfolobus solfataricus G.....IRGE IEK...LEPK NDIEKGLKRV LEREYKYYDA VPEELDMKLH RITSEATVVW RNAKAKGDFN AFKPYLEQIL EIKREIAHKL GYK.......
Bacillus subtilus SDRMKELIDV LYER..FDDL SEDTKKAVEL AKKEYEENKK IPEAEYKEYV ILCSKAETAW EEAKGKSDFS LFSPYLEQLI EFNKRFITYW GYQ.......
Bacillus halodurans STEMGEYLHA LQEEGVQAEL SEITRKTVEE CAKEYERNVK IPKEEYKEYV VLQNQSESIW EEAKEKADFE TFRPNLEKLV EFKKRFIGYW GYD.......
Thermus aquaticus DPRIGEWLEK VEGSSLVEDP LSDAAVNVRA WRRAYERARA IPERLAVELA QARSEGETAW EALRPRDDWQ GFLPYLKRLF ALAKEEAEIL MAVGPDPLDP
Halobacterium sp. DQQVADWLDD LADA....DL TDEQAAVVRE IRREHERAAR VPSDLIADLS EKTSEAHPKW KQAREEDDFS IFAPVLEDLI DLKREYANHI DPD.......
Deinococcus radiodurans DAGYGKLLDA ASSRS...DL SPEQARMVQV ARQDFEKATR IPAEFVREFS GHVGQSYSAW TEARPANDFG RMVPYLEKTL DLSLQAASYF PEF.......
Rickettsia rickettsii SPILKELLSK AKEE..SKNL DEWQNGNIRE IERKVTDANC IDEQLQKKLV AATTKAALVW REARKHNDYN LFKSHLQKVL DYTKEVAKVR ADAFN.....
Leishmania major SPSTKALLDE AETAKA..EL TTVQQANLRE LRRIYASQAA LPTELRVLKT KLSATTPLIW AKCRSNNDFA TFLPALKEMI ALARREAQYR SAATG.....
Vibrio cholerae QPQLADWFAE AES..E..TL NIEQQATLRE MKRQWQQATV LPEALVEAQS LAGSKCEHAW RSQRKNNDWT GFEKNWAEVV KLSQEEAQIR AEATG.....
Rhizobium loti APEIADWIAA AEG..E..AL DDEQRAALGE LRRQYTNLTC LPVEFVERQT TARMRCEQLW RDLRAKNDWA GFQPALEGVV ALVREEAAMR SDVLG.....
160 170 180 190 200 210 220 230 240
α7
α8
α9
α10
β1
β2
Pyrococcus furiosus ..EEPYDALL DLYEEGLRTR DVEKMFEVLE KKLKPLLDKI LEEGKVPREH .....PLEKE KYEREWMERV NLWILQKFGF PLGTRARLDV SAHPFTTEFG
Pyrococcus horikoshii ..EEPYDALL DLYEEGLRTR DVVKMFDVLE KELRPLLEKI LEEDKVPREH .....PLEKE KYEREWMEKV NLWVLEKFGY PLGTRARLDV SAHPFTTEFG
Pyrococcus abyssi ..EEPYDALL DLYEEGLRTR DVVKMFEVLE KKLRPLLDKI LEEGKVPREH .....ELEKE KYEREWMERV NLWVLEKFGY PLGTRARLDV SAHPFTTEFG
Aeropyrum pernix ..GHPYNALL DLHEEGLTVG DVDRIFDSIV PTLRRALDRV LSTGSYPRSH ....SLESVE .YEQGVMEAV NREVLDLLGF PW.DRGRLDV SPHPFTISMG
Sulfolobus solfataricus ..DHPYSALL DRYEEGFTVT DAERVFNELL PGLSKILNKI DDKFT..RKY ....HFEDEK YDVFQMSKTI EAIAYEVLKM PK.DRFRIDV SPHPFTVSMS
Bacillus subtilus ..EHPYDALL DLFEPGVTVK VLDQLFAELK EAIIPLVKQV TASGNKPDTS ......FITK AFPKEKQKEL SLYFLQELGY DF.DGGRLDE TVHPFATTLN
Bacillus halodurans ..KHPYNTLL DDYEPGVFVD TIDEVFGTLR EQLIPLVKGV TESATQPEAD ......VLFE RFPKANQQAL SEAVLREIGY DF.EAGRLDE TVHPFAIGIN
Thermus aquaticus PYGELYDALL DGYEPGARAR DLEPLFRELS SGLKGLLDRI LGSGRRPDVG ......VLHR HYPKEAQRAF ALELLQACGY DL.EAGRLDP TAHPFEIAIG
Halobacterium sp. ..ADAYEVLF ADYEPYLDLD TAERVLERLR DELVPLIADI RDSDVTLADP ......FAGS QFDVDTQEDL ARDVLDTLGY DW.EHGRIDT APHPFSSGTP
Deinococcus radiodurans ..GDPLDYYI NESDEGMTAE QVGQVFAELR AALVPLADAV IAAGAPRTD. ......FLGR GFAQERQLAF GERVIRDYGY DF.RRGRQDL THHPFMTRLG
Rickettsia rickettsii ..CELYNSLI DMFDPSRKSS EIKQVFSVLK KELPQLINKV LEKQKTEKEL ......VKHS KLAPEMQKRI GKRIMEIMQF DL.TKGRLDE STHPFCGGTP
Leishmania major ..KPLYEALF NQYESGMTLE TLEKILLDVK SWLPELLQKI LAAQRDA... .GLEVVAPEA PFPKDKQEAL SRHLMEVWGF DF.ESGRLDV SEHPFTGMVK
Vibrio cholerae ..KTPYDAML ELYEPGTTTE QLNRVFSDVK TWLPSLIDVV IEKQSSESF. ...IAP..KG HYPAESQKAL GLDVMKLLQF DF.NHGRLDE SVHPFCGGVP
Rhizobium loti ..LAPYDALM EQFDPGNRTA DITPVFADLK AFLKGFVPEA LAIQEARLRK HPLKPL..SG TYAIDKQREL GLAMMAAVGF DL.THGSLSV SHHPFCGGVP


250 260 270 280 290 300 310 320 330 340
β3 α11 α12 α13 α14 α15 α16
Pyrococcus furiosus IRDVRITTRY EGYDFRRTIL STVHEFGHAL YELQQDE.RF MFTPIAGGVS LGIHESQSRF WENIIGRSKE FVELIYPVLK ENL.PFM.SN YTPEDVYLYF
Pyrococcus horikoshii IRDVRITTRY EGFDFRRALL STIHEFGHAL YELQQDE.RF MFTPIAGGVS LGIHESQSRF WENIIGRSKE FVEFIYPVLK ENLPFM..EK YTQEDVYLYF
Pyrococcus abyssi IRDVRITTRY EGFDFRRTLL STVHEFGHAL YELQQDE.RF MFTPIAGGVS LGIHESQSRF WENIIGRSRE FVELIYPVLK ENLPFI..DK YTQEDVYLYF
Aeropyrum pernix IGDVRITTRY EGRDFRHTLF AVIHEFGHAL YELQIDE.RL KASPLAEGAS LGVHESQSRF WENVVARSRS FIPLVTPILR KRLPFL..KD YSDDDVYLYF
Sulfolobus solfataricus RNDVRITVRY EGYDFKRVLY SLVHESGHAI YELQIDP.SL EYSPLANAPS MGLHESQSRF WENVVGRSYG FIKTIYPLLN V........K DSIDDVYYYV
Bacillus subtilus RGDVRVTTRY DEKDFRTAIF GTIHECGHAI YEQNIDE.AL SGTNLSDGAS MGIHESQSLF YENFIGRNKH FWTPYYKKIQ EASPVQFK.D ISLDDFVRAI
Bacillus halodurans PNDVRVTTKY NERDFRVSLF GTIHEGGHAL YEQNISK.EL IGTPLCTGTS MGIHESQSLF WEKFVGKHYG FWQRNYELLK QHASGQFD.N VSLDDFYFAV
Thermus aquaticus PGDVRITTRY YEDFFNAGIF GTLHEMGHAL YEQGLPE.AH WGTPRGEAAS LGVHESQSRT WENLVGRSLG FWERFFPRAK EVFSS.LA.D VRLEDFHFAV
Halobacterium sp. .FDARVTTRF DPEKPVDALL STIHEFGHAR YTLGLPR.AE YGTPLGESRD LTVHESQSRL WENHIGRSPA FWEAFLPEMK TQFPQ.VE.D ATPQAAFEAA
Deinococcus radiodurans GHDVRITTRV KEQDPTDALY STLHEAGHAL YEQGVDA.AF LGTPLGGGVS AGVHESQSRL WENLVGRSRA FWAAYFGDWR DTFPEQLA.G VTEEEMYRAV
Rickettsia rickettsii .NDIRLTTRY DKDNFISGLM GIIHETGHAL YEQNLPE.MY KGQPVGLAKG MAFHESQALF MEMQVGRSRE FTEFLAKLLR DEFALKS.EE YSAASLYRKI
Leishmania major .EDSRITTAY DLQDFAKGLF ATIHETGHSK YETNCGPMEM RGQPVCEARS MTIHESQSRF AEVVIGHSSA FLEFLTPLLK EYFGDQP..A FSLENVRLMN
Vibrio cholerae .SDVRITTRY NESEFVQSLM GIVHETGHAR YEQGLPK.HL AGTPAGEARS MGIHESQSLF FEMQVGRSPA FIAHLAELAG KHFSAMNDPV FRVENIQKLY
Rhizobium loti .SDVRITTRY KTSDFLSALM GVLHETGHAL YEQNLPK.AW SHWPLGKARG MAVHESQSLF VEKQIGRNPA FWRWALPVVE KHLGE....A WSLDDLLPHV
350 360 370 380 390 400 410 420 430
α17 α18 α19 α20 α21 α22
Pyrococcus furiosus NIV.RPDFIR TEADVVTYNF HILLRFKLER LMVSEEIKAK DLPEMWNDEM ERLLGIRPR. KYSEGILQDI HWAHGSIG.Y FPTYTIGTLL SAQLYYHIKK
Pyrococcus horikoshii NMV.RPDFIR TEADVVTYNF HVLLRFKLER MMVSEDVKAK DLPEMWNDEM ERLLGIRPK. TYREGILQDI HWAHGSIG.Y FPTYSIGTIL SAQLYYHIKK
Pyrococcus abyssi NMV.RPDFIR TEADVVTYNF HIILRFKLER LMVSEEVKAG DLPEMWNEEM ERLLGIRPK. SYKEGILQDI HWAHGSIG.Y FPTYSIGTIL SAQLYYHIKR
Aeropyrum pernix NIV.RPGLIR VEADELTYNF HIYLRYTLEK GLIEGSIKVD DLPELWNSMM EDLLGVRPK. SYRDGVLQDI HWSHGSIG.Y FPTYTLGTVL SAQIKAAIDG
Sulfolobus solfataricus NGV.KRQPIR VDADEVTYNF HIAIRYEIEK RAIEGSLEAS EFPSLFNDLM DKYLNIRPK. NDGEGVLQDV HWSQGSFG.Y FPTYTLGNVI AGMVYYHMKS
Bacillus subtilus NES.KPSFIR VEADELTYPL HIIIRYEIEK AIFSNEVSVE DLPSLWNQKY QDYLGITPQ. TDAEGILQDV HWAGGDFG.Y FPSYALGYMY AAQLKQKMLE
Bacillus halodurans NQA.KPSLIR IEADELTYCL HIILRYELEK GLFENSIEVK DLPGLWNEKM EEYLGLRPS. HDGEGVLQDV HWSFGAFG.Y FPSYALGYIY SAQLKEAILK
Thermus aquaticus NAV.EPSLIR VEADEVTYNL HILVRLELEL ALFRGELFLE DLPEAWREKY RAYLGVAPR. DYKDGVMQDV HWSGGMFG.Y FPTYTLGNLY AAQFFAKAQE
Halobacterium sp. NEIFEDNLIR VEADELTYHM HIVVRFEIER ELIEGDLDVA DVPEAWNDKY EEYLGVRPE. TDAEGCLQDI HWSQGSIG.Y FPTYSLGSVL AAQIFAALEA
Deinococcus radiodurans NTV.SRSLIR TDADELTYNL HVITRFELER EMLAGKLAVR DLADAWHAAY EQNLGLRAP. SDVDGALQDV HWYFGPIGGS FQGYTIGNVL SAQFYAAAEA
Rickettsia rickettsii TRV.TPDFIR VDADELTYPM HVMLRFEIEE MLINSDLNLD ELPSCWDSKM QEYLGVKPV. SFSNGCLQDI HWSHGNFG.Y FPAYTNGAII ASMVMKKVTE
Leishmania major QTV.KPGFIR IRADEVCYPL HILLRYEIER ALIEGTMEAE DIPRVWNEKM KAYLGLETEG RDEIGCLQDI HWSMGAFG.Y FPTYSLGSMF AAQLMVTIKN
Vibrio cholerae TRV.QKDFIR VDADELTYPA HVILRFEIER DLMNGKIKHT DVPELWDQKM QAYLGLSTKG NDQNGCMQDI HWTDGSFG.Y FPSYTLGAMY AAQFMAAMKK
Rhizobium loti HRV.ERGLIR VDADEVTYPL HVILRYELEQ ELVSGRLEVA DLPEAWDAKM RDYLGLSTLD NPADGPMQDV HWPGAAFG.Y FPSYTLGAMM AAQQWAALTR
441 450 460 470 480 490
α23 α24 α25 α26
Pyrococcus furiosus DIP..DFEEK VAKAEFDPIK AWLREKIHRW GSIYPPKELL KKAIGED.MD AEYFVRWVKE KYL.....
Pyrococcus horikoshii DIP..DFEEK VAKGEFEPIK SWLREKIHRW GSIYPPKELL KKAIGED.VN AEYFVRWIKE RYL.....
Pyrococcus abyssi DIP..DFEEK VAKGEFEPIK VWLREKIHRW GSIYPPKELL RRAIGEE.MD AEYFIRWIEE RYL.....
Aeropyrum pernix SLG..GLYRL VENGELDRVR TFLRENIHRY GSTYAPKELV ERGLGEP.LN PEYYNQYITE KFYRG...
Sulfolobus solfataricus ERG..... FDISNIEGIK NWLRERIHKY GSIYSPKELQ MRSFGEA.YN PSRLLDYMRE KYNA....
Bacillus subtilus DLP..EFDAL LERGEFHPIK QWLTEKVHIH GKRKKPLDII KDATGEE.LN VRYLIDYLSN KYSNLYLL
Bacillus halodurans DIP..SFDRL IQSGDYAPIR EWLTKKVHQY GKMKKPTEII QDITGGG.ID SAPLIQYLNE KYRRLYHF
Thermus aquaticus ELG..PLEPL FARGEFTPFL DWTRRKIHAE GSRFRPRALV ERVTGSP.PG AQAFLRYLEA KYGALYGF
Halobacterium sp. DIG..PVDER VREGDFDVVG DWLESNVHQH GARYTTPDLV REATGEA.FT ADYFLDHVHE KYGELYDL
Deinococcus radiodurans ANP..GLEAD FARKDFSRLH GWLRENVYRH GRRWTPGELI ERATGQA.LT AGPYLKYLRG KYGELYGV
Rickettsia rickettsii MHS..NIKDD ILKGDFSNLN NYLNKNFRNL GSLKNSADLL KSASGEEKIS PEVYIRYLEG KYL.....
Leishmania major ELGEDTVDKC IRTGQMEPIF EKQREKIWSQ GCLYDTEDLI LKATGEA.LN PKHFREYLER RYLRQEG.
Vibrio cholerae TV...DVDGA IRRGDLSPIF TWLSENIWSK GSLFSTDELV KQATGET.LN PEHFKAHLSQ RYLK....

Rhizobium loti EHP..TADDD LAKGNFAAIN GWRREKIWSQ GSRWSTPNLL ERATGEK.LN AVHFVNHLKQ RYGG....

Fig. 20 (previous two pages) Sequence alignment of the M32 family of metalloproteases with ClustalW (Thompson, Higgins et al. 1994). The numbering and secondary structure above the alignment corresponds to PfuCP. Residues conserved between all the homologs are in bold and highlighted. Conserved motifs appear at 238-240 (HPF), DXRXT (248-252), HESQ (298-301), IRXXAD (350-355) and GXXQDXHW (405-412). Accession numbers with percent identity to PfuCP are as follows: B. subtilus, p50848, 36%; T. aquaticus, p42663, 37%; Halobacterium sp., q9mb7, 34%; D. radiodurans, q9rrr3, 33%; R. rickettsii, cac33677, 30%; L.major, cac37108, 33%; V. cholerae, q9ks46, 31%; R. loti, bab50714, 32%.

Chapter 4

Steady-State Kinetics (Activity) and Simulations

Abstract

Although preliminary experiments at saturating concentrations of Co^{2+} and substrate (Z-Ala-Arg) were consistent with simple Michaelis-Menten kinetics (Cheng, Ramakrishnan et al. 1999), varying the amounts of substrate and metal over a non-saturating concentration range revealed complex, multiphasic behavior. As the enzyme has not proven amenable to routine biophysical study due to weak association of metal and reaction coordinate analogs, simulations assuming rapid-equilibrium conditions were fit to metal activity profiles (activity vs. [Co^{2+}]) and substrate activity profiles (activity vs. [ZAR]) to obtain binding and kinetic constants for the non-hyperbolic behavior. Allostery between multiple bound metals and substrates is apparent and it is revealed that the activity profiles arise from an ensemble of active species. An absolute requirement for several substrates to be bound before full activity suggests a functional role for the large active site groove in terms of extensive, obligatory substrate-enzyme contacts. The lack of activity in the presence of Zn^{2+} is explained as tight, uncompetitive inhibition by multiple Zn^{2+}. The simple elegance of the kinetic approach belies the power of simulations to unravel the complex behavior of catalytic systems.

Since the crystal structure of PfuCP (Arndt, Hao et al. 2001) reveals the archetypical ligand constellation of mono-zinc metalloproteases (consensus metal binding motif HEXXH), the mechanism is assumed to be metal-assisted peptide hydrolysis that is sequential in nature, rather than ping-pong as typified by serine and cysteine proteases wherein a covalent 'acyl-enzyme' intermediate is generated.

The key issues that steady-state kinetic analysis of activity might then be able to address include the following:

1) whether there is a rapid pre-equilibrium such that chemical steps are rate-limiting,
2) the binding affinities of ligands, allosteric interactions, and other kinetic constants,
3) whether the enzyme behaves as a dimer or a monomer,
4) the nature of the apparent lack of zinc activation or zinc inhibition,
5) the effects of varying the temperature,
6) the effects of varying the pH (identification of critical ionizing amino acid residues).

As a first approach, the experimental data was fit to a rapid equilibrium scheme since it is not known *a priori* which binding steps might have slow rates similar to the chemical steps. This also allowed for a smaller number of parameters to be fit than for a full steady-state analysis and hence easier interpretation of allosteric effects on ligand binding and rate.

There is much evidence to suggest that the dimer form is stable: SDS-PAGE (without sample heating), crystallography, gel filtration, small-angle dynamic light scattering, MALDI-MS have all indicated a dimeric form at room temperature. However, it is unknown whether the enzyme behaves as an active dimer or monomer at higher temperatures (or a mixture of the two) since thermal inactivation has a second-order dependence which suggest intersubunit interactions yet the crystal structure indicates a tenuous interface between subunits and SDS-PAGE with pre-heating of the sample indicates a mixture of monomer and dimer depending on the conditions. If the enzyme is dimeric at high temperatures, then any allostery between subunits might distinguish the kinetics of the monomer from that of the dimer.

To this end, models for rapid equilibrium schemes of monomeric and dimeric PfuCP were assumed, general velocity equations derived, and simulations generated to compare with the observed kinetics. A brief overview of the experimental data follows, which serves as a guide for the development of the model.

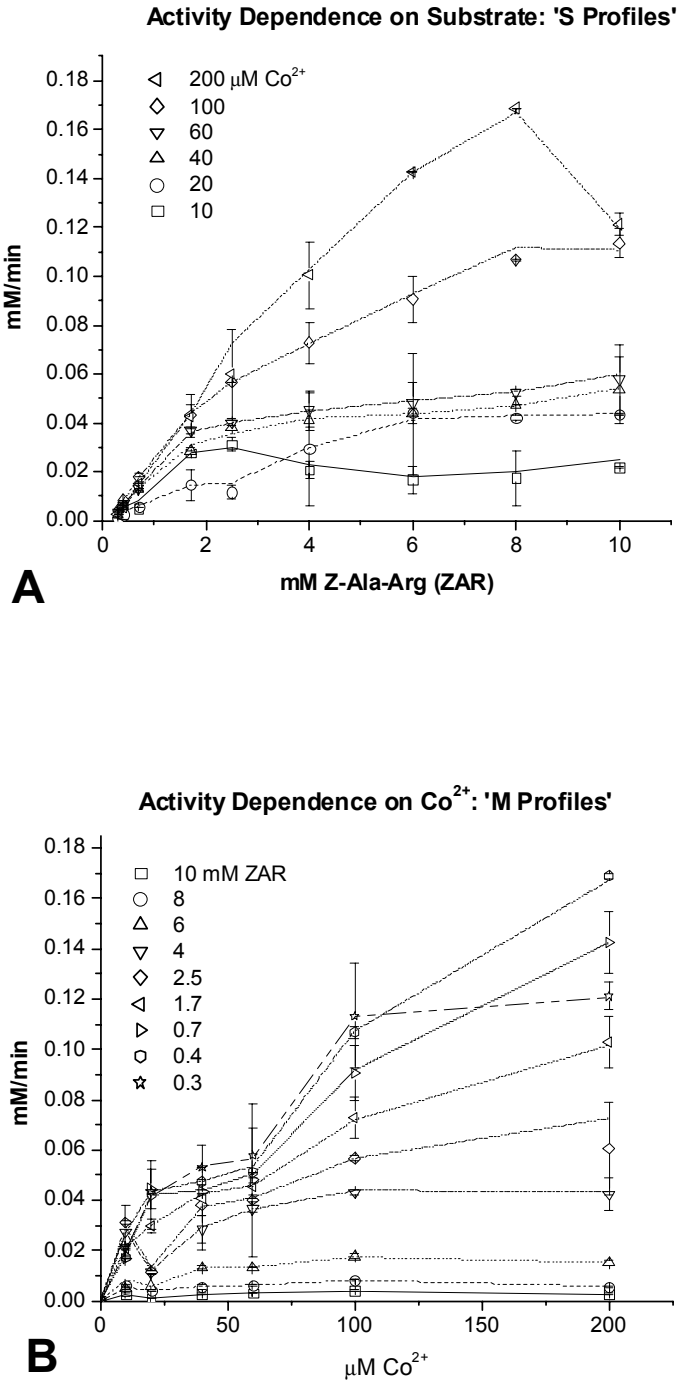


Fig. 21 Dependence of initial velocity on the substrate (A) or Co^{2+} concentration (B) from a single set of experimental measurements (0.1 M KMes pH 6.5, [E] = 6.7 x 10^{-7} mM, 80 ^{o}C, Ionic strength = 0.1 M with less than 10% variation in ionicity over course of titrations). Duplicate data points appear with error while simulations from Equation 3 appear as lines.

S Profiles (Activity vs. [Z-Ala-Arg])

The S (substrate) activity profiles in Figure 21 A show the variation in initial velocity measurements as substrate (ZAR) is titrated in against a constant background of Co^{2+}. As the profiles move upwards, the background metal concentration is increasing from 10-200 μM. In the bottommost trace with a constant background of 10 μM Co^{2+}, the initial increase in activity is apparently followed by substrate inhibition and then a slight substrate reactivation or grow in of a new active species at high [ZAR]. This immediately suggests a noncompetitive or uncompetitive inhibition mode since competitive inhibition would depress the initial rise and flatten the overall profile but would not cause a downturn. Moving up to 20 μM Co^{2+}, the initial hump seen previously at 10 μM Co^{2+} has been inhibited and diminished by the greater background of metal while the rising second phase at higher substrate concentration could be reactivation by substrate or the appearance of a new enzyme species. At 40 and 60 μM Co^{2+}, it appears that the initial hump which was inhibited at 20 μM Co^{2+} has been either reactivated by higher metal or again a new active species grows in. Finally, the profiles at 100 and 200 μM appear to be either an aggregate of contributions from the species already present or from more species that grow in at higher concentration of metal.

M Profiles (Activity vs. Co^{2+})

The M (metal) activity profiles in Figure 21 B show the variation in initial velocity measurements as Co^{2+} is titrated in against a constant background of substrate[12]. Derived from the same data set as the S profiles, recasting the data in this way is equivalent to taking vertical slices of the S profiles which yields a clearer picture of the activity dependence on metal concentration. Moving up the M profiles from a background substrate concentration of 0.3 to 2.5 mM ZAR, the traces have a similar shape and reveal an initial peak of activity at 10 μM Co^{2+}, followed by a sharp non-competitive/uncompetitive inhibition (abrupt loss of active species) at 20 μM Co^{2+} then two more phases of activity growth from 40-100 μM Co^{2+} that could either be reactivation of the inhibited species or the appearance of one or two new active species. There is a slight but real decrease in activity for these profiles (0.3-2.5 mM ZAR) beyond 100 μM Co^{2+} which suggests uncompetitive metal inhibition.

The main feature in the activity profiles moving from 4 to 10 mM ZAR is either a release of the metal inhibition which causes the initial dip in activity at 20 μM Co^{2+} in a lower background of substrate, or grow in of a new species which obscures the dip.

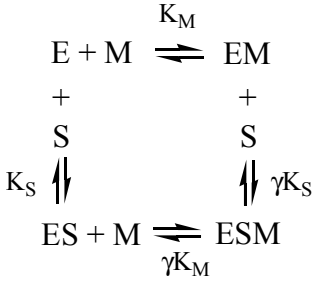
Thus overall, in both M and S profiles, noncompetitive/uncompetitive inhibition by both substrate and inhibitor are for certain; however, it is not at first obvious how either reactivation of inhibited species or grow in of new species contributes to the various rising phases of activity, especially since an apparently uniform phase may

[12] The absence of the initial peak then dip in an earlier study on metal dependence (Cheng 1999) is attributed to greater error in the Co^{2+} stocks at low concentration.

actually obscure contributions from an ensemble of active species. Simulations with minimal assumptions were found to limit the possibilities.

Basic models for monomer and dimer kinetics were derived along the lines of the Simple Sequential Interaction Allosteric Method (Adair 1925; Pauling 1935) assuming rapid equilibria (Segel 1975), then the velocity equations derived from these schemes were fit separately but by similar methods using Microsoft Excel™ 2002 to both the M and S profiles (see Appendix D for details). Since basic sequential monomer (Fig. 22) or dimer models (Fig. 24) predict simple hyperbolic behavior, more terms were added to the basic equations to account for the shifts in activity due to inhibition, activation, or grow in of active species.

Monomer Model



$$1)\quad v = \frac{k_{cat}[E]_t\left(\frac{\overbrace{[M][S]}^{EMS}}{\gamma K_M K_S}\right)}{\frac{1}{E} + \underbrace{\frac{[M]}{K_M}}_{EM} + \underbrace{\frac{[S]}{K_S}}_{ES} + \underbrace{\frac{[M][S]}{\gamma K_M K_S}}_{EMS}}$$

$$2)\quad v = \frac{k_{cat}[E]_t[S]}{\gamma K_S\left(\frac{K_M}{[M]}+1\right)+[S]\left(\frac{\gamma K_M}{[M]}+1\right)}$$

Fig. 22 Basic Monomer Model: the general velocity expression (Eq. 1) can be rearranged to show the explicit hyperbolic dependence on S (Eq. 2). The dependence on M has the same form as in (Eq. 2) but with M and S switched.

Since the amino acid sequence and crystal structure indicate only one apparent active site per monomer, the simplest active species would be EM'S' (enzyme + catalytic Co^{2+} + substrate being cleaved); however, it was found that multiple, additionally bound 'effector' ligands (both metal and substrate) were required in order to explain the observed non-linear dependence of activity on ligand concentration (i.e. in order to introduce the required sigmoidicity which causes species to grow in 'later' in the profile at higher ligand concentration). A similar occurrence of multiple bound ligands (substrate and inhibitors) has long been observed in CPA (Lumry, Smith et al. 1951; Quiocho and Richards 1966). Overall, it was found that a minimum of six 'core populations' of active species were required to fit the data to the monomer model, each with a different and distinct number of bound effectors. In concert, these ligands act essentially as activators in that they are required to be bound before any activity is

observed; however, microscopically some of these effectors might be inhibitory in terms of their individual allosteric effect.

Core populations can be arbitrarily categorized by the number of effector metals bound (see Tables 6 and 7). Binding of additional inhibitor or activator metals and substrates to these core species give derivative species (b, c) within each core population. Physically, these effectors are likely additional Co^{2+} and substrate molecules that interact and bind in the ample space of the active site groove in PfuCP. Their individual binding sites are likely different in terms of both affinity and allosteric effect on other bound ligands, but since no specific information is available on each site, and obtaining these parameters spectroscopically or by other means does not appear feasible due to the complex and heterogeneous mixture of core populations, the effectors on each active species are represented as an average, 'dummy' ligand which binds to several identical sites with a defined binding affinity and allosteric effect on other bound ligands (these effector ligands are asterisked and have numeric subscripts to indicate the number bound). For the same reasons, individual dummy constants could not be reported but are listed in Tables 6 and 7 as a figure of merit (linear product of individual constants) that is equivalent to the product of the true binding constants and their interaction factors. If the true interaction factors are all close to 1 meaning allosteric effects are minimal, all the values reported will be close to the actual values[13].

It was found that fitting of distinct features in the activity profiles was possible as the effects of different core populations were sufficiently separated (Fig. 23) such that at lower ambient ligand concentrations, active species with lower numbers of these bound

[13] See Appendix D for simulation details

'effectors' are higher in number and dominate in terms of their effect on overall activity, while at higher ambient ligand concentrations, species with more of these bound effectors control the shape of the profiles. Species that control the shape of the profile have the largest terms in the velocity equations and dominate both the denominator and numerator in Eq. 1.

General equations derived and fit separately for both the M and S profiles were compared, checked for consistency, and merged into the following single equation which completely describes the steady-state kinetic behavior of PfuCP for Co^{2+} and the N-blocked dipeptide Z-Ala-Arg.

Eq. 3 -General Expression for Rapid Equilibrium Steady-State Velocity of PfuCP + Co^{2+} + Z-Ala-Arg

$$v = \frac{k_{cat}[E]_t \left(\overbrace{\frac{m_{1a}[M][S]^3}{K_{M'}K_{S'}K_{S^*}^{\ 2}}}^{TERM\,1a} + \overbrace{\frac{m_2[M]^8[S]^9}{K_{M'}K_{S'}K_{M^*}^{\ 7}K_{S^{**}}^{\ 8}}}^{TERM\,2} + \overbrace{\frac{m_3[M]^{12}[S]^6}{K_{M'}K_{S'}K_{M^{**}}^{\ 11}K_{S^{***}}^{\ 5}}}^{TERM\,3} + \overbrace{\frac{m_{4a}[M]^{17}[S]^9}{K_{M'}K_{S'}K_{M^{***}}^{\ 16}K_{S^{****}}^{\ 8}}}^{TERM\,4a} + \overbrace{\frac{m_{5a}[M]^{22}[S]^{12}}{K_{M'}K_{S'}K_{M^{****}}^{\ 21}K_{S^{*****}}^{\ 11}}}^{TERM\,5a} + \overbrace{\frac{m_{6a}[M][S]^2}{K_{M'}K_{S'}K_{S^{*****}}}}^{TERM\,6a} \right)}{\left(\begin{array}{l} 1 + \underbrace{\frac{[M][S]^3}{K_{M'}K_{S'}K_{S^*}^{\ 2}}}_{TERM\,1a} + \underbrace{\frac{[M]^4[S]^3}{K_{M'}K_{S'}K_{S^*}^{\ 2}K_{Mi1}^{\ 3}}}_{TERM\,1b} + \underbrace{\frac{[M][S]^6}{K_{M'}K_{S'}K_{S^*}^{\ 2}K_{Si1}^{\ 3}}}_{TERM\,1c} + \underbrace{\frac{[M]^8[S]^9}{K_{M'}K_{S'}K_{M^*}^{\ 7}K_{S^{**}}^{\ 8}}}_{TERM\,2} + \underbrace{\frac{[M]^{12}[S]^6}{K_{M'}K_{S'}K_{M^{**}}^{\ 11}K_{S^{***}}^{\ 5}}}_{TERM\,3} \\ + \underbrace{\frac{[M]^{17}[S]^9}{K_{M'}K_{S'}K_{M^{***}}^{\ 16}K_{S^{****}}^{\ 8}}}_{TERM\,4a} + \underbrace{\frac{[M]^{18}[S]^9}{K_{M'}K_{S'}K_{M^{***}}^{\ 16}K_{S^{****}}^{\ 8}K_{Mi2}}}_{TERM\,4b} + \underbrace{\frac{[M]^{22}[S]^{12}}{K_{M'}K_{S'}K_{M^{****}}^{\ 21}K_{S^{*****}}^{\ 11}}}_{TERM\,5a} + \underbrace{\frac{[M]^{22}[S]^{23}}{K_{M'}K_{S'}K_{M^{****}}^{\ 21}K_{S^{*****}}^{\ 11}K_{Si2}^{\ 10}}}_{TERM\,5b} \\ + \underbrace{\frac{[M][S]^2}{K_{M'}K_{S'}K_{S^{*****}}}}_{TERM\,6a} + \underbrace{\frac{[M]^4[S]^2}{K_{M'}K_{S'}K_{S^{*****}}K_{Mi3}^{\ 3}}}_{TERM\,6b} + \underbrace{\frac{[M][S]^5}{K_{M'}K_{S'}K_{S^{*****}}K_{Si3}^{\ 3}}}_{TERM\,6c} \end{array} \right)}$$

Table 6 -Kinetic Parameters from Simulations of PfuCP 'S Profiles'

	µM Co^{2+} (Back-ground)	Product of 'Dummy' Constants for Co^{2+} and Substrate	Terms and k_{cat} Multiplication Factors (m_n) for Active Species				Inhibitor Contants (mM)	
Core Population 1 **$ES^*_2M'S'$**		K_{S^*} (mM)	**ES^*_2MS** Term 1a	**$ES^*_2M'S'(M_{i1})_3$** Term 1b	**$ES^*_2M'S'(S_{i1})_3$** Term 1c		K_{Mi1}	K_{Si1}
	10	2	1				Function	3.5
	20	2	1				of	3.5
	40	2	1				S	3.5
	60	2	1				(see	3.5
	100	2	1				Table 7)	3.5
	200	2	1					3.5

Core Population 2 **$EM^*_7S^{**}_8M'S'$**		$(K_{M^*})(K_{S^{**}})$ $(mM)^2$	**$EM^*_7S^{**}_8M'S'$** Term 2									
			0.3	0.4	0.7	1.7	2.5	4	6	8	10	← mM ZAR
	10	0.0462	0.1	0.1	0.1	0.1	0.1	0.35	0.37	0.37	0.37	
	20	0.0462	0.1	0.1	0.1	0.1	0.1	0.35	0.37	0.37	0.37	
	40	0.0462	0.1	0.1	0.1	0.1	0.1	0.35	0.37	0.4	0.45	
	60	0.0462	0.1	0.1	0.1	0.1	0.1	0.4	0.45	0.45	0.45	
	100	0.0462	0.1	0.1	0.1	0.1	0.1	0.4	0.45	0.45	0.45	
	200	0.0462	0.1	0.1	0.1	0.1	0.1	0.4	0.45	0.45	0.45	

Table 6 -S Profiles Cont'd

	μM Co^{2+} (Back-ground)	Product of 'Dummy' Constants for Co^{2+} and Substrate	Terms and k_{cat} Multiplication Factors for Active Species			Inhibitor Constants (mM)	
Core Population 3 **$EM^{**}_{11}S^{***}_{5}M'S'$**	10 20 40 60 100 200	$(K_{M^{**}})(K_{S^{***}})$ $(mM)^2$ f(S)	**$EM^{**}_{11}S^{***}_{5}M'S'$** Term 3 f(S)				
Core Population 4 **$EM^{***}_{17}S^{****}_{8}M'S'$**	10 20 40 60 100 200	$(K_{M^{***}})(K_{S^{****}})$ $(mM)^2$ f(S)	**$EM^{***}_{16}S^{****}_{8}M'S'$** Term 4a f(S)	**$EM^{***}_{16}S^{****}_{8}M'S'(M_{i2})$** Term 4b		K_{Mi2} f(S)	

Table 6 -S Profiles Cont'd

	μM Co^{2+} (Back-ground)	Product of 'Dummy' Constants for Co^{2+} and Substrate	Terms and k_{cat} Multiplication Factors for Active Species			Inhibitor Contants (mM)	
Core Population 5 **$EM^{****}{}_{21}S^{*****}{}_{11}MS$**	10 20 40 60 100 200	$(K_{M****})(K_{S*****})$ $(mM)^2$ f(S)	**$EM^{****}{}_{21}S^{*****}{}_{11}M'S'$** Term 5a f(S)	**$EM^{****}{}_{21}S^{*****}{}_{11}M'S'(S_{i2})_{10}$** Term 5b		K_{Si2} f(S)	
Core Population 6 **$ES^{*}M'S'$**	10 20 40 60 100 200	K_{S*****} (mM) 1 1 1 1 1 1	$ES^{******}M'S'$ Term 6a f(S)	$ES^{******}M'S'(M_{i3})_3$ Term 6b	$ES^{******}M'S'(S_{i3})_3$ Term 6c	KM_{i3} (mM) f(S)	K_{Si3} (mM) 3 3 3 3 3 3

Table 7 -Kinetic Parameters from Simulations of PfuCP Activity Profiles (M profiles)

	mM Substrate (Background)	Product of 'Dummy' Constants for Co^{2+} and Substrate	Terms and k_{cat} Multiplication Factors for Active Species			Inhibitor Constants (mM)		
Core Population 1 **$ES^*_2M'S'$**		K_{S*} (mM)	**$ES^*_2M'S'$** Term 1a	**$ES^*_2M'S'(M_{i1})_3$** Term 1b	**$ES^*_2M'S'(S_{i1})_3$** Term 1c	K_{Mi1}	K_{Si1}	
	0.3	2	1			0.01	3.5	
	0.4	2	1			0.01	3.5	
	0.7	2	1			0.01	3.5	
	1.7	2	1			0.01	3.5	
	2.5	2	1			0.009	3.5	
	4	2	1			0.008	3.5	
	6	2	1			0.007	3.5	
	8	2	1			0.004	3.5	
	10	2	1			0.0017	3.5	
Core Population 2 **$EM^*_7S^{**}_8M'S'$**		$(K_{M*})(K_{S**})$ $(mM)^2$	**$EM^*_7S^{**}_8M'S'$** Term 2					
	0.3	0.0462	f(S) and f(M)					
	0.4	0.0462						
	0.7	0.0462	See Table 6 above					
	1.7	0.0462						
	2.5	0.0462						
	4	0.0462						
	6	0.0462						
	8	0.0462						
	10	0.0462						

Table 7 –M Profiles Cont'd

Core Population 3 **$EM^{**}_{11}S^{***}_{5}M'S'$**		$(K_{M^{**}})(K_{S^{***}})$	**$EM^{**}_{11}S^{***}_{5}M'S'$**			
		$(mM)^2$	Term 3			
	0.3	0.0225	0.025			
	0.4	0.0225	0.05			
	0.7	0.025	0.11			
	1.7	0.05	0.30			
	2.5	0.05	0.33			
	4	0.05	0.35			
	6	0.05	0.35			
	8	0.05	0.35			
	10	0.05	0.45			
Core Population 4 **$EM^{***}_{17}S^{****}_{8}M'S'$**		$(K_{M^{***}})(K_{S^{****}})$	**$EM^{***}_{16}S^{****}_{8}M'S'$**	**$EM^{***}_{16}S^{****}_{8}M'S'(M_{i2})$**	K_{Mi2}	
		$(mM)^2$	Term 4a	Term 4b		
	0.3	0.0018	0.5		0.01	
	0.4	0.0018	1		0.01	
	0.7	0.0024	1		0.029	
	1.7	0.0086	1.2		0.085	
	2.5	0.0092	1.4		0.17	
	4	0.0094	1.4		0.18	
	6	0.0094	1.4		0.19	
	8	0.0094	1.4		0.19	
	10	0.0094	1.6		0.19	

Table 7 -M Profiles Cont'd

	mM Substrate (Background)	Product of 'Dummy' Constants for Co^{2+} and Substrate	Terms and k_{cat} Multiplication Factors for Active Species			Inhibitor Constants (mM)	
Core Population 5 **$EM^{****}{}_{21}S^{*****}{}_{11}M'S'$**		$(K_{M****})(K_{S*****})$	**$EM^{****}{}_{21}S^{*****}{}_{11}M'S'$**	**$EM^{****}{}_{21}S^{*****}{}_{11}M'S'(S_{i2})_{10}$**		K_{Si2}	
		$(mM)^2$	Term 5a	Term 5b			
	0.3	0.0835	0.5			8	
	0.4	0.0835	0.5			8	
	0.7	0.0835	0.5			8	
	1.7	0.0835	0.5			8	
	2.5	0.0835	0.5			8	
	4	0.0835	1			8	
	6	0.087	1.5			8	
	8	0.088	1.75			9	
	10	0.088	2.2			9.85	
Core Population 6 **$ES^{*}M'S'$**		$K_{S******}$	$ES^{******}M'S'$	$ES^{******}M'S'(M_{i3})_3$	$ES^{*}M'S'(S_{i3})_3$	K_{Mi3}	K_{Si3}
		(mM)	Term 6a	Term 6b	Term 6c	(mM)	(mM)
	0.3	1	0.8			0.0025	3
	0.4	1	0.8			0.005	3
	0.7	1	0.6			0.005	3
	1.7	1	0.6			0.0095	3
	2.5	1	0.6			0.01	3
	4	1	0.6			>0.01	3
	6	1	0.6			>0.01	3

	8 10	1 1	0.6 0.6			>0.01 >0.01	3 4

-the nominal values for the dissociation constants of the 'active' ligands M' and S' are Km = 0.01 mM and Ks = 1 mM respectively for all simulations
-k_{cat} is 2 x 10^5 min^{-1}
-the concentration of monomeric PfuCP is 6 x 10^{-7} mM

Tables 6 and 7 contain the parameters derived from fitting either the M or S profiles and although all values are redundant, separate tables better illustrate the dependence of various parameters on background M and S respectively.

The total enzyme concentration (6×10^{-10} M) was determined from A_{280} measurements on the purified enzyme ($\varepsilon = 97720\ M^{-1}\ cm^{-1}$) (Gill and von Hippel 1989) and throughout all of the simulations a k_{cat} of $2 \times 10^5\ min^{-1}$ was found to give the best fit. Also included in the simulation parameters are k_{cat} multiplier factors which allow for variation in the k_{cat} for individual enzyme species due to the allosteric effect of different ligands. The interaction factors inherent to the Simple Sequential Interaction Model of Allostery accounts for the homo- or heterotropic allosteric effect that a given ligand has on the binding of another when both are coupled in a thermodynamic cycle and equilibrium is assumed. These interaction factors were collapsed into the dummy constants since the separate allosteric effects cannot be discerned from the experimental data[14] such that in cases where a given ligand is exerting a negative allosteric effect and causing other ligands to bind less tightly (interaction factor >1), the value determined for the binding constant of this particular ligand represents an upper limit to the actual value, whereas when a given ligand is causing all other ligands to bind tighter (interaction factors < 1), the simulated, effective binding constant represents a lower limit. As the strength of the allostery decreases, all interaction factors approach 1 and the effective dissociation constants approach their intrinsic values.

All terms in the general velocity equation contain K_M (0.01 mM), K_S (1 mM), and gamma (1) (nominal values in parentheses), which describe the active core EM'S',

[14] see appendix D for details

and although these individual constants are estimates derived from inspection of the concentration ranges in the respective M or S activity profile, their product is exactly equal to the product of the true binding constants in the absence of allosteric effects. Gamma is set to one for simplicity essentially collapsing any allosteric effects into the product of $K_{M'}$ and $K_{S'}$. Similarly, for other terms where separate constants cannot be simulated in a meaningful way, their product is reported and fit as an overall single variable.

In the simulations undertaken here, two general types of dissociation constants, K_M and K_S, describe the binding of metal and substrate respectively (here K_M is not the Michaelis-Menten constant, K_m, which is only defined for the simpler case of hyperbolic kinetics which are linear with respect to ligand concentrations). If the ligand is inhibitory, the subscript 'i' is appended (e.g. K_{Mi1}, K_{Si2}), where the number represents a unique binding site. Since there is apparently only one active site per subunit, there is only one pair of 'active' metal and substrate denoted as M' and S', while other ligands that inhibit or activate are labeled accordingly (e.g. M_{i1}).

Downturns in both M and S profiles could only be simulated by uncompetitive inhibition by M or D. For M profiles, a rise in activity could only be modeled as a new species growing in rather than reactivation by activator binding to an inhibited species, since the latter introduced extraordinarily large terms in the simulations that skewed the profiles. Inactive terms without the complete active M'S' core (e.g. EM', ES', $EM^{*}_{7}S^{**}_{8}$, or with less ligands bound $EM^{*}_{4}S^{**}$) had a negligible effect on the activity when modeled in which suggests that cooperativity involving the active core is large and effector ligands do not bind in its absence.

The kinetic parameters appear to become dependent on ligand concentration at higher levels of substrate and metal. Two possible explanations are that i) this is a non-specific effect of unbound and bound ligands interacting, and ii) the change in parameter values represents a shift in the contributions of microscopically different enzyme species. Since the dummy ligands actually represent the 'average' effect of different sites with different binding affinities, increasing the ligand concentration will first populate the most tightly bound sites followed by more labile ones. If these sites differ in their effects on k_{cat} as well as on dissociation constants and interaction factors of other ligands, the kinetics will reflect this shift. In cases where binding constants do not appear to vary with ligand concentration, it is also possible that different microscopic sites are progressively occupied, however, the different allosteric effects cancel so that there is no net effect apparent. It should be noted that additional terms could not be added to the general velocity equation (Eq. 3) in order to account for the observed kinetics while maintaining all parameters constant.

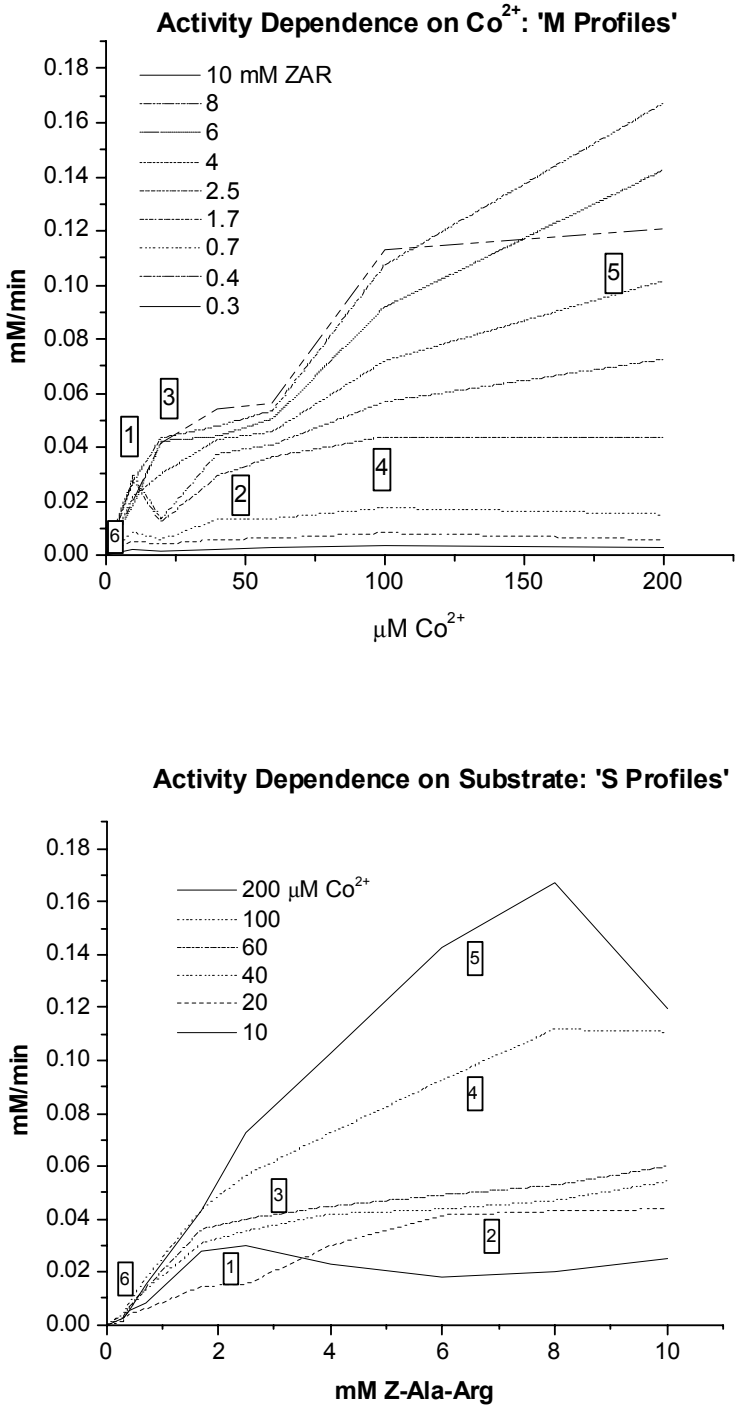


Fig. 23 Monomer simulations of M and S profiles with regions labeled by the predominant core population present.

The activity profiles can now be revisited for a more meaningful interpretation of the activity.

Core species 1, $ES^{*}_{2}M'S'$, is a low metal, low substrate form of the enzyme that is inhibited by both metal (Term 1b) and substrate (Term 1c) at higher ambient concentrations. The simulations further suggest that metal inhibition of this species increases at higher levels of substrate (K_{Mi1} decreases) indicating synergistic interaction between substrate and metal.

Core species 2, $EM^{*}_{7}S^{**}_{8}M'S'$, grows in at slightly higher concentrations of Co^{2+} and substrate, and again synergy between metal and substrate is evident from the simulations (Tables 6 and 7) which show that the k_{cat} multiplication factor (hence the effective k_{cat} for this species) increases with either Co^{2+} or substrate concentration, and in the particular case of Co^{2+}, significant activation by this metal only occurs at progressively higher background substrate beyond an apparent threshold at 2.5 mM ZAR. This illustrates a positive, cooperative response resulting in an overall enhancement of activity at higher ligand concentrations.

Core species 3, $EM^{**}_{11}S^{***}_{9}M'S'$, grows in at even higher metal and substrate concentration but it appears that the binding of either metal or substrate is reduced as the background substrate concentration increases (Table 7), due to the negative allosteric effects of substrate molecules that bind 'later' (Abashkin, Burt et al.). The k_{cat} multiplier for species 3 increases with substrate, however, suggesting that the allosteric effects of 'later' substrate extends to either boosting the activity or neutralizing the inhibitory effects of 'earlier' effector substrate.

Core species 4, $EM^{***}{}_{17}S^{****}{}_{8}M'S'$, dominates the activity at higher ligand concentrations, and as for species 3, it appears that binding of additional substrate reduces the binding affinity of other ligands including an explicit inhibiting metal (K_{Mi2}), while simultaneously increasing the k_{cat}.

Core species 5 acts at higher metal and substrate and similar to Terms 3 and 4, an increase in substrate apparently increases k_{cat} and reduces ligand binding, including that of an explicit inhibitor substrate (K_{Sa2}).

The M profiles with higher concentrations of 'background' ZAR (2.5-10 mM) were fit first because of the greater activity, larger 'signal to noise', and more distinct transitions in the traces, however more information can be gained by extending the model to the lower substrate profiles (0.3-1.7 mM ZAR). In this range it was necessary to introduce a new core species (ES*M'S' or Term 6) to account for the slight activity peak at 10 μM Co^{2+} (Fig. 23) that is predicted in the 5 term model to disappear since less ES*2MS (species 1) is present. Core population 6 is similar to core population 1 except that one of the effector substrates is missing, which reflects the lower concentration of ambient substrate. The dummy binding constant for this single effector substrate ($K_{S******} = 1$) is less than that of species 1 ($K_{S*} = 2$) since the former describes binding to a single site while the latter describes binding to two inequivalent sites. Also, the higher k_{cat} multiplication factors for species 1 compared with species 6 suggests that more than two bound substrates are required for full activity. Thus, binding a second effector substrate to give $EMS^{*}{}_{2}S$ correlates with reduced ligand binding affinity as well as simultaneous higher activity, a trend repeated in Terms 3, 4, and 5. Similar to species 1,

species 6 is modeled as being inhibited by both metal and substrate (Terms 6b and 6c). The apparent drop in activity for species 6 from 0.4 to 0.7 mM background ZAR (the k_{cat} multiplier decreases from 0.8 to 0.6) has no obvious explanation, although a generic interaction between the increasing amounts of ambient substrate and the ligand bound enzyme complex may be at work.

Quite notably, in all of the simulations no amount of EM'S' (an active species without any bound effectors) can be modeled to fit the data since this would predict more activity appearing at very low concentrations of substrate and Co^{2+} than is observed.

Dimer Model

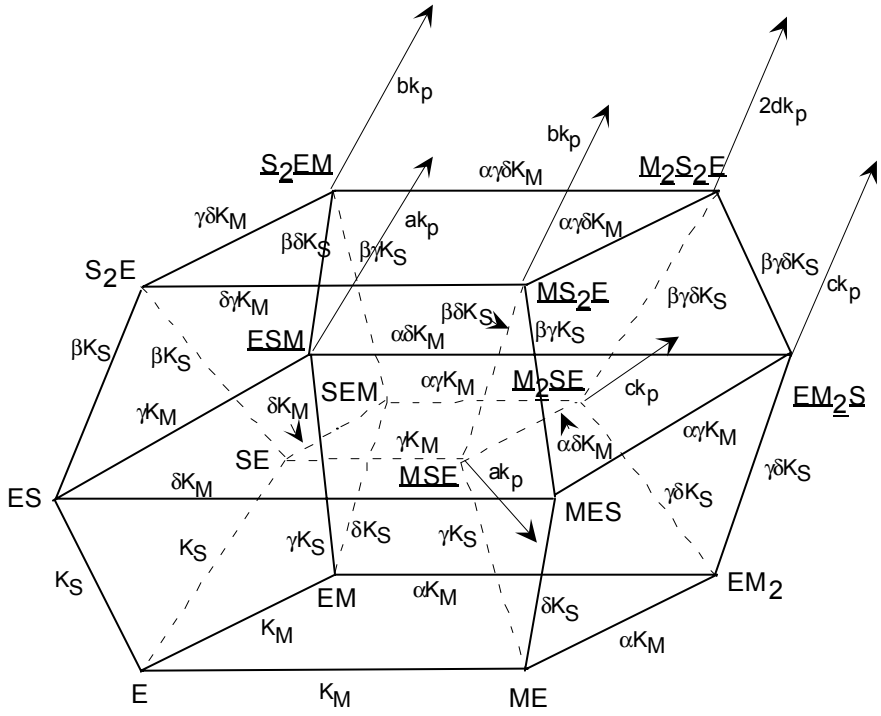


Fig. 24 Basic Dimer Model: the subunits are assumed to be identical with only one active site each. Rapid equilibrium conditions are assumed with no inhibition or activation and allostery is manifested in the interaction factors.

The simplest representation for dimeric PfuCP is a 16 corner hypercube (Fig. 24). Moving in a horizontal plane, either right or back, represents binding of Co^{2+} (M) to either subunit in the dimer (M or S on the left and right sides of E signifies one subunit or the other). Moving up vertically represents the addition of substrate (S) to either subunit. Symmetry in the binding of ligands yields symmetrical pairs for partially occupied complexes (e.g. EMS and MSE). The underlined enzyme-substrate-metal complexes represent product forming species: <u>EMS</u> (and <u>MSE</u>), <u>EM_2S</u> (and <u>M_2SE</u>), <u>EMS_2</u> (and <u>MS_2E</u>), and <u>EM_2S_2</u> with their respective k_{cat} multipliers 1, b, c, and d. Although SEM and its counterpart MES have both substrate and metal bound, the ligands occupy different subunits and hence do not generate product. The interaction factors (α, β, γ, and δ) represent potential allosteric interactions.

Eq. 4 –Basic Velocity Equation for Dimeric PfuCP (assuming rapid equilibrium and no inhibition or activation)

$$v = \frac{2k[E]_t\left(\overbrace{\frac{[M][S]}{\gamma K_m K_s}}^{EMS} + \overbrace{\frac{b[M][S]^2}{\beta\gamma\delta K_m K_s^{\,2}}}^{EMS_2} + \overbrace{\frac{c[M]^2[S]}{\alpha\delta\gamma K_m^{\,2} K_s}}^{EM_2S} + \overbrace{\frac{d[M]^2[S]^2}{\alpha\beta\gamma^2\delta^2 K_m^{\,2} K_s^{\,2}}}^{EM_2S_2}\right)}{1 + \underset{E}{} \underbrace{\frac{2[M]}{K_m}}_{EM} + \underbrace{\frac{2[S]}{K_s}}_{ES} + \underbrace{\frac{2[M][S]}{\gamma K_m K_s}}_{EMS} + \underbrace{\frac{2[M][S]}{\delta K_m K_s}}_{MES} + \underbrace{\frac{[M]^2}{\alpha K_m}}_{EM_2} + \underbrace{\frac{[S]^2}{\beta K_s^{\,2}}}_{ES_2} + \underbrace{\frac{2[M]^2[S]}{\alpha\gamma\delta K_m^{\,2} K_s}}_{EM_2S} + \underbrace{\frac{2[M][S]^2}{\beta\gamma\delta K_m K_s^{\,2}}}_{EMS_2} + \underbrace{\frac{[M]^2[S]^2}{\alpha\beta\gamma^2\delta^2 K_s^{\,2} K_m^{\,2}}}_{EM_2S_2}}$$

The terms in Tables 6 and 7 are empirical fits to the data independent of the model chosen, so while core population 1 has an obligatory third-order substrate dependence in order to properly fit the 10 μM S profile, the assumption of one metal bound for Term 1 (just the active metal) was made purely on the grounds of minimal

assumptions. It is however possible to fit Term 1 (and Term 6) with one additional effector metal yielding $EM^{*****}S^{*}_{2}M'S'$ (and $EM^{*****}S^{*****}M'S'$). Given this, all of the terms in Tables 6 and 7 can be derived from any of the core active terms in the basic equation for the dimer (Eq. 4) by addition of multiple M and S. Since it is not possible to tell *a priori* which active dimer terms correspond to which empirically observed active species, the model can only be fit by collapsing all the binding constants and interaction factors into the dummy constants.

The key difference between monomer and dimer is that the dimer has an active species with two fully occupied active sites; however, it is apparent that there are no distinguishing features for this in simulations. Since there are four times the number of active species compared with the monomer, assignment of observed species (Table 6 and 7) to derivative forms of the above dimer terms becomes increasingly ambiguous as does the meaning of the individual parameters so any given fit of the dimer model is not particularly meaningful. But although the simulations cannot discriminate between monomer and dimer, a predominantly monomeric form is plausible given the low concentration of enzyme in the assays (6×10^{-10} M monomer) and the fact that in order to fit the dimer model to the some of the empirically derived terms, ligand would have to be distributed asymmetrically between the two subunits which seems unlikely given that the subunits in the crystal structure appear identical; for example, if Term 2, $EM^{*}_{7}S^{**}_{8}M'S'$, were to be a dimeric species, two of the metals and two of the substrates would form the active core in each subunit leaving 6 metals and 7 substrates to distribute between the two subunits (the same situation would arise for Term 4).

A further complication could be that the reaction is actually a mixture of active monomers and dimers, in which case the activity profiles would represent a sum of fractional contributions from either form, the ratio of contributions dependent on the monomer-dimer equilibrium under the reaction conditions.

Regardless of the interpretation, the general velocity equation derived (Eq. 3) is representative of the kinetic behavior of the enzyme independent of oligomeric state and armed with a working model of the steady-state kinetics, the effects of metal inhibition (Zn^{2+}) can now be evaluated.

Zn^{2+} Inhibition

Most metalloproteases can be 'superactivated' with divalent cobalt when the native zinc has been replaced and although zinc is catalytically essential it has also been found that excess zinc inhibits activity. In PfuCP, this appears to have been taken to an extreme such that once the enzyme has lost any native metal during purification, it is only activated by the addition of Co^{2+}, not Zn^{2+}, nor any other metal (Ch. 2). This apparent preference for a metal which is not readily bioavailable and the rejection of one that is was explored through steady-state kinetics of Zn^{2+} inhibition.

PfuCP was incubated at a saturating concentration of Co^{2+} (200 μM) in the presence of increasing amounts of Zn^{2+}. The resulting activity profiles (Fig. 25) were simulated using the monomer model derived above (Eq. 3) by adding and adjusting terms to model rapid equilibrium competitive as well as non-competitive/uncompetitive inhibition.

Simulations of competitive inhibition by Zn^{2+} against substrate or Co^{2+} (which introduces $E(Zn^{2+})_n(M\ effector)_x(S\ effector)_y$ terms in the denominator without an M'S' 'active' core) cannot explain the experimental deactivation profiles (Fig. 25). In a qualitative sense, competitive inhibition would progressively flatten the activity profiles to zero from the left yielding a steeper left-hand shoulder, which is clearly not the case (competitive inhibition is apparent for the recombinant however see Fig. 26).

Since the core active terms 3, 4, and 5 dominate the velocity at high (Co^{2+}), the effects of any non-competitive/uncompetitive zinc inhibition are largely due to binding to these core species which gives three additional terms in the denominator of Eqn. 3.

$$\underbrace{\frac{[M]^{12}[S]^{6}[Zn^{2+}]^{2}}{K_{M'}K_{S'}K_{M**}{}^{11}K_{S***}{}^{5}(K_{iZn^{2+}})_{3b}{}^{2}}}_{TERM\ 3b}+\underbrace{\frac{[M]^{17}[S]^{9}[Zn^{2+}]^{2}}{K_{M'}K_{S'}K_{M***}{}^{16}K_{S****}{}^{8}(K_{iZn^{2+}})_{4c}{}^{2}}}_{TERM\ 4c}+\underbrace{\frac{[M]^{22}[S]^{12}[Zn^{2+}]^{2}}{K_{M'}K_{S'}K_{M****}{}^{21}K_{S*****}{}^{11}(K_{iZn^{2+}})_{5c}{}^{2}}}_{TERM\ 5c}$$

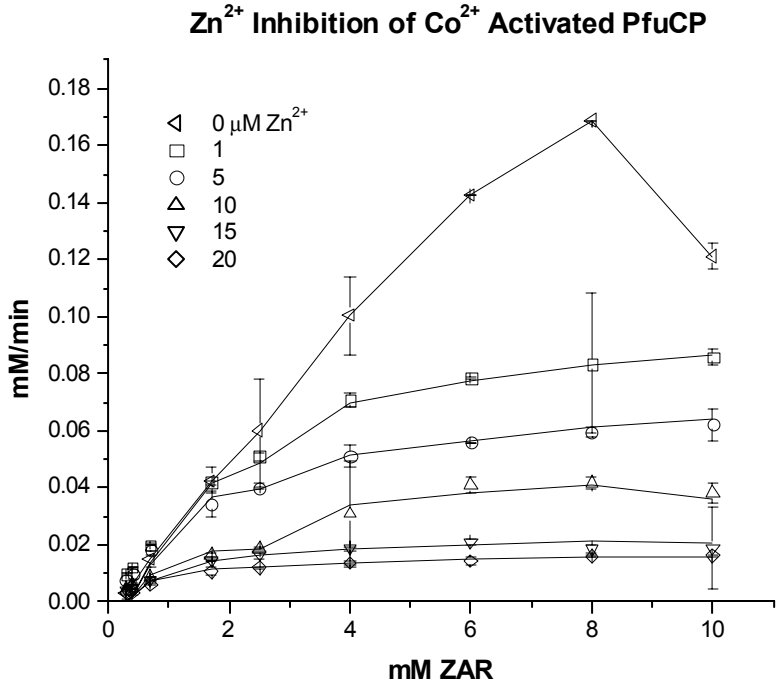


Fig. 25 Uncompetitive inhibition by Zn^{2+}: simulated activity profiles appear as lines joining duplicate experimental data points with standard error (200 μM $CoCl_2$, 0.1 M KMes pH 6.5, [E] = 6 x 10^{-7} mM).

Simulations of uncompetitive inhibition where Zn^{2+} binds and deactivates core species 3, 4 and 5 suggest that i) the k_{cat} multiplication factors for Term 5 are modulated by the presence of Zn^{2+} (Table 8); and ii) the Zn^{2+} dissociation constants for each core species increase as the concentration of Zn^{2+} rises (Table 8), which is reasonable providing more than one Zn^{2+} binds to inequivalent sites in each core species such that Zn^{2+} ions which add later (at higher [Zn^{2+}]) exert a negative allosteric effect on zinc already bound, perhaps due to electrostatic effects. These simulations (Fig. 25) assume that two Zn^{2+} ions bind to each core species, described by an 'average' binding constant, but the exact number of Zn^{2+} bound remains unknown, since the change in dissociation constants with [Zn^{2+}] prevents an unambiguous determination. By comparison, ZnCPA

inhibited by lead citrate has 2 Pb^{2+} bound, one in the corresponding Zn^{2+} inhibitory site and a second further liganded by nonprotein attachment to the citrate carboxyls (Larsen and Auld 1991; Lipscomb, Coppola et al. 1966).

Table 8 -Kinetic Parameters from Simulations of Uncompetitive Inhibition by Zn^{2+}

μM Zn^{2+}	$(K_{Zn2+})_{3b}$ (Term 3b) (mM)	$(K_{Zn2+})_{4c}$ (Term 4c) (mM)	$(K_{Zn2+})_{5c}$ (Term 5c) (mM)
1	$<1 \times 10^{-2}$	1×10^{-3}	1×10^{-3}
5	1×10^{-2}	3.4×10^{-3}	3.8×10^{-3}
10	1×10^{-2}	3.7×10^{-3}	5.8×10^{-3}
15	1×10^{-2}	6×10^{-3}	5.8×10^{-3}
20	1.3×10^{-2}	6×10^{-3}	6.5×10^{-3}

Table 9 -Adjusted Kinetic Parameters for Term 5 from Fit of Zn^{2+} Uncompetitive Inhibition

	k_{cat} multiplier
mM ZAR	$EM^{****}{}_{21}S^{*****}{}_{11}M'S'Zn_2$
0.3	0.9
0.4	0.9
0.7	0.9
1.7	0.9
2.5	1
4	1.2
6	1.3
8	1.4
10	1.5

-also, K_{Si2} increases from 10 to 13 mM at $[Zn^{2+}] < 10$ μM

If Zn^{2+} can indeed bind to the metal ligands in the catalytic site (there is no evidence to the contrary), then there is likely a small component of competitive behavior with Co^{2+}; however, the activity profiles for Zn^{2+} inhibition seem to suggest that the behavior of Zn^{2+} is predominantly uncompetitive.

Millimolar excess of zinc is known to inhibit several zinc proteases including CPA (Vallee, Rupley et al. 1960; Larsen and Auld 1989), thermolysin (Holmquist and Vallee 1974), neutral endopeptidase from rabbit kidney brush border (Kerr and Kenny 1974), human neutrophil collagenase (Mallya and Van Wart 1989), angiotensin converting enzyme (Bunning, Holmquist et al. 1983), and CPA. Crystallographic studies on CPA (Gomez-Ortiz, Gomis-Ruth et al. 1997) and thermolysin (Holland, Hausrath et al. 1995) clearly show that inhibition from excess Zn^{2+} results from a second Zn^{2+} that binds in a tetrahedral geometry between the conserved active site HEXXH glutamate and the catalytic Zn^{2+} (an aquo bridges the two metals). Although the mode of inhibition is competitive in CPA, the uncompetitive nature of the inhibition in PfuCP is manifest in the absence of any reactivation by Zn^{2+} and the much lower levels of ambient zinc required for inhibition. Compared with the equilibrium binding constants for other HEXXH metalloproteases, those estimated for PfuCP (Table 7) are higher in keeping with the general trend towards loose binding of metals ($K_i = 7 \times 10^{-7}$ M for CPA, (Larsen and Auld 1991). The apparent lack of competitive character for Zn^{2+} inhibition of PfuCP further suggests that the second inhibitory zinc bridges to a Co^{2+} as opposed to another Zn^{2+} ion in the catalytic metal site; and in fact, a similar Co^{2+}-Zn^{2+} binuclear complex has

been suggested from kinetic inhibition studies on CPA (Larsen and Auld 1991). Alternately, the inhibitory Zn^{2+} in PfuCP may bind in an as yet unknown mode altogether, but the uncompetitive character suggests an absolute requirement for the active EM'S' core for inhibition which is reasonable if the inhibiting Zn^{2+} is near the catalytic metal. By extension, the lack of reactivation by other divalent metals and their relative inhibition strength (Ch. 2) may reflect their similar binding to this secondary inhibition site. The enthalpy of the inhibitory metal-hydroxide bond has also been correlated with the strength of inhibition (Larsen and Auld 1991) and hence is predicted to roughly follow the Irving-Williams series. It is possible that Co^{2+} does not bind effectively to this inhibitory site due to the instability of placing two Co^{2+} in close proximity in tetrahedral dispositions since binding relative to the free hexaaquo ion is already disfavored due to a loss in ligand field stabilization energy (Berg and Merkle 1989).

If zinc or another metal is indeed the native, catalytic metal in PfuCP, then there must be a mechanism to reduce the binding of Zn^{2+} and other inhibitory metals to this secondary metal site. One possibility is that as the native protein folds, the metal is placed in a kinetic trap and the electrostatic effects of the localized charge and electrostriction on nearby residues prevents the second inhibitory site from being occupied. *In vivo*, chaperones, accessory proteins, or other ligands that bind PfuCP in a complex may favor this disposition; molecules that enhance the thermostability of enzymes have been discovered in hyperthermophiles such as cyclic-2,3-diphosphoglycerate in methanogens (Hensel and Konig 1988) and di-myo-inositol-1,1'-phosphate and mannosyl glycerate in *P. furiosus* (Martins and Santos 1995). However,

when native conditions are lost (when the protein is isolated for *in vitro* studies), the altered conformation of the protein will favor binding of the metal in the inhibitory site which for the same reasons mentioned above will disfavor binding of the catalytic metal in a truly native mode. In support of this, the CP from *Thermus aquaticus* could not be reactivated by Zn^{2+} after having been deliberately isolated as the *apo*enzyme, but upon expression in *E.coli*, the recombinant was found to contain a bound Zn^{2+} despite the presence of 1 mM $CoCl_2$ in the culture media (Lee, Taguchi et al. 1994). Zn^{2+} inhibition in *Taq* CP differs from PfuCP however in that mM levels of Zn^{2+} are required and no inhibition is observed at 10 μM.

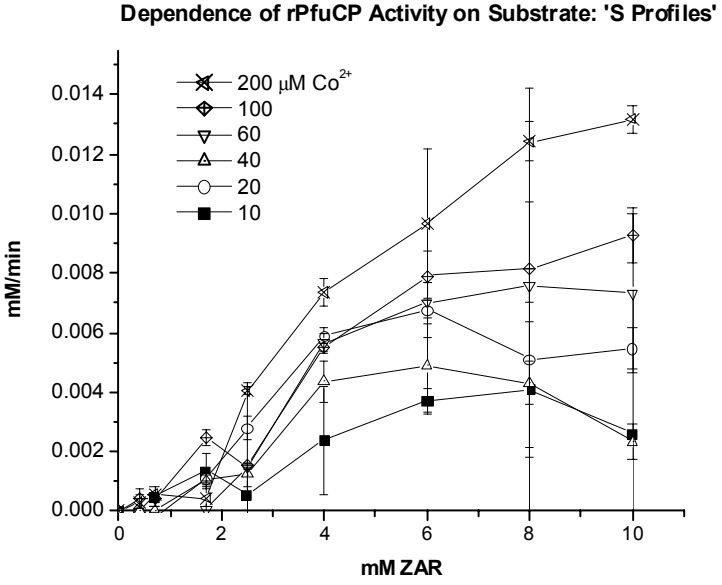
This kinetic hypothesis coincides with earlier observations that newly purified PfuCP appears to be very thermostable (Fig.13) with an alternate conformation (higher dimer band in SDS-PAGE, Fig. 12) but upon removal of stabilizing metal(s) both thermostability and the stabilizing conformation are lost. It is possible that the stabilizing metal is actually the native catalytic metal, and that the inability to reconstitute the enzyme (Ch. 2) demonstrates a kinetic trap that involves a metastable conformation.

If, however, Co^{2+} is the native metal then there must be mechanisms to incorporate it into the enzyme without exposure to sulfides which might combine with Co^{2+} in insoluble complexes. Similar proteins that import and ferry Co^{2+} within archaebacteria are known in the cobalamin synthetic pathway (Frausto da Silva and Williams 2001).

So to generalize, there appear to be two separate issues regarding the relative activities of different metal substituted isoforms: i) the relative affinity of the metal for (uncompetitive) inhibition sites, and ii) the intrinsic ability of the metal to lower the

energy barrier to reaction. When the enzyme is studied in the presence of excess metal, opportunistic inhibition sites come into play, whereas in studies where an apoenzyme has been reconstituted with metal and washed clean of excess, the relative activities of different metalloforms reflect differences in intrinsic catalytic properties, an issue that will be explored in the following chapter (Ch. 5).

Steady-State Kinetics for the Recombinant (rPfuCP)



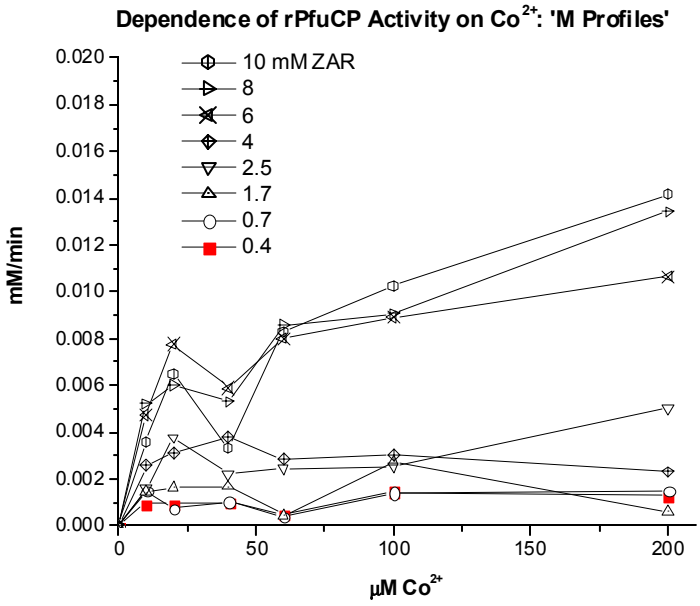


Fig. 26 Dependence of initial velocity on substrate and Co^{2+} for semi-purified rPfuCP from on-resin refolding. The same set of duplicate data points is shown in both graphs without simulations. Lines are only a visual guide and standard error is not shown in the M profiles for clarity. Conditions are 0.1 M KMes pH 6.5, 80 ^{o}C, ionic strength = 0.1 M under otherwise standard conditions.

Refolded, semi-purified rPfuCP from the Co^{2+} IMAC column eluant was used in these kinetic studies since the removal of the His-Tag by enterokinase currently yields insufficient amounts of stable enzyme for studies (cleavage is inefficient). These profiles (Fig. 26) were not simulated due to the unknown effects of other trace proteins and the His-Tag itself; anomalies for the recombinant with His-Tag were also observed in the temperature profiles for activity (Fig. 9). Nevertheless, the overall activity patterns are similar to those of the native enzyme (Fig. 21), illustrating the presence of multiple active species with varying amounts of metal and substrate bound. Differences in both the distribution of species and apparent composition of bound ligands might also result from structural artifacts from the refolding process.

One notable difference from native PfuCP is that the initial lefthand segment of the 200 μM 'S' profile of the recombinant (Fig. 26) is flattened which indicates competitive inhibition at high concentration of Co^{2+} between substrate and Co^{2+} where one excludes the other.

Conclusions

A model describing rapid pre-equilibrium ligand binding can be fit to the kinetic behavior of PfuCP; however, confirmation of a chemical step as rate-limiting awaits time-resolved kinetic studies. The complexity of the steady-state kinetics also prevents an unambiguous analysis of the pH activity profiles which can reflect changes in binding constants, interaction factors, and allosteric effects in addition to ionizations of critical active site residues. M and S profile studies carried out at lower temperatures (50 and 70 ^{o}C, not shown) do not reveal progressive changes in binding constants and k_{cat} multiplication factors but rather necessitate the construction of entirely new velocity equations with new terms and different numbers of bound effectors. Preliminary simulations do suggest, however, that the binding of substrate and metal **increases** as the temperature is lowered. Aside from the added flexibility that enzymes are presumed to exhibit at higher temperature (Laderman, Davis et al. 1993), it is possible that a lower dielectric at elevated temperatures would increase repulsion between multiple metals in close proximity, reducing binding affinities.

In fact, the physiological relevance of small, synthetic substrates as well as high *in vitro* concentrations of metal are unclear. Longer, natural substrates (unblocked and

unmodified) may not exhibit such complex kinetic behavior, and high enough free ion concentrations to give multiply liganded species may not be feasible, nevertheless, these simulations do yield insight into the molecular behavior of PfuCP.

The k_{cat} of 2 x 10^5 min^{-1} is larger than that for a similar substrate of CoCPA (k_{cat} = 12.3 x 10^3 for Z-Gly-Phe), which suggests perhaps an intrinsic difference in catalytic efficiency between hyperthermo and mesophilic CPs.

The inhibition by Zn^{2+} is clearly uncompetitive and the shift in allosteric effects as the concentration of Zn^{2+} increases suggests that two or more Zn^{2+} bind per active species.

Both substrate and metal can inhibit the enzyme as well as activate it, and in fact, many alternate configurations of substrate and metal allow for activity. This synergism between substrate and metal is evident in the variations in k_{cat} and ligand binding affinities as either the metal or substrate concentration is ramped. This allows for selectivity in that the activity of the enzyme can be modulated by the particular characteristics of ambient ligands which may be important for integrating the function of PfuCP into the various internal metabolic pathways of *Pyrococcus furiosus* as well as coupling its reponse to ligands that may arise from an unstable, dynamic external environment. This also indicates a robust ability to maintain activity despite numerous inhibitory configurations of substrate and metal, which may reflect the situation *in vivo* where the intracellular space is crowed with similar ligands. The apparent drastic changes in allosteric behavior with temperature further integrates a thermal response into the kinetic behavior of PfuCP. The simulations also suggest that varying the Co^{2+} concentration has less of an effect on the enzyme than substrate does, which suggests a

larger role for ‘softer’ ligands in modulating activity due to perhaps the increased richness of interactions such molecules can afford, which enables them to be more varied and subtle messengers of cellular information.

The minimum requirement for two substrates to be bound for any activity (and for three before full activity) suggests that more extensive substrate contacts along the active site groove are essential, although additional studies are needed to confirm the location of binding. The apparent high cooperativity between effector binding and formation of the active EM’S’ core suggests that the most plausible location for abundant interactions between multiple ligands would be in the active site groove (highly cooperative interactions and contacts would serve to reduce nonproductive binding).

The general trends from Tables 6 and 7 indicate that as more substrate binds (more extensive substrate-enzyme contacts), the affinity for all ligands decreases while the k_{cat} is augmented. This may reflect a more subtle functional role for the active site groove in addition to simple size selection of substrates. Larger substrates that do not bind as tight are less likely to be caught in kinetic traps, which might discourage processivity, and if the binding tightens as the substrate is shortened due to reduced enzyme-substrate contacts, then these shortened substrates will not be ‘lost’ as easily from the active site and hence have a ‘competitive advantage’.

C-terminal protein sequencing studies using PfuCP (Ch. 6) demonstrate that not all peptides can be cut to the same extent even if they are largely unstructured and small enough to fit inside the groove. The progressively reduced ligand binding demonstrated in these simulations suggests a general mechanism for the enzyme to disfavor binding of random substrates compared with peptides that might have specific, favorable contacts

with the enzyme groove. Larger, random peptides will not bind as tightly unless there are sequence specific contacts such as if the peptide happens to be a special metabolic target of PfuCP.

Chapter 5

Implications for Mechanism, Specificity, and Rate Theory

Abstract

Although plausible reaction schemes for metalloproteases have been proposed based on high-resolution structural information, the details appear more subtle when additional biochemical data is considered and a universal and unequivocal reaction mechanism remains elusive. The structural and biochemical aspects of PfuCP are now placed within the context of current thought on metalloprotease mechanisms, while the characteristics which make the enzyme unique (a loosely bound catalytic metal, activation only by cobalt) are explored for the new insight they offer regarding metallohydrolase behavior and enzyme catalysis in general.

Introduction

Specific mechanisms that have been proposed can be grouped into the following general pathways, although the particular strategy used by any given metalloprotease may vary depending on the structural details of both substrate and active site.

Possible Reaction Pathways

1) Anhydride
2) Activated-Carbonyl
3) Water-Promoted

4) Reverse Protonation

The anhydride mechanism asserts that a conserved glutamate sidechain near the catalytic metal (from the HEXXH or HXXE motif) acts as a nucleophile to attack the scissile peptide bond, generating an anhydride intermediate (acyl-enzyme) which is then hydrolyzed to release the truncated peptide. However, evidence for the mixed anhydride intermediate has come from only a handful of experiments, mostly on ester substrates which have higher inherent reactivity than peptide bonds towards nucleophilic attack (Makinen, Wells et al. 1984) (Kuo and Makinen 1985) (Britt and Peticolas 1992) (Mustafi and Makinen 1994). There is also a glaring absence of steady-state kinetic characteristics that typify such an intermediate despite extensive studies on the archetypical metalloproteases thermolysin and CPA; these include

1) observation of a common k_{cat} for a series of substrates that differ only in their C-terminal leaving group (if deacylation is rate limiting),
2) capture of the intermediate by nucleophiles (eg. NH_2OH, N_2H_4, alcohols),
3) observation of partitioning of products in the presence of a mixture of competing 'acceptor' nucleophiles (if acylation is rate limiting).

The above kinetic criteria have all been established for serine proteases which have well-characterized acyl-enzyme intermediates (Galdes, Auld et al. 1983).

Further assorted arguments against the anhydride pathway include the absence of an unambiguous 'burst' phase in pre-steady-state studies corresponding to the

stoichiometric release of leaving group (amino acids for exoproteases and peptides for endoproteases), the observation of a random order for substrate addition in the reverse reaction for thermolysin (peptide synthesis) whereas an acyl-enzyme intermediate would require an ordered mechanism (Wayne and Fruton 1983; Riechmann and Kasche 1986), and studies where incorporation of labeled water during peptide synthesis (transpeptidation) with CPA require the presence of nucleophilic amine; in this final case, the water is incorporated following addition of the amine then re-hydrolysis of the amide bond, whereas in an anhydride pathway, the labeled water should add directly to the anhydride and label product in the absence of amine (Breslow and Wernick 1977; Breslow, Chin et al. 1983). Further, reaction coordinate analogs with ketone or aldehyde moieties bind to the metal ion in CPA as *gem*-diol(ate)s rather than bonding covalently to the putative nucleophile glutamate (Christianson and Lipscomb 1986; Christianson, David et al. 1987; Shoham, Christianson et al. 1988).

Activated Carbonyl and Water Promoted Pathway

The catalytic metal can also act as a Lewis acid when bound to the oxygen of either a substrate carbonyl (peptide bond) or water, whereby the metal polarizes the carbonyl bond enhancing nucleophilic attack in the former and promotes ionization of water to a nucleophilic hydroxide in the latter. Precedents for both of these processes exist in various organic and inorganic model systems (Mock 1998) and in the case of water bound to Zn^{2+}, the pKa is lowered to 5 or 6 compared to free hexaaquo-zinc in solution which has pKa around 9 (Perrin 1982). Whether the metal ion polarizes either the carbonyl or water (or both) in general remains unresolved; however, evidence that the

mechanism (in CPA at least) cannot be purely water polarization arises from synergy between the metal and substrate in reactions with thionopeptide substrates (such as Bz-Gly-NH-CH_2-C(=S)NH-Phe) wherein the activity of CdCPA shows higher activity compared with ZnCPA, which conversely shows higher activity against normal carboxamide peptide substrates (Mock, Chen et al. 1981; Bartlett, Spear et al. 1981; Campbell and Nashed 1982; Bond, Holmquist et al. 1986). This was interpreted in terms of the polarizability of atoms (Pearson series, (Pearson 1963)) where favorable interactions occur between 'hard' atoms such as zinc and oxygen, and similarly between 'soft' atoms such as sulfur and cadmium.

However, in all metalloprotease structures to date the catalytic metal has been observed to bind at least one water in addition to its protein ligands.

Reverse Protonation

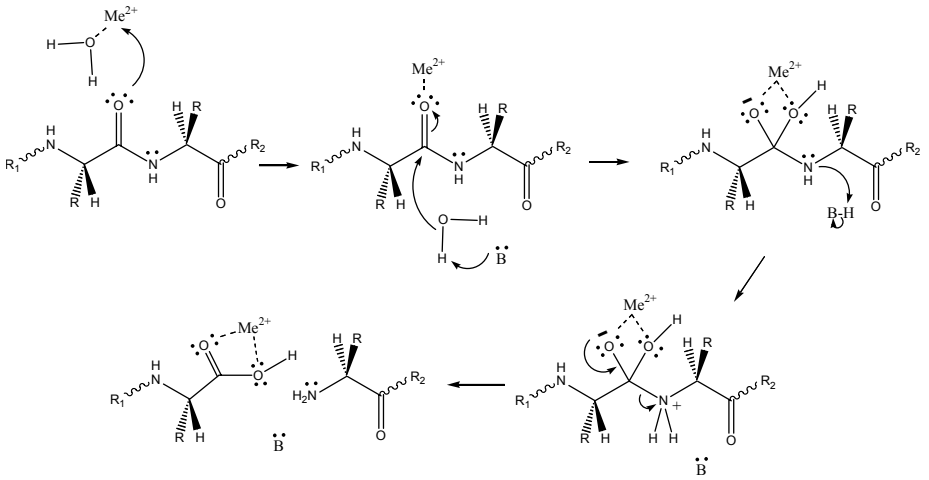


Fig. 27 Reverse protonation where metal bound water ($pK_a \sim 6$) is displaced by the scissile carbonyl and the nucleophilic hydroxide is generated by a general base other than the metal (either the substrate C-terminus or an active site residue such as histidine).

In reverse protonation (Fig. 27), the water initially bound to the metal is not the nucleophile that attacks the substrate, but rather a transient ligand that simply occupies a vacant coordination position prior to being displaced by the substrate carbonyl. Evidence for this mechanism has largely arisen from pH dependence studies on thermolysin where the two main residues responsible for the bell-shape in the k_{cat}/K_m vs. pH profiles (Fig. 28) are assigned pKa values which are opposite to what is commonly assumed, reversing the assignment of relative proton affinities and hence the name.

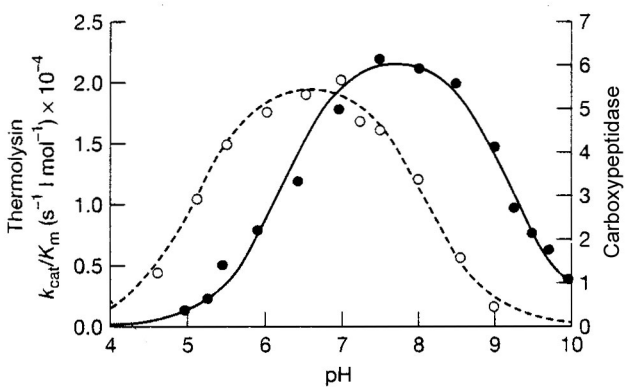


Fig. 28 pH profile for thermolysin and CPA (adapted from (Mock and Tsay 1988) and (Mock and Aksamawati 1994)).

Typically the alkaline limb (pKa ~8 or 9 for thermolysin and CPA respectively) is assigned to the metal-bound water since similar values have been reported for free zinc aquo species (Perrin 1982) while the acidic limb (pKa ~5 or 6 for thermolysin and CPA respectively) is linked to the active site glutamate implicated in general base catalysis.

Since His 231 in thermolysin has been established as one ionizing group (pKa ~8) by chemical modification studies with varying pH (Burnstein, Walsh et al. 1974; Kunugi, Hirohara et al. 1982), the other pKa ~5 is ascribed to water.

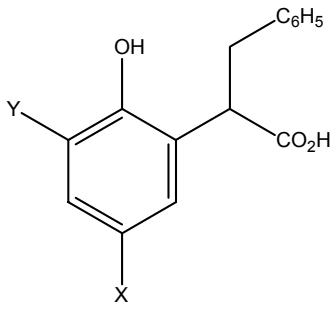


Fig. 29 Phenolic inhibitors that display competitive inhibition consistent with a reverse protonation pathway in thermolysin (Mock 1998); Y=H, Cl and X=NO_2, CN, and Cl.

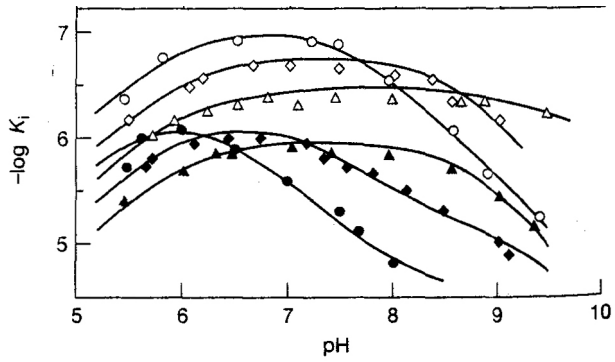


Fig. 30 pH dependence of competitive inhibition constants K_i for phenolic substrate analogs. Phenol ring substitutents (○), p-NO_2; (◊) p-CN; (Δ), p-Cl; (●), o-Cl, p-NO_2; (♦), o-Cl, p-CN; (▲), o, p-Cl_2 (adapted from (Mock 1998))

Equilibrium inhibition studies with phenolic substrate analogs (Fig. 30), where the phenolic hydroxylate binds directly to Zn^{2+}, have also been taken to be consistent with reverse protonation (Mock 1998). The pH dependence of the K_i (Fig. 30) links the inhibition strength with the phenolate (since the alkaline limb varies with phenol ring substitutent); however, a greater abundance of phenolate at higher pH should yield greater inhibition from increased ligation of the metal. The rationale for reverse protonation is that the drop off in inhibition strength at alkaline pH arises from

deprotonation of metal bound water that becomes harder to displace by phenolate whereas at lower pH the hydroxo is converted to a more labile water but then simultaneously protonated phenolic hydroxyl loses its nucleophilicity. At intermediate pH, a balance is struck whereby a maximum amount of phenolate ligation can occur with sufficient amounts of protonated and labile water.

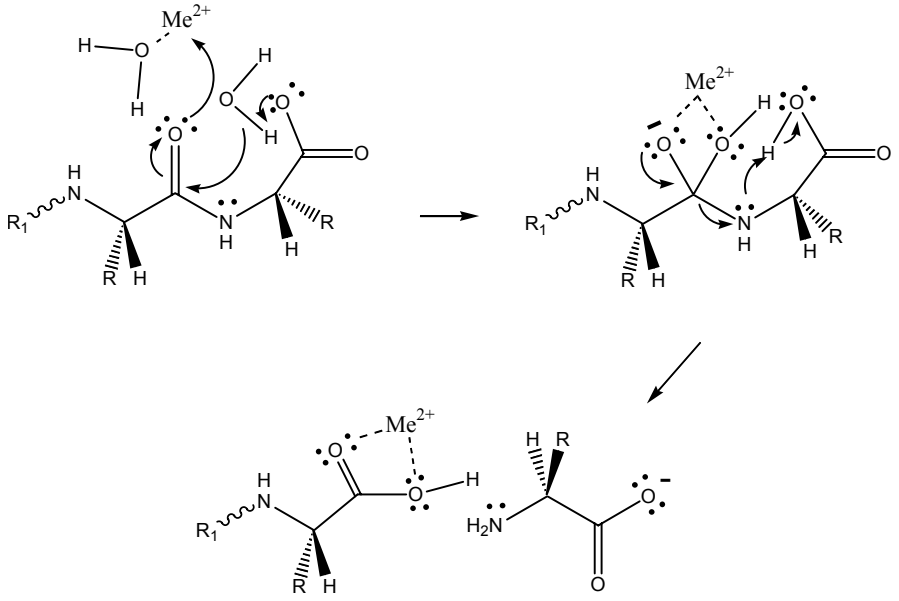


Fig. 31 Reverse protonation mechanism in CPA where the C-terminus acts intramolecularly as the general base/proton shuttle.

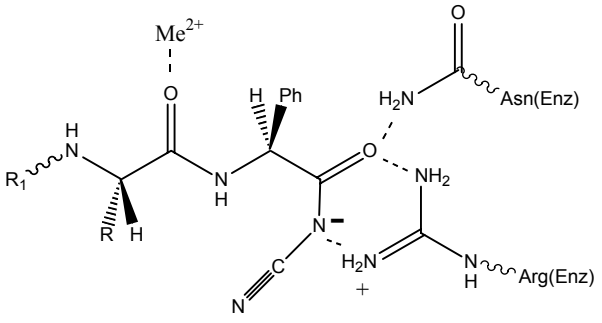


Fig. 32 Binding mode of cyanamide substrates to CPA (Mock 1998).

Since carboxypeptidases such as CPA lack a catalytically convenient histidine similar to that in thermolysin, it has been proposed that the C-terminal carboxyl of the substrate itself serves as a general base to deprotonate the attacking water (Fig. 32) (Mock and Xu 1994; Mock and Zhang 1991). Studies with acylcyanamide substrates (Fig. 33) have been taken to support this assertion since their binding mode is predicted to block intramolecular general-base catalysis (no lone pairs available unlike normal C-terminii) and only allow cleavage if Glu-270 acts as a general base; since these substrates are hydrolyzed on the order of 10^5 more slowly, this was taken to suggest that a competent C-terminal carboxyl as opposed to Glu-270 is the catalytically essential general-base (Mock and Xu 1994). Furthermore, structural and kinetic studies with chiral sulfoximine inhibitors which mimic a tetrahedral intermediate (Fig. 34) are suggestive of *si* face attack of water which is consistent with intramolecular general-base catalysis rather than by Glu-270 which would promote *re* face attack. In these inhibitors, the lower pK_a for the *si* diastereomer amine (left) is taken to indicate preferential polarization of the analogously located oxygen in tetrahedral intermediates of substrates (due to asymmetric binding interactions in the active site). The sufoximine oxygen would then correspond to the hydroxo nucleophile and the stereochemistry of attack can be defined (Mock and Zhang 1991).

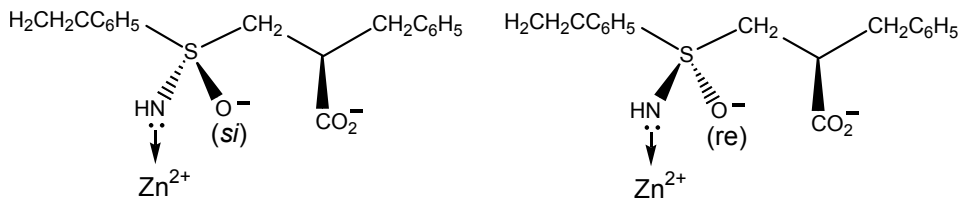


Fig. 33 Diasteromeric sulfoximine inhibitors of CPA.

All the mechanistic schemes discussed above are ambiguous with regard to proton transfer steps, other transient intermediates, and perhaps rate-limiting conformational changes which are not so easily accessible by experiment; however, computational studies (where the nucleophilic water is generally set as being water-bound) have suggested additional subtleties although these tend to be energetically small and in general a rate-limiting step of nucleophilic hydroxide attack with an activation barrier of about ~20 kcal mol^{-1} is identified (Abashkin, Burt et al. 1996; Alvarez-Santos, Gonzalez-Lafont et al. 1998).

Mechanism of PfuCP

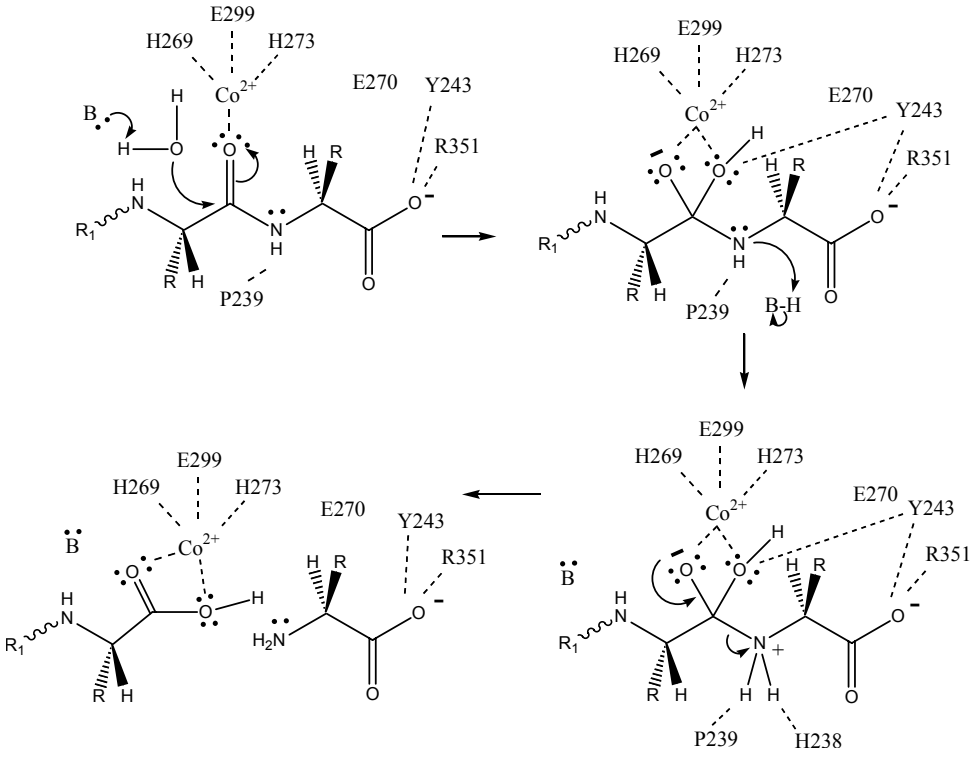


Fig. 34 Generalized reaction pathway of PfuCP.

As with most metalloproteases, the crystal structure of PfuCP (Ch. 3) is consistent with a catalytic metal acting as a Lewis acid to polarize a substrate carbonyl and/or a metal-bound water. An anhydride pathway, however, is not indicated since reactions carried out in the presence of 1 mM hydroxylamine do not trap any intermediates as determined by mass spectrometry, and the absence of esterase activity is at odds with experiments on CPA that demonstrate anhydride intermediates only in the presence of ester substrates.

Although there is a conserved active site histidine (His 298), it is located too far away and situated too 'low' relative to the plane of metal ligands to act as a proton

shuttle, and in this regard, Glu270 or the substrate C-terminus may deprotonate water to generate the activated nucleophile. Modeling of a decapeptide into the active site groove in a similar orientation relative to active site residues in thermolysin and astacin reveals that water-activation via Glu270 would yield *re* face attack, while deprotonation of water by the substrate carboxylate would yield *si* face attack. Studies similar to those for thermolysin with acylcyanamide substrates and sulfoximine inhibitors might help clarify the nature of the proton shuttle and resolve the ambiguity between metal-activated water attack or reverse protonation (Mock 1998).

Incidentally, there has been much debate over whether or not the mechanisms are the same for ester and peptide hydrolysis in CPA, as there is ambiguous evidence from chemical modification (Coleman, Pulido et al. 1966; Davies, Auld et al. 1968; Davies, Riordan et al. 1968; Glovsky, Hall et al. 1972; Bunting and Myers 1975; Sebastian and Lo 1978; Sebastian, Hinks et al. 1987) and other kinetic studies (Galdes, Auld et al. 1986; Auld and Vallee 1987). As PfuCP has no apparent esterase activity these issues could not be explored; however, the absence of activity does suggest that the mechanisms are intrinsically different possibly involving the mode of substrate binding rather than being simply a difference in rate-determining step.

The steady-state kinetics of PfuCP (Ch. 4) revealed that aside from the mechanics of peptide cleavage, allostery in terms of metal and substrate interactions (between inorganic and organic ions) is important and that the loose association of ligands for rapid binding and release might play a role in regulation. This aspect of loose binding might also have a more subtle function in catalysis by modulating reaction barriers, and this possibility is discussed below.

Theories of Enzyme Catalysis

There appear to be many ways in which enzymes can achieve their remarkable rate enhancement over uncatalyzed reactions.[15] Conventional thought is that enzymes can accelerate reactions by

1) bringing together and orienting reactants in productive geometries to lower both the activation entropy and enthalpy (template effect / preorganization of dipoles),
2) inducing strain in substrates in order to weaken bonds (trading bond enthalpy for binding enthalpy),
3) stabilizing the transition state and lowering its energy,
4) stabilizing and increasing the concentration of key intermediates.

Strategies 1-3 reduce the intrinsic barrier (E_a^o) and hence overall barrier to reaction and in the Polanyi or Marcus formalisms this amounts to reducing the slopes of the potential energy surfaces for reactants and/or products (Fig. 35 and 36).

[15] Rate enhancement = rate of enzyme catalyzed reaction / rate of uncatalyzed (solution) reaction

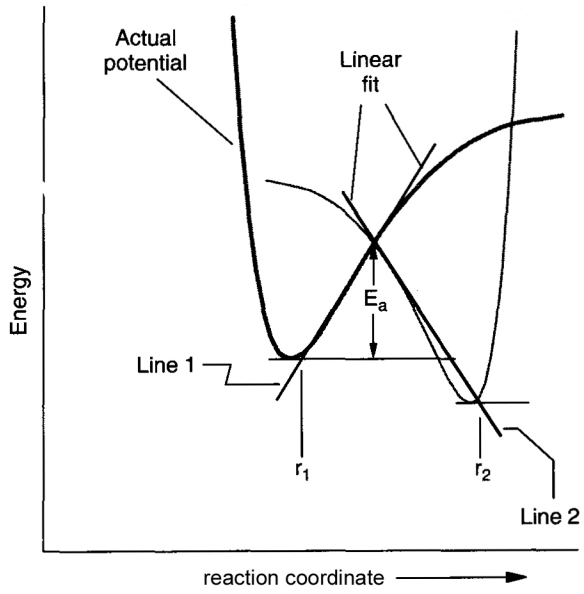


Fig. 35 Polanyi linear approximation to reactant and product potentials (Adapted from (Masel 2001)).

Polanyi Relationship: $E_a = E_a^o + \gamma_P \Delta G_r$

where $\gamma_P = \dfrac{|Sl_1|}{|Sl_1| + |Sl_2|}$

and $E_a^o = \left(\dfrac{|Sl_1||Sl_2|}{|Sl_1| + |Sl_2|} \right)(r_2 - r_1)$

and ΔG_r is the reaction free energy, Sl_1 and Sl_2 the tangents to the reactant and product potential surfaces, respectively, r_1 and r_2 the normalized reaction coordinates of the reactants and products, E_a the overall activation barrier, γ_P the transfer coefficient, and E_a^o the intrinsic barrier.

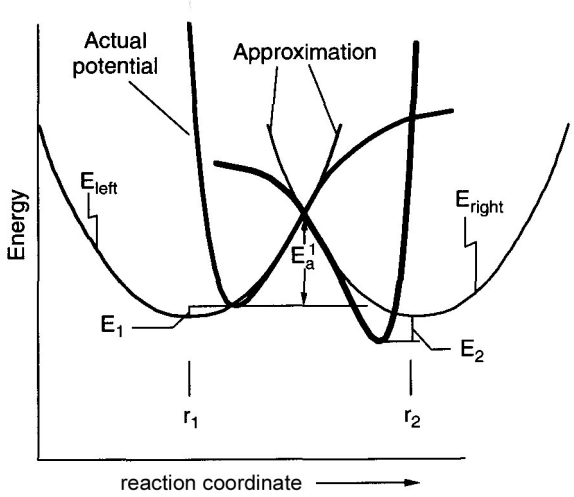


Fig. 36 Marcus parabolic approximation to reactant and product potentials (Adapted from (Masel 2001)).

Marcus relation: $$E_a = \left(1 + \frac{\Delta G_r}{4E_a^o}\right)^2 E_a^o$$

where $$E_a^o = \frac{(r_2 - r_1)^2}{4} SS_1$$

and E_a^o is the intrinsic barrier, ΔG_r the free energy of reaction, r_1 and r_2 the normalized reaction coordinates of the reactants and products, respectively, and SS_1 a fitting factor ($E_a = E_a^1$ since work terms omitted)

The rate is then proportional to the concentration of reactants (reactive intermediates) raised to their stoichiometric coefficients and $e^{-Ea/kT}$ by the conventional formulation of Transition State Theory and the Arrhenius Relation (where k is the Boltzmann constant, and T the absolute temperature).

Preorganization of dipoles in the active site of an enzyme lowers both entropic and enthalpic contributions to the activation barrier which would otherwise arise in an

uncatalyzed solvent environment from slow re-organization of solvation cages and necssary rearrangements in substrate conformation throughout the reaction coordinate (Cannon, Singleton et al. 1996). In effect, the 'fixed' disposition of electrostatic charges and dipoles in the protein matrix and the forced restriction on substrate conformational possibilities within the confines of the active site 'pre-pay' some of these energetic costs. This shape-selective nature of enzyme-substrate binding can also stabilize strained conformations of substrates or transition state structures, lowering the enthalpic barrier to reaction through a 'rack' mechanism. Stabilizing reactant conformations close to the transition state of the rate-determining step can also increase the concentration of productive species and hence the rate.

Accurate and quantitative predictions of activation barriers are still limited to atom transfers in the gas-phase due to the difficulty in calculating intrinsic barriers for macromolecular systems which have complex and unknown dependence on substrate and active site geometries as well as other phenomena such as energy transfer. Nevertheless, qualitative analysis of more complicated reactions such as enzymatic hydrolysis can serve as a guide for future studies when more advanced modeling methods become available.

Implications for Co^{2+} activated Zn^{2+} metallohydrolases[16]

[16] The absence of strong evidence for an anhydride mechanism and the clear indications that the catalytic metal in metalloproteases does indeed ligate the substrate carbonyl suggest that for the purposes of discussion the former can be ignored and the latter assumed. The remaining ambiguity in the mechanism would then be whether or not the attacking aquo species is bound to the catalytic metal. In any case, the identity and nature of the catalytic metal has a direct effect on the rate of reaction.

A long-standing and unanswered question for metallohydrolases, is exactly how metal ions contribute to rate acceleration. In some cases, the *apo*enzyme has been reported to bind substrate without metal (Christianson and Lipscomb 1989) while in others the catalytic metal is loosely bound (Ghosh, Grunden et al. 1998; Green, Ginsburg et al. 1991; Ferretti, Luchinat et al. 1995; Wingfield, Graber et al. 1989; Cheng, Ramakrishnan et al. 1999), calling into question the role of the metal as a template to bring reactants together. Replacement of the native catalytic metal generally leads to minimal disturbance of either the active site residues or the global structure (carbonic anhydrase, (Hunt, Ahmed et al. 1999); CPA, (Feinberg, Greenblatt et al. 1993); thermolysin, (Holland, Hausrath et al. 1995); astacin, (Gomisruth, Grams et al. 1994)) suggesting no dramatic effect in orienting dipoles. Similarities in the binding disposition of reaction coordinate analogs among the different metalloforms of CPA has also been taken as to suggest that the differences in activity are not caused by altered modes of substrate binding and hence differences in induced strain or substrate orientation (Feinberg, Greenblatt et al. 1993). The apparent minimal effect of metal on the protein matrix, and evidence to the contrary (that it may in fact be the matrix which is constraining the energy of the metal and the disposition of its coordination sphere through 'entasis' (Vallee, Rupley et al. 1960)), suggest that there are intrinsic, electronic properties of the metal that influence the rate of reaction by perhaps stabilizing transition states or key intermediates.

A noteworthy and common observation is the rate-acceleration when a native Zn^{2+} is replaced by Co^{2+}. Both ions are the same size (74 pm, Pauling radii) with similar polarizabilities yet Co^{2+}-enzymes can be up to 223 % more active (Davies, Riordan et al.

1968) and to date, no explanation has been proposed. The following discussion examines this issue with reference to current understanding of pertinent bioinorganic chemistry and introduces some new hypotheses which may have universal relevance to how metals contribute to rate enhancement in metalloenzymes.

If the aquo nucleophile in the metalloprotease is and remains bound to metal ion throughout the reaction coordinate, then the change from 4 to 5 coordination upon binding substrate may be slightly more favorable for Co^{2+} over Zn^{2+} due to a slight increase in ligand field stabilization energy (LFSE) for Co^{2+}. This would argue for Co^{2+} as a better template; however, this effect is likely to be small and can estimated to be less than ~ 7 kcal mol^{-1} (the approximate difference in LFSE due to a switch from pure tetrahedral to octahedral Co^{2+}) (Trachtman, Markham et al. 2001; Orgel 1960); Zn^{2+} has no such ligand field stabilization effects due to its filled d-shells. If the mechanism is reverse protonation, the issue of changing coordination is even less relevant as the interaction of substrate with metal amounts to simple water displacement, and in this regard, Zn^{2+} appears to exchange ligands more readily (experimental water exchange rates for Zn^{2+} and Co^{2+} are respectively 2 $x10^7$ and 3 x 10^6 s^{-1} (Lippard 1994) and the activation enthalpies for water exchange are on the order of 10 kcal mol^{-1} (Hunt, Ahmed et al. 1999)). This however would not be in line with the observation than Co^{2+} yields the better catalyst.

The stronger Lewis acidity of Zn^{2+} (higher charge density (z/r) and ionization enthalpy) compared with Co^{2+} would presumably polarize and strain the substrate carbonyl (and bound water) to a greater degree prior to nucleophilic attack. This effect is

embodied in the Irving-Williams series which also describes trends in metal complexation constants: $Ca^{2+} < Mg^{2+} < Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$. A stronger Lewis acid would pull more of electron density in the carbonyl toward the coordinated oxygen, reducing the contribution of Pauli repulsions to the activation barrier as the nucleophilic hydroxide attacks. However, this is again at odds with the higher activity for the weaker Lewis acid, Co^{2+}.

On the basis of Lewis acidity or electrostatics alone, Zn^{2+} should also be more effective at stabilizing the electron rich, gem-diolate transition state, but again this cannot explain the higher activity for Co^{2+}.

Considering a water-bound nucleophile, if the metal serves to stabilize and increase the concentration of key intermediates, Zn^{2+}-hydroxide is expected to be a more abundant nucleophile than Co^{2+}-hydroxide on grounds of electrostatics (for comparison, the pKa for the free aquo ions of Zn^{2+} and Co^{2+} are 9.65 and 8.96, respectively (Baes and Mesmer 1976)). This also does not match with the higher activity seen for Co^{2+} assuming that the rate-determining step is attack of the aquo species[17].

[17] If other steps such as substrate and product binding and release were rate-limiting, Zn^{2+} would be expected to be the better catalyst due to its ligand exchange properties. The issue of a rate-limiting conformational change is discussed below.

However, in a reverse protonation mechanism a more potent Lewis acid would yield a higher concentration of catalytically incompetent metal-hydroxide species[18] which is more difficult to displace by the scissile carbonyl. This would be consistent with the observed higher activity of Co^{2+} enzyme which is predicted to have about 5 times less hydroxide bound in which case, the mechanism of rate acceleration of Co^{2+} over Zn^{2+} would be destabilization of a non-productive side reaction. However, calculations of the enthalpies of both water and monohydroxide dissociation indicate that these values are very similar between Co^{2+} and Zn^{2+} (enthalpies of water dissociation for Co^{2+} and Zn^{2+} respectively, 96.6 and 94.7 kcal mol^{-1} (Trachtman, Markham et al. 1998); enthalpies of hydroxide dissociation for Co^{2+} and Zn^{2+} respectively, 416.4 and 421.1 kcal mol^{-1}, (Trachtman, Markham et al. 2001)) disqualifying the above argument for significant rate enhancement of Co^{2+} over Zn^{2+} as it pertains to reverse protonation.

Thus, the preponderance of evidence as it relates to conventional theories of enzyme catalysis seem to suggest that by all accounts, Zn^{2+} and not Co^{2+}, should give the larger rate acceleration. Some new ideas are proposed here to address this issue.

[18] The relative pKa values of the free aquo ions of Co^{2+} and Zn^{2+} (hexaaquo) are representative of the relative pKa's of enzyme bound metal hydrates so that the difference in pKa between the free Co^{2+} and Zn^{2+} aquo complexes (first ionization, pKa = 9.65 and 8.96 respectively) gives the ratio of fractional ionization of hexaaquo Co^{2+} relative to hexaaquo Zn^{2+} (Appendix C).

The less tightly bound valence electrons of Co^{2+} might participate in more covalent interactions with the carboxamide antibonding orbitals, weakening the scissile carbonyl and lowering the activation barrier. Analogous quantum effects are known for π-backbonding between filled metal d-orbitals and the 2 π* antibonding orbital of CO where the strength of π donation was correlated to the ionization potential of the metal d-electrons (Lupinetti, Fau et al. 1997).

The weaker binding of Co^{2+} implies weaker force constants that hold the metal in place which in turn increases the 'degrees of freedom' the metal can explore. Peptide substrates have additional degrees of freedom in their extended peptide chain which likely give rise to specific conformations and interactions with the protein matrix lending a certain ruggedness to the reactant potential energy surface (PES) with perhaps equilibrium stabilization of non-productive configurations; incidentally, if the ruggedness is not characterized by fast relaxations, an additional complication would be that the rate might be stochastically, temporally gated. In either case, additional mobility in the metal can compensate for ruggedness by flattening the reactant PES. If the transition state is also described as an ensemble of near degenerate states rather than as a canonical structure, greater positional freedom for the metal can both broadening the transition state surface (TSS) (Fig. 37) and flatten any ruggedness from transient non-productive configurations (similar to the ground state).

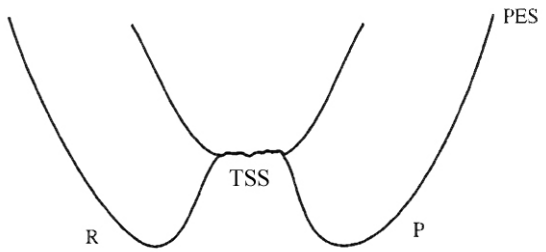


Fig. 37 A represenation of a broadened transition state hypersurface that reduces the activation barrier (adapted from (Ma, Kumar et al. 2000)).

In statistical mechanical terms, the partition function of the transition state will be larger for the Co^{2+} enzyme than for Zn^{2+}. However, a looser metal would also increase the degrees of freedom for the reactants and hence offset any net gain in the activation entropy which is achieved by increasing the configuration space of the transition state ensemble alone. Since the transition state for the rate determining attack of hydroxide is associative, this suggests that rather than increasing activation entropy, a looser metal may act to decrease the obligatory loss in entropy, thus reducing the intrinsic barrier. From another perspective, added mobility in the metal might also allow for a greater number of 'near-attack-conformers' (NACs) in the ground state by flattening the PES with the effect of decreasing the enthalpy of activation by allowing the reaction coordinates of reactant and product to move closer together. In some cases, catalysis is thought to arise primarily if not solely from ground state effects rather than by other mechanisms such as transition state stabilization (Bruice and Benkovic 2000). For both ground and transition state structures, increasing the mobility of a metal can compensate not only for ruggedness due to conformations in a given substrate but also for different amounts of ruggedness between different protein sequences, in effect introducing a non-

specificity which balances the specifity engendered in other mechanisms such as size-selection of substrates or particular enzyme-substrate contacts.

A less tightly bound Co^{2+} would also contribute to a smaller loss in vibrational entropy as the carbonyl becomes more polarized towards the metal. The greater Lewis acidity of Zn^{2+} generates a larger polarization of carbonyl electron density which leads to a stronger metal-oxygen bond but also a stronger C-O electrostatic attraction within the carbonyl and hence higher force constants and less vibrational entropy in the transition state.

A less tightly confined metal might also couple vibrational modes more easily to the substrate (and perhaps protein matrix) for intramolecular energy transfer and in switch more easily between modes as the reaction coordinate progresses since it experiences more of an inhomogeneous and anharmonic electrostatic environment than a pseudoharmonically bound metal. This ease in moderating energy transfer throughout the reaction coordinate might arise in the case where nucleophilic water and carbonyl are simulataneously metal bound such that the added mobility of Co^{2+} might work better at alternately weakening the OH bond of bound water for generation of the nucleophile and polarizing the scissile carbonyl for attack, two competing processes (since they use the same Lewis acid) that might otherwise be less favorable for a more tightly constrained Zn^{2+}. Further along the reaction coordinate (Fig. 15), electron density needs to shift to the scissile amide nitrogen for protonation before the stretching and breaking of the peptide bond, and again a less 'electrostrictive' metal might be better at modulating these steps where energy needs to be redirected into new vibrational modes. Also, a rate-determining conformational change might occur more easily with Co^{2+} if the greater

electrostrictive effect of Zn^{2+} on the protein matrix reduces the degrees of freedom for reaching a transition state configuration; and further, since the active site is large enough to contain many waters, the smaller electrostrictive effect of Co^{2+} might also lower any contribution to the reaction barrier from solvent re-organization[19].

In contrast, the cobalt isoform of carbonic anhydrase is approximately 80 % as active as the native Zn^{2+} enzyme despite the fact that Co^{2+} binds with lower affinity (Tu and Silverman 1985). One explanation might be that the smaller, rigid substrate in CA (CO_2 / bicarbonate) has a smaller TSS to begin with which cannot and does not significantly change throughout the reaction coordinate leaving Lewis acid polarization (electrostatics) as the dominant effect.

The observation can also be made that Mn^{2+} binds to CPA less tightly than Co^{2+} and has faster exchange kinetics (rate of water exchange is 2×10^7 s^{-1}) yet its activity is 8 % that of the Zn^{2+} metalloform (Table 4). Mn^{2+} also has no ligand field stabilization energy effects due to the half-filled d-shells. In addition to its weaker Lewis acidity (compared with either Co^{2+} or Zn^{2+}), weaker ability to polarize and stabilize species along the reaction coordinate, and the apparent displacement of the coordinated water towards Glu270 for a tighter hydrogen bond (Feinberg, Greenblatt et al. 1993), it is possible that the larger size of the ion (80 pm) cancels the effects of added mobility by narrowing the TSS in an as yet unknown manner. Enzyme metal-binding sites in the ground state are known to have specific preferences for metal ions through size-selection

[19] Although direct coupling of catalysis to protein dynamics is generally considered to be minimal on the grounds that the energy will be dissipated into a large number of modes before it can reach the critical reactive modes, direct coupling to a metal co-factor has been implicated in the 'blocking' of energy transfer

and the disposition of second-shell residues attached to or near the metal ligands (Hunt, Ahmed et al. 1999; Lesburg and Christianson 1995; Kiefer, Paterno et al. 1995), and it is entirely plausible that this selectivity extends into the transition states.

Aside from Zn^{2+} and Co^{2+} metalloproteases, this concept of looser metals literally translating into looser transition states might also explain the behavior of arginase (Christianson 1997), which has full activity with two Mn^{2+} but which exhibits half the activity when the one labile Mn^{2+} is lost. In this case, the TSS would be reduced by the absence of an additional metal cofactor.

As a final note, the looseness of the metal may arise from the nature of the protein matrix itself wherein second shell residues modulate metal affinity through networks of hydrogen bonds or hydrophobic contacts or even by changing the nature of the local dielectric through the sizes and shapes of cavities. The dynamics of the protein may then further modulate the coordination distances of metal ligands and hence affect not only their Lewis acidity but their effect on the TSS.

---

to the protein matrix in order to keep substrates 'hot' by localizing the energy transfer pathways (Solc 1999).

Chapter 6

Biotechnological Applications and C-terminal Protein Sequencing

Abstract

The capability of PfuCP to sequence peptides and proteins from their C-terminii at elevated temperatures was evaluated. The speed of the procedure is remarkable; however, the substrate size amenable to sequencing appears to be restricted by the shape of the active site cavity. This represents the first report of protein sequencing at high temperatures.

Introduction

Enzymes from extremophiles have found use in demanding applications where their resistance to harsh chemical and thermal environments, born from their native niche, enables them to carry out reactions under conditions that would otherwise denature their mesophilic counterparts. Hyperthermophilic enzymes, glycosidases and DNA polymerases, have now become workhorse enzymes in sugar conversion (Laderman, Davis et al. 1993) and PCR (polymerase chain reaction) (Lundberg, Shoemaker et al. 1991).

Various uses have been proposed for carboxypeptidases and their reverse reactions (defouling in bioremediation, debittering of protein hydrolysates (Umetsu, Matsuoka et al. 1983), activation of pro-drugs in antibody-directed (ADEPT) therapeutics (Rowsell, Pauptit et al. 1997)), attachment of chemical groups to polypeptides by

transpeptidation and other organic transformations including (trans)esterification (Nashed and Kaiser 1986; Owusu and Cowan 1991)).

Protein sequencing, however, remains a largely undeveloped application since success has proven sporadic with mesophilic CPs, limited to relatively small peptides and a few anomalous proteins (Thiede, Salnikow et al. 1997). A (hyper)thermophilic CP has the potential to enable sequencing at elevated temperatures or under otherwise denaturing conditions (chaotropes, organic solvents) which would allow more extensive digestion of larger polypeptides and intact proteins by removing any interfering secondary (higher order) structure. The benefits to proteomics are clear since current techniques such as peptide mass fingerprinting require multiple steps of fragmentation, chromatographic separation, and mass spectrometric measurement followed by database query and statistical analysis in order to identify a single protein, which also requires that the sequence be already known. A *de novo* technique that yields more extensive sequence information on larger, perhaps intact proteins or more sizeable fragments, would lead to higher-throughput with attendant reductions in cost and time without requiring a complete catalog of the relevant proteins. Applications of a more efficient and effective protein identification technique can be found in quality control for purified or synthetic peptides/derivatives/peptidomimetics, screening for protein markers of disease or metabolic response to drugs and environmental factors (e.g. agrobiotech). If specificities for post-transcriptional and post-translational modifications such as phosphorylation can be engineering in, information could also be obtained that would not be evident from DNA expression data or genomics methods.

Edman sequencing (chemical) is useful for N-terminal characterization but fails when the amino terminus is chemically blocked. Endoproteolytic fragmentation is the standard solve but an effective C-terminal method would provide an alternative. Chemical C-terminal methods analogous to Edman degradation have been developed but their efficiency is much less and limited to a few callable residues (Hardeman, Samyn et al. 1998). Fragmentation methods are not sufficiently robust for *de novo* sequencing and generally less than 10 residues can be sequenced on peptides less than 25 kDa in size.

Enzymatic methods have been developed which monitor the time-lapsed appearance of released amino acids by chemical (ninhydrin) methods, but a more powerful approach is ladder sequencing wherein a carboxypeptidase is used to generate a set of differentially cleaved peptides that can be visualized in a mass spectrum (Chait, Wang et al. 1993). Mass differences between adjacent peaks correspond to the molecular masses of individual amino acids that have been released. The potential advantages of enzymatic sequencing are many:

ii) Longer sequence reads (the amount of sequence information is limited only by how far the enzyme can cut and the resolution of the mass spectrometer);

iii) Speed (laddering is finished on the order of a few minutes and ease of interpretation for the mass spectra and potential for automation can give a higher throughput than other methods; by comparison Edman sequencing is on the order of hours);

iv) Cost (an installed base of mass spectrometers which is used for other purposes can be adapted for sequencing).

Difficulties that have limited the usefulness of enzymatic sequencing include a narrow specificity (particularly for the C-terminal amino acid) and a resistance of native protein substrates to digestion at mesophilic temperatures, presumably because of higher order structure. Hyperthermophilic CPs, which have less specificity and which are stable against extreme chemical and thermal conditions, have the potential to circumvent these problems.

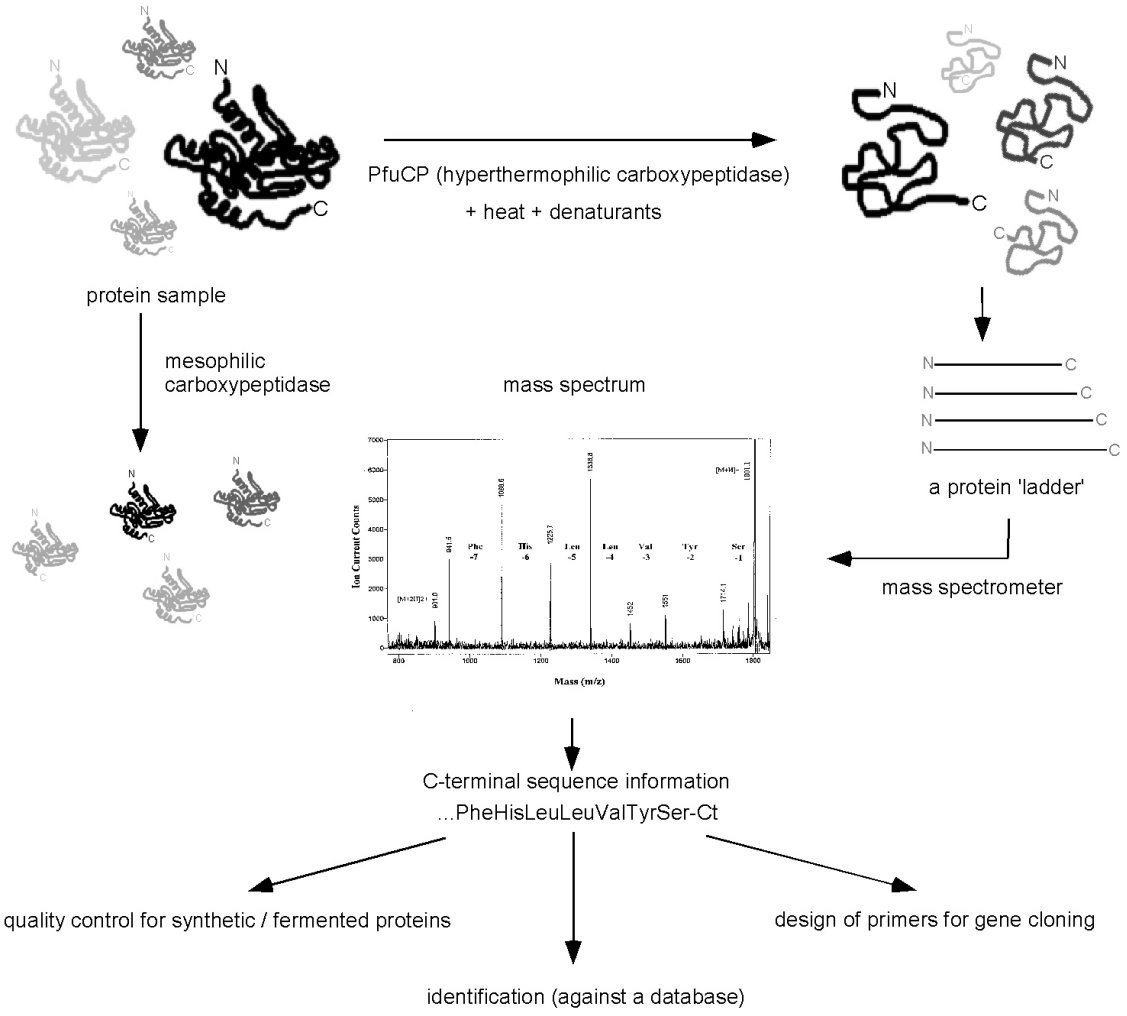


Fig. 38 C-terminal sequencing methodology for a hyperthermophilic carboxypeptidase.

Table 10 -C-terminal Sequencing Trials

Substrate	MW (kDa)	Sequence
N-acetyl renin substrate	1.801	NAc-DRVYIHP**FHLLVYS**
Protein Kinase C substrate	1.271	PLSRTLS**VAAKK**
Fibroblast Growth Factor Basic Fragment 106-120	1.963	YRSRKYSSQ**YYVALKR**
TyrOMe36 Neuropeptide Y 18-36	2.458	ARYTSALR**HYINLITRQRY**
Fibronectin 1377-1388	1.357	HSRNSIT**LTNLT**
Bradykinin	1.06	RPPGFSP**FR**
β-endorphin 1-27	3.022	YGGFMTSEKSQTPLVTLFKNAIIKNA**Y**

NB: residues in bold were cleaved by PfuCP

Several test peptides (Table 10) were sequenced by PfuCP, but all substrates longer than approximately 30 residues showed no sign of digestion despite exhaustive trials with denaturing conditions (e.g. cytochrome c, 11.7 kDa; myoglobin, 18.8 kDa; BSA, 66.3 kDa). It is probable, that the deep active site groove size-limits the substrates that can be cut, and that the C-terminii of longer polypeptides simply cannot be accommodated. C-terminal sequencing of a typical peptide such as N-acetyl-renin substrate at 80 ^{o}C involved a one-minute digest followed by MALDI-TOF MS analysis of the reaction quenched on ice (Cheng et al. 1999).

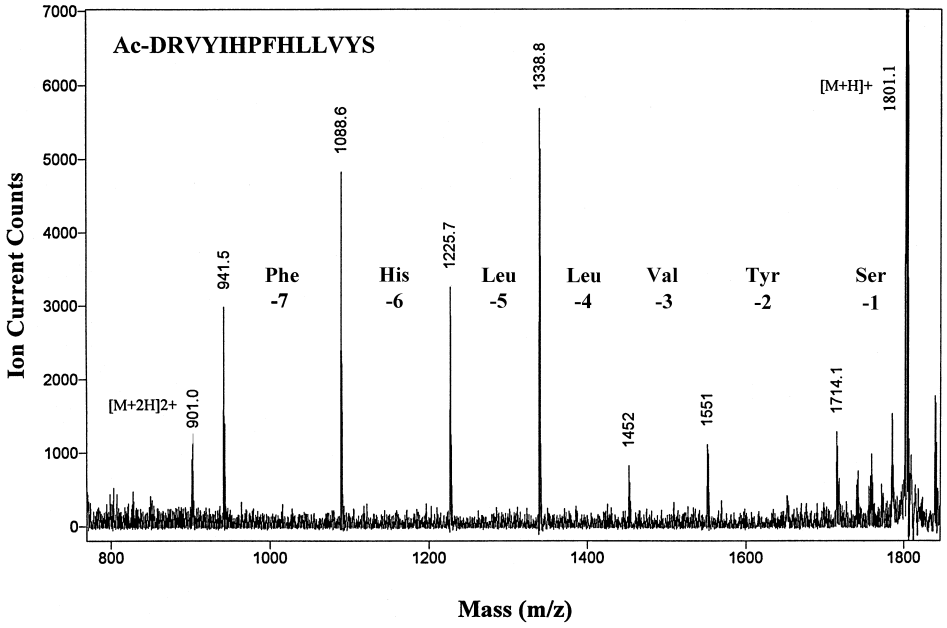


Fig. 39 Representative mass spectrum of a standard C-terminal sequencing digest of N-acetyl-renin substrate by PfuCP (5 mM KMes pH 6.5, 50 µM N-acetyl-renin substrate, 0.4 mM $CoCl_2$, and 150 nM *apo*PfuCP, total volume 5 µL, digest time 1' at 80 ^oC followed by quenching on ice and addition of 0.1% TFA prior to MALDI-TOF MS).

Chapter 7

Conclusions and Future Directions

The carboxypeptidase, PfuCP, has traits which suggest that it hails from ancient roots on the evolutionary tree. The catalytic metal binding motif (HEXXH) is better known to endoproteases rather than carboxypeptidases and the broad substrate specificity and a possible requirement for cobalt (more prevalent in primordial history) all hail back to earlier biogeochemical times. Analysis suggests (Ch. 5) that if the native metal is indeed Co^{2+}, the evolutionary trend towards Zn^{2+} in metallohydrolases may not represent a drive towards stronger Lewis acidity (Mock 1998), but a trade off between a loss of intrinsic catalytic ability balanced by the gain of a more 'stable' catalytic entity (longer half-life for an active holoenzyme and immunity to redox behavior), hallmarks of a more specialized role in multicellular, aerobic organisms with increasingly complex metabolic pathways. As archaea are thought to have evolved slowly, PfuCP takes carboxypeptidases a step closer to a progenitor protease that may have had multiple activities and functions.

Various experimental observations have suggested that the details of purification can significantly impact the behavior of the purified enzyme as compared with *in vivo* conditions. Investigations into native PfuCP prior to loss of stabilizing metal(s) should determine the cause of the apparent thermostability with regards to both the dimer structure and activity at high temperature as well as the kinetic nature of the binding for the native catalytic metal. It will also be interesting to see how the nature of the inhibition for Zn^{2+} or other metals might differ in a metastable holoPfuCP. Continued fermentation (D'souza and Holz 1999) and refolding experiments in the presence of different metals may give an indication as to what the native metal might be as well as

perhaps reducing or eliminating inclusion body formation if certain metals are required for proper folding.

The complicated steady-state kinetics of PfuCP may be related to a metabolic role which requires sensitivity to fluctuations in ambient peptide and metal concentrations as well as an integrated temperature response. Alternate and perhaps longer substrates should inform on the nature of the both the ensemble of active species and the apparent allostery between metals and substrate observed in the steady-state kinetics and further activity studies in the presence of other cellular components and nonprotein ligands may begin to unravel the niche this enzyme occupies in *Pyrococcus furiosus*. If the recombinant free of His-Tag proves to be monomeric at all temperatures (as suggested by SDS-PAGE on the partially purified recombinant), it may be possible to dissect the steady state kinetics into contributions from inter- as well as intrasubunit interactions, distinguishing dimer from monomeric. The current data however is consistent with monomeric PfuCP at higher temperatures since the enzyme in the steady-state assays was at low concentration (6×10^{-10} M) and the enzyme corresponds to a form of native enzyme that been treated with DTT to remove native dimer-stabilizing metal(s).

Substrates with less kinetic complications should also prove useful for pre-steady-state turnover experiments to better characterize the chemical steps and additional studies akin to those in characterizing thermolysin (Mock 1998) may clarify the relevance of the reverse protonation mechanism if any to the hydrolytic mechanism of PfuCP.

The structure of PfuCP is unlike that of the majority of known proteases and the ability to bind many substrates and ligands in its spacious active site may in part mitigate the effects of mass transfer by ‘collecting’ compatible substrates and ligands into a

'sticky' groove prior to catalysis. The overall positive cooperative response in activity with the obligatory binding of multiple substrates for full activity suggests a functional role for the active site groove in addition to simple size-selection.

It is apparent that the response of enzymes to different metal ions is separated into two issues: inhibition and other allosteric effects, and the intrinsic properties of the metal.

PfuCP is inhibited by at least two Zn^{2+} in an uncompetitive manner most likely due the first Zn^{2+} binding to an inhibitory site homologous to other HEXXH proteases and then bridging to a catalytic Co^{2+}. Binding to this inhibitory site is presumably prevented in the native enzyme by binding of alternate stabilizing metal(s) which are removed by prolonged incubation of fully purified PfuCP with DTT. The stabilizing metal(s) might in fact be the native catalytic metal itself or metals bound to other sites which are effectively kept in a kinetic trap until release by DTT which occurs with a conformational change in the dimer.

A loose association of a catalytic Co^{2+} to the demetallated PfuCP has been demonstrated by reconstitution experiments that cannot incorporate a stoichiometric Co^{2+} and optical spectra that show an apparent octahedral coordination. A loose metal can reduce the free energy of activation by decreasing the loss in entropy from ground to transition states (broadening the transition state surface) and by decreasing the enthalpy barriers that arise from conformational rearrangements required to bring substrates closer to the transition state in effect allowing for more 'near-attack' conformers (Bruice and Benkovic 2000). This loose movement of a catalytic metal coupled to the overall dynamics of the protein matrix may in part explain the 'design flaw' in rigid model enzyme systems which limits their activity since greater stochastic movement of a protein

matrix and attached components enables a wider sampling of conformational space for both reactant and transition state species.

By extension, other hydrolases which act on larger, heterogeneous, and more dynamic substrates (proteins, DNA, lipids, and sugars) may use this positional flexibility of metals to fine-tune the reaction landscape and optimize both specificity and activity. Furthermore, in enzymes where the catalytically essential metal is loosely associated, the concept of entasis does not necessarily apply in that the metal is not fixed in a constrained ligand cage of high relative energy approaching the transition state configuration; rather pre-organization is restricted to the protein matrix.

Other aspects of Co^{2+} that may contribute to its rate enhancement over Zn^{2+} (Ch. 5) will need to be quantified with the appropriate experiments and computational studies before any definite conclusions can be drawn.

It will be interesting to see how more distant correlated motions of the protein matrix (essentially a variable dipolar and dielectric field) will couple to the localized metal properties in the catalytic metal site, and further, how these effects will differ between mesophiles and hyperthermophiles given the expanded configurational space of the latter and issues of thermoactivation.

Understanding the fundamentals of enzyme catalysis will enable the design of faster and better biocatalytic tools. With its broad specificity and rapid digestion, PfuCP has shown promise in this regard, and future improvements with additionally engineered capabilities are anticipated.

Appendix A -Cloning and Subcloning of PfuCP

Cloning from *P. furiosus* genomic DNA

Using standard molecular biology techniques (Maniatis and Sambrook 1989; Harwood 1996; Brown 2000; Tuan 1997), genomic DNA was prepared from cells fermented as in (Cheng, Ramakrishnan et al. 1999) and an Xba size-selected library created by PCR (Taq polymerase), ligation into a pET24a cloning vector (Bluescript), and transformation into *E. coli* BL21 DE3 lyse S cells. The PCR primers were constructed from 'reverse' translation of amino acid sequence tags obtained from N-terminal sequencing of native protein fragments allowing for redundancy in the nucleic acid sequence.

1) aa(a/g)ta(c/t)tcngc(a/g)tccat(a/g)tc(c/t)tc

2) atgga(a/g)gtntt(t/c)ca(a/g)aa(t/c)ga

3) a(a/g)(a/g)ttnac(a/g)aa(c/t)tcngg(a/g)tg

4) ga(a/g)ga(a/g)gtntt(t/c)ca(a/g)aa(t/c)ga(a/g)ac

Two clones which each contained one halve of the PfuCP gene were excised and ligated together before subsequent subcloning into the following expression vectors in *E. coli* TOP10F cells:

1) N-terminal His-Tag (Qiagen, pQE, T5 promoter)
2) C-terminal His-Tag (Qiagen, pQE, T5 promoter)
3) GST-tag (Invitrogen, T7 promoter)

Subcloning into *E. coli* PROTet.E133™ 6xHN

The PfuCP gene in expression vector 2) was subcloned through standard PCR techniques using Deep Vent™ polymerase (New England Biolabs). Primers 1) and 2) were constructed with overhangs with Not I and Sal I sites at either end:

1) ATA TGT CGA CAT GGA AGA GGT ATT CCA AAA CGA AA

2) ATA TGC GGC CGC ATT AGA GAT ATT TCT CCT TCA CCC AC

PCR was carried out for 30 cycles in the presence of 5 mM $MgSO_4$:

60s 95 ^{o}C melting

30s 55 ^{o}C annealing

90s 72 ^{o}C extension

final cycle 60 s 95 ^{o}C, 60s 60 ^{o}C, 5 min 72 ^{o}C then hold 4 ^{o}C

PCR product was isolated from 1% agarose and ProTet.E133 were cut with Sal I and Not I and ligated before transformation into competent BL cells. Positive transformants were selected on agar containing 34 μg/mL chloramphenicol and 50 μg/mL spectinomycin and the inserts verified by restriction analysis using Afl II and Not I which have unique sites within the recombinant vector.

Induction

Cells were grown at 37 ^{o}C to an A_{600} of 0.5 in a 15 L fermentor from 300 mL overnight cultures and induced with 1 μg/mL anhydrotetracycline. Cells were pelleted by centrifugation at 3000 RCF for 10' (wet mass 58.76 g) and lysed by incubation with 400 μg/mL lysozyme and 100 μg/mL DNase I at 37 ^{o}C in 200 mL 50 mM Tris pH 8 + 100 mM NaCl for 15' followed by 10' of sonication. The lysate was centrifuged at 3000 RCF for 15' yielding 16 g of crude inclusion bodies which were washed twice with 20 mM MOPS pH 8 + 100 mM NaCl before storage at –80 ^{o}C.

Appendix B -Refolding of rPfuCP from Inclusion Bodies

All methods used a FPLC LC-500 system and columns from Pharmacia. The Co^{2+} IMAC resin (TALON) was purchased from Clontech and packed into a 25 mL column. Urea solution A consists of 10 M urea + 20 mM MOPS pH 8 + 100 mM NaCl and urea was deionized by vacuum filtration through on IRA 900 cation and anion exchange resin (Sigma). Refolding buffer was 20 mM MOPS pH 7 + 100 mM NaCl and elution buffer had the same composition with an additional 200 mM imidazole.

Method 1: Simple Dilution (~ 9 h/cycle)

A 50 mL aliquot of concentrated inclusion body (80 mg/mL) in urea solution A was injected at 0.1 mL/min together with refolding buffer (+1% PEG 600 + 1 mM PMSF) at 9.9 mL/min into the rapid mixing attachment and outflow was then directed onto a Co^{2+} IMAC column in order to recover refolded PfuCP. The dilution factor of 100 reduced the concentration of urea and total protein to 100 mM and 800 μg/mL respectively. Elution of refolded PfuCP was accomplished with three column volumes of elution buffer at 5 mL/min.

Method 2: Gel filtration (3 h/cycle)

A concentrated 250 μL aliquot of inclusion body (40 mg/mL) in urea solution A was injected onto a 24 mL Superdex G500 column followed by elution at 0.4 mL/min with two column volumes of 20 mM MOPS pH 7 + 150 mM NaCl + 1 mM PMSF.

Method 3: On-resin (~1 h/cycle)

A 50 mL aliquot of concentrated inclusion body (80 mg/mL) in urea solution A was loaded onto a 25 mL Co^{2+} IMAC column at 10 mL/min then washed with one column volume urea solution A at 10 mL/min followed by buffer exchange with two column volumes of refolding buffer at 5 mL/min. Refolded protein was then eluted with two column volumes of elution buffer. Glycerol (10%),1 mM EDTA, and 1 mM PMSF were then added to the collected fractions prior to storage at –80 ^{o}C. To remove the His-Tag, aliquots of eluted rPfuCP were concentrated with a 30 kDa cutoff ultrafiltration unit (Microcon), exchanged twice with 20 mM MOPS pH 7.4 + 50 mM NaCl + 2 mM $CaCl_2$ and then incubated with 42 ng/mL enterokinase (Boehringer Mannheim) at 20^{o}C for 16 h. The digest was concentrated with a 30 kDa cutoff ultrafiltration unit (Microcon) and again passed through the Co^{2+} IMAC column to removed undigested recombinant. The pass through was then diluted into 45 mL of 50 mM Tris pH 8 + 10% glycerol + 1 mM PMSF + 1 mM EDTA and loaded at 1.75 mL/min onto a pre-packed 2 mL MonoQ ion exchange column. Recombinant was eluted with a salt gradient from 0-0.35 M NaCl at 0.5 mL/min and the peak which eluted at ~0.27 M NaCl was stored at –80 ^{o}C.

Appendix C -Assorted Calculations

Binding of Co^{2+} to the catalytic metal site of PfuCP

$$E\cdot Co^{2+} \overset{K_{Co^{2+}}}{\rightleftharpoons} E + Co^{2+}$$

$$K_{Co^{2+}} = \frac{[E][Co^{2+}]}{[E\cdot Co^{2+}]} \text{ or } [E\cdot Co^{2+}] = \frac{[E][Co^{2+}]}{K_{Co^{2+}}}$$

$$\text{fraction of enzyme bound} = \frac{[E\cdot Co^{2+}]}{[E]+[E\cdot Co^{2+}]} = \frac{[Co^{2+}]}{K_{Co^{2+}}+[Co^{2+}]}$$

such that the enzyme becomes saturated with Co^{2+} (all E in $E\cdot Co^{2+}$ form) once [Co2+] >> $K_{Co^{2+}}$

Relative Concentrations of Hydroxo Species Bound to Zn^{2+} and Co^{2+}

For two hexaaquo metal cations, Co^{2+} and Zn^{2+}, their Ka's and their ratio for the first ionization is as follows:

$$Ka_{Co^{2+}} = \frac{[H^+][Co(H_2O)_5OH^+]}{[Co(H_2O)_6{}^{2+}]}$$

$$Ka_{Zn^{2+}} = \frac{[H^+][Zn(H_2O)_5OH^+]}{[Zn(H_2O)_6{}^{2+}]}$$

$$\frac{Ka_{Co^{2+}}}{Ka_{Zn^{2+}}} = \frac{10^{-pKa_{Co^{2+}}}}{10^{-pKa_{Zn^{2+}}}} = 10^{pKa_{Zn^{2+}} - pKa_{Co^{2+}}} = 10^{8.96-9.65} = 0.20 = (5)^{-1}$$

Since the fraction of hexaaquo metal (Me^{2+}) ionized to monohydroxide is

$$= \frac{[Me(H_2O)_5OH^+]}{[Me(H_2O)_6{}^{2+}] + [Me(H_2O)_5OH^+]}$$

which $\sim \frac{[Me(H_2O)_5OH^+]}{[Me(H_2O)_6{}^{2+}]}$ when $[Me(H_2O)_5OH^+]$ is small,

the ratio of fractionally ionized hexaaquo cations is $\sim \frac{Ka_{Co^{2+}}}{Ka_{Zn^{2+}}}$

Therefore, hexaaquo Zn^{2+} dissociates to give approximately five times the population of hydroxide bound metal as Co^{2+}.

Experimental pKa values are from (Baes and Mesmer 1976).

Appendix D –EPR Methods

Procedures for standard activity assays and collection of optical, CD, fluorescence and mass spectra are as described previously (Cheng, Ramakrishnan et al. 1999).

EPR

Spectra were recorded on a Varian E-line Century X-band spectrometer at a frequency of 9.23 GHz, 5 mW and modulation of 1. Sample temperature was maintained at 4.2 K with an ESR-900 Oxford Instruments (Oxford, England) liquid helium cryostat. Samples were prepared aerobically or anaerobically by sealing 200 μL of a 100 μM protein solution in 0.1 M KMes pH 6.5 under an argon atmosphere in quartz EPR tubes. Samples prepared in tubes sealed with septa were rapidly frozen in liquid nitrogen prior to analysis.

Appendix E -Methods for Simulations of Steady-State Kinetics

Following the principles of the Simple Sequential Interaction Allosteric Method (Adair 1925; Pauling 1935), terms representing newly appearing enzyme species were added to the basic equations for dimer and monomer kinetics (Eqs. 3 and 4) as required to model the shifts in activity during 'titration' of metal or substrate in the M or S activity profiles respectively. The general equations so derived were fit separately but by similar methods using Microsoft Excel™ 2000. The term for a given species is simply the product of ligand concentrations raised to their stoichiometric contribution to the complex over the product of interaction factors between all possible pairs of ligands and binding constants for each ligand raised to their stoichiometric contribution (Segel 1975). Several assumptions were made however to simplify the simulations without compromising the conclusions that could be drawn from them:

1) binding events are fast relative to the chemical steps resulting in rapid pre-equilibria;
2) if possible, all multiplication factors of k_{cat} for different active enzyme species were set to 1 as a first approximation;
3) multiple, bound 'effector' ligands can be modeled as a single type of hypothetical 'dummy' ligand;
4) reasonable initial values for K_M and K_S can be estimated from the M and S profiles by inspection of the experimental ligand concentration ranges.

A further simplification included collapsing the interaction factors into the binding constants eg. the product of $K_{M'}$ and $K_{S'}$ in Term 1 includes gamma. If gamma is close to 1, the product of $K_{M'}$ and $K_{S'}$ is approximately (0.01 mM) x (1 mM) = 0.01 mM^2 or the product of the true binding constants if they were known.

Since not enough is currently known about either the chemical steps of the enzyme or the specific binding behavior of the ligands, rapid equilibrium conditions were assumed to a first approximation rather than modeling a more general steady-state scheme. In the absence of independently determined values for various binding events (e.g. from spectroscopic titration, equilibrium dialysis, isothermal calorimetry) initial values were estimated for $K_{M'}$ and $K_{S'}$ by inspection of the respective concentration ranges of ligands in the activity profiles (Fig. 21), and all subsequent parameters flowed from this initial assumption. Separate derivations of velocity equations for either M or S profiles better revealed the kinetic dependence on metal and substrate concentration alone, and comparison between both fits enabled refinement of the model and consolidation of terms into the general velocity equation Eq. 3.

The Simple Sequential Method invokes interaction factors to account for the homo- or heterotropic allosteric effect that a given ligand has on the binding of another, when both can be coupled in a thermodynamic cycle and equilibrium assumed. Pairwise interactions between all bound ligands have interaction factors that appear in the denominator of each term in the kinetic equations. An interaction factor greater than 1 represents reduced binding of a ligand in the presence of the other (negative cooperativity) while a value less than one represents increased affinity of the enzyme for

a ligand in the presence of another (positive cooperativity). This simple methodology can model allostery in any rapid equilibrium kinetic scheme of arbitrary complexity.

In the simulations reported here, the addition of ligands to form derivative terms from core species may in fact change the binding constants for ligands already bound; for example, in Terms 3, 4, and 5 of Tables 6 and 7, the binding constant for either the effector metals or the effector substrates increases with a greater background of substrate (seen as an increase in the product of the two constants for the 'dummy' effector metal and the effector substrate, where this product is simply a figure of merit which indicates that the binding properties of both effectors are being modeled as a unit). In fact, as background substrate increases, the dissociation constants for 'active' metal and substrate may be increasing ($K_{M'}$ and $K_{S'}$) which would have a similar effect to increasing the dummy constants on Term 3, its concentration, and the shape of the activity profile. Since it is not possible to know what is being bound less tightly, for the sake of simplicity, these allosteric effects are modeled as being confined to the effector ligands, such that $K_{M'}$ and $K_{S'}$ retain their nominal values.

Terms were added to the basic equations by modeling from the left of each activity profile (where species of lower 'ligand order' predominate) to the right (where higher order terms control the shape of the profile). Simulations within each set of profiles (M or S) were also carried out from low to high 'order'; for example, the set of M profiles was simulated beginning with the lowest background substrate (0.3 mM) to the M profile with the highest background substrate (10 mM). This allowed for cumulative effects of various species on the profile, kept the number of terms required to a minimum, and

allowed new features (terms) to arise when possible from addition of ligands to 'lower order' terms that already exist or by progressive changes in their binding constants.

For terms where the data required a certain order of ligand dependence (i.e. higher order exponent) in order to give the observed sigmoidicity, the presumably heterogeneous collection of effector ligands was modeled as several identical 'dummy' ligands binding to identical sites. This dummy dissociation constant gives in a sense the 'average' binding affinity provided the individual binding sites do not vary too much in their affinities and allosteric effects. Since individual dummy constants could not be ascertained, the binding behavior of effector metals and substrates on a given species were simulated in tandem as a unit whose overall binding behavior is represented in Tables 6 and 7 by a figure of merit, the linear product of the individual constants (which contain the appropriate interaction factors). The closer the interaction factors are to 1, the closer the simulated binding constants become to to their actual values.

Furthermore, since there is not enough information inherent in the data to assign the exact values of interaction factor values or individual binding constants, meaningful error could not be assigned to the determined parameters nor were corrections made for complexation of free Co^{2+} by buffer molecules.

References

Abashkin, Y. G., S. K. Burt, et al. (1996). Density Functional Modeling of Ligand-Metal Interactions and Enzymatic Activities in Metalloenzymes. *Metal-Ligand Interactions*. N. Russo and D. R. Salahub. Netherlands, Kluwer Academic Publishers**:** 1-22.

Adair, G. S. (1925). *J. Biol. Chem.* **63**: 529.

Adams, M. W. W. and R. M. Kelly (1998). "Finding and using hyperthermophilic enzymes." *Trends in Biotechnology* **16**: 329-332.

Alvarez-Santos, S., A. Gonzalez-Lafont, et al. (1998). "Theoretical study of the role of arginine 127 in the water-promoted mechanism of peptide cleavage by carboxypeptidase A." *New J. Chem.* **22**: 319-325.

Arndt, J. W., B. Hao, et al. (2002). "Crystal structure of a novel carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*." *Structure* **10**: 215-225.

Auld, D. S. and B. I. Vallee (1987). *Hydrolytic Enzymes*. A. Neuberger and K. Brocklehurst. New York, Elsevier**:** 201-255.

Baes, C. F. and R. E. Mesmer (1976). *The Hydrolysis of Cations*. New York, Wiley.

Bartlett, P. A., K. L. Spear, et al. (1981). *Biochem.* **21**: 1608-1611.

Berg, J. M. and D. L. Merkle (1989). *J. Am. Chem. Soc.* **111**: 3759-3761.

Bond, M. D., B. Holmquist, et al. (1986). *J. Inorg. Biochem.* **28**: 97-105.

Breslow, R., J. Chin, et al. (1983). *Proc. Nat. Acad. Sci. U.S.A.* **80**: 4585-4589.

Breslow, R. and D. L. Wernick (1977). *Proc. Nat. Acad. Sci. U.S.A.* **74**: 1303-1307.

Breyer, W. A. and B. W. Matthews (2001). *Prot. Sci.* **10**: 1699-1711.

Britt, B. M. and W. L. Peticolas (1992). *J. Am. Chem. Soc.* **114**: 5295-5303.

Brown, C. K., K. Madauss, et al. (2001). *Proc. Natl. Acad. Sci. U.S.A.* **98**: 3127-3132.

Brown, T. A., Ed. (2000). *Essential Molecular Biology: A Practical Approach*. Oxford, Oxford University Press.

Bruice, T. C. and S. J. Benkovic (2000). *Biochem.* **39**: 6267-6274.

Bunning, P., B. Holmquist, et al. (1983). *Biochem.* **22**: 103-110.

Bunting, J. W. and C. D. Myers (1975). *Can. J. Chem.* **53**: 1993-2004.

Burnstein, H. B., K. A. Walsh, et al. (1974). *Biochem.* **13**: 205-210.

Campbell, P. and N. T. Nashed (1982). *J. Am. Chem. Soc.* **104**: 5221-5226.

Cannon, W. R., S. F. Singleton, et al. (1996). *Nat. Struct. Biol.* **3**(10): 821-833.

Chait, B. T., R. Wang, et al. (1993). "Protein ladder sequencing." *Science* **262**: 89.

Cheng, T., V. Ramakrishnan, et al. (1999). "Purification and characterization of a cobalt-activated carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*." *Protein Science* **8**: 2474-2486.

Christianson, D. W. (1997). *Prog. Biophys. Mol. Biol.* **67**: 217-252.

Christianson, D. W., P. R. David, et al. (1987). *Proc. Nat. Acad. Sci. U.S.A.* **84**: 1512-1515.

Christianson, D. W. and W. N. Lipscomb (1986). *J. Am. Chem. Soc.* **108**: 4998-5003.

Christianson, D. W. and W. N. Lipscomb (1989). 22: 62-69.

Coleman, J. E., P. Pulido, et al. (1966). *Biochem.* **5**: 2019-2025.

Colombo, S., S. D'Auria, et al. (1992). "Purification and characterization of a thermostable carboxypeptidase from the extreme thermophilic archaebacterium *Sulfolobus solfataricus*." *Eur. J. Biochem.* **206**: 349-357.

Davies, R. C., D. S. Auld, et al. (1968). *Biochem. Biophys. Res. Commun.* **31**: 628-633.

Davies, R. C., J. F. Riordan, et al. (1968). *Biochem.* **7**: 1090-1099.

D'souza, V. M. and R. C. Holz (1999). *Biochem.* **38**: 11079-11085.

Feinberg, H., H. M. Greenblatt, et al. (1993). *J. Chem. Inf. Comput. Sci.* **33**: 501-516.

Feinberg, H., H. M. Greenblatt, et al. (1993). "Structural studies of the role of the active site metal in metalloenzymes*." J. Chem. Inf. Comput. Sci.* **33**: 501-516.

Ferretti, S., C. Luchinat, et al. (1995). "Polymetallic hydrolytic zinc enzymes -probing the site of nuclease P1 through cobalt(II) substitution." *Inorg. Chim. Acta* **234**: 9-11.

Fiala, G. and K. O. Stetter (1986). "*Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic arcahebacteria growing optimally at 100C." *Arch. Microbiol.* **145**: 56-61.

Frausto da Silva, J. J. R. and R. J. P. Williams (2001). *The biological chemistry of the elements: the inorganic chemistry of life.* Oxford, Oxford University Press.

Galdes, A., D. S. Auld, et al. (1986). *Biochem.* **25**: 646-651.

Galdes, A., D. S. Auld, et al. (1983). "Cryokinetic studies of the intermediates in the mechanism of carboxypeptidase A." *Biochemistry* **22**: 1888-1893.

Ghosh, M., A. Grunden, et al. (1998). "Characterization of native and recombinant forms of an unusual cobalt-dependent proline dipeptidase (prolidase) from the hyperthermophilic archaeon *Pyrococcus furiosus*." *J. Bacteriol.* **180**: 4781-4789.

Gill, S. C. and P. H. von Hippel (1989). *Anal. Biochem.* **182**: 319-326.

Glovsky, J., P. L. Hall, et al. (1972). *Biochem. Biophys. Res. Commun.* **47**: 244-247.

Gomez-Ortiz, M., F. X. Gomis-Ruth, et al. (1997). *FEBS Lett.* **400**: 336-340.

Gomisruth, F. X., F. Grams, et al. (1994). *J. Biol. Chem.* **269**: 17111-17117.

Grams, F., V. Dive, et al. (1996). "Structure of astacin with a transition-state analogue inhibitor." *Nat. Struct. Biol.* **3**: 671-675.

Green, S., A. Ginsburg, et al. (1991). "Roles of metal ions in the maintenance of the tertiary and quaternary structure of arginase form *Saccharomyces cerevisiae*." *J. Biol. Chem.* **266**: 21474-21481.

Guex, N. and M. C. Peitsch (1997). *Electrophoresis* **18**: 2714-2723.

Harada, S., T. Kinoshita, et al. (1995). *Eur. J. Biochem.* **233**: 683-686.

Hardeman, K., B. Samyn, et al. (1998). "An improved chemical approach toward the C-terminal sequence analysis of proteins containing all natural amino acids." *Prot. Sci.* **7**(7): 1593-1602.

Harwood, A. J., Ed. (1996). *Basic DNA and RNA protocols. Methods in Molecular* Biology. Totowa, N.J., Humana Press.

Harwood, V. J., J. D. Denson, et al. (1997). "Overexpression and characterization of a prolyl endopeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*." *J. Bacteriol.* **179**: 3613-3618.

Hensel, R. and H. Konig (1988). "Thermoadaptation of methanogenic bacteria by intracellular ion concentration." *FEMS Microbiol. Lett.* **49**: 75-79.

Hess, D., K. Kruger, et al. (1995). "Dimeric 3-phosphoglycerate kinases from hyperthermophilic archaea -cloning, sequencing and expression of the 3-phosphoglycerate kinase gene of Pyrococcus woesei in Escherichia coli and characterization of the protein." *Eur. J. Biochem.* **233**: 227-237.

Higgins, S. J. and B. D. Hames, Eds. (1999). *Protein Expression: A Practical Approach.* Oxford, Oxford University Press.

Holden, H. M., D. E. Tronrud, et al. (1987). *Biochem.* **26**: 8542.

Holland, D. R., A. C. Hausrath, et al. (1995). *Protein Sci.* **4**: 1955-1965.

Holland, D. R., A. C. Hausrath, et al. (1995). *Prot. Sci.* **4**: 1955-1965.

Holm, L. and C. Sander (1998). *Nucleic Acids Res.* **26**: 316-319.

Holmquist, B. and B. L. Vallee (1974). *J. Biol. Chem.* **249**: 4601-4607.

Hooper, N. M. (1994). "Families of zinc metalloproteases." *FEBS Letters* **354**: 1-6.

Hunt, J. A., M. Ahmed, et al. (1999). *Biochem.* **38**: 9054-9062.

Ishikawa, K., H. Ishida, et al. (2001). *Appl. Environ. Microbiol.* **67**: 673-679.

Jiang, W. and J. S. Bond (1992). *FEBS Lett.* **312**: 110-114.

Kengen, S. W. M. and A. J. M. Stams (1994). "Formation of L-alanine as a reduced end-product in carbohydrate fermentation by the hyperthermophilic archaeon *Pyrococcus furiosus*." *Arch. Microbiol.* **161**: 168-175.

Kengen, S. W. M., A. J. M. Stams, et al. (1996). "Sugar metabolism of hyperthermophiles." *FEMS Microbiol. Rev.* **18**: 119-137.

Kerr, M. A. and A. J. Kenny (1974). *Biochem. J.* **137**: 489-495.

Kiefer, L. L., S. A. Paterno, et al. (1995). *J. Am. Chem. Soc.* **117**: 6831-6837.

Kohlhoff, M., A. Dahm, et al. (1996). "Tetrameric triosephosphate isomerase from hyperthermophilic archaea." *FEBS Lett.* **383**: 245-250.

Kunugi, S., H. Hirohara, et al. (1982). *Eur. J. Biochem.* **124**: 157-163.

Kuo, L. C. and M. W. Makinen (1985). *J. Am. Chem. Soc.* **107**: 5255-5261.

Laderman, K. A., B. R. Davis, et al. (1993). "The purification and characterizaition of an extremely thermostable a-amylase from the hyperthermophilc archaebactrium *Pyrococcus furiosus*." *J. Biol. Chem.* **268**: 24394-24401.

Larsen, K. S. and D. S. Auld (1989). *Biochem.* **28**: 9620-9625.

Larsen, K. S. and D. S. Auld (1991). *Biochem.* **28**: 9620-9625.

Larsen, K. S. and D. S. Auld (1991). *Biochem.* **30**: 2613-2618.

Lee, S. H., E. Minagawa, et al. (1992). "Purification and characterization of a thermostable carboxypeptidase (carboxypeptidase Taq) from Thermus aquaticus YT-1." *Biosci. Biotech. Biochem.* **56**: 1839-1844.

Lee, S.-H., H. Taguchi, et al. (1994). "Carboxypeptidase Taq, a thermostable zinc enzyme, from Thermus aqauticus YT-1: molecualr cloning, sequencing, and expression of the encoding gene in *Escherichia coli*." *Biosci. Biotech. Biochem.* **58**: 1490-1495.

Legrain, C., V. Villeret, et al. (1997). "Biochemical characterisation of ornithine carbamoyltransferase from *Pyrococcus furiosus*." *Eur. J. Biochem.* **247**: 1046-1055.

Lesburg, C. A. and D. W. Christianson (1995). *J. Am. Chem. Soc.* **117**: 6838-6844.

Lippard, S. J. (1994). *Principles of Bioinorganic Chemistry.* Mill Valley, CA, University Science Books.

Lipscomb, W. N., J. C. Coppola, et al. (1966). *J. Mol. Biol.* **19**: 423.

Lumry, R., E. L. Smith, et al. (1951). *J. Amer. Chem. Soc.* **73**: 4330.

Lundberg, K. S., D. D. Shoemaker, et al. (1991). "High-fidelity amplification using a thermostable DNA-polymerase isolated from *Pyrococcus furiosus*." *Gene* **108**: 1-6.

Lupinetti, A. J., S. Fau, et al. (1997). *J. Phys. Chem. A* **101**: 9551-9559.

Ma, B., S. Kumar, et al. (2000). "Transition-state ensemble in enzyme catalysis: possibility, reality, or necessity?" *J. Theor. Biol.* **203**: 383-397.

Makinen, M. W., G. B. Wells, et al. (1984). *Adv. Inorg. Biochem.* **6**: 1-69.

Mallya, S. K. and H. E. Van Wart (1989). *J. Biol. Chem.* **264**: 1594-1601.

Maniatis, T. and T. Sambrook (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory.

Martins, L. O. and H. Santos (1995). "Accumulation of mannosyl glycerate and di-myo-inositol-phosphate by *Pyrococcus furiosus* in response to salinity and temperature." *Appl. Environ. Microb.* **61**: 3299-3303.

Masel, R. I. (2001). *Chemical Kinetics and Catalysis*. New York, John Wiley & Sons, Inc.

Mock, W. L. (1998). *Zinc proteinases, Academic Press Ltd*.**:** 425-453.

Mock, W. L. and M. Aksamawati (1994). *Biochem. J.* **302**: 57-68.

Mock, W. L., J.-T. Chen, et al. (1981). *Biochem. Biophys. Res. Commun.* **102**: 389-396.

Mock, W. L. and J.-T. Tsay (1988). *J. Biol. Chem.* **263**: 8635-8641.

Mock, W. L. and X. Xu (1994). *Bioorg. Chem.* **22**: 373-386.

Mock, W. L. and J. Z. Zhang (1991). *J. Biol. Chem.* **266**: 6393-6400.

Mustafi, D. and M. W. Makinen (1994). *J. Biol. Chem.* **269**: 4587-4595.

Nashed, N. T. and E. T. Kaiser (1986). 108: 2710-2715.

Orgel, L. E. (1960). *An Introduction to Transition-Metal Chemistry: Ligand Field Theory*. New York, John Wiley & Sons, Inc.

Owusu, R. K. and D. A. Cowan (1991). *Enz. Microbiol. Tech.* **13**: 158-163.

Pauling, L. (1935). *Proc. Nat. Acad. Sci. U.S.A.* **21**: 186.

Pearson, R. G. (1963). *J. Am. Chem. Soc.* **85**: 3533.

Perrin, D. D. (1982). *Ionization Constants of Inorganic Acids and Bases in Aqueous Solution*. Oxford, Pergamon Press**:** 136-138.

Perrin, D. D. (1982). *Ionization Constants of Inorganic Acids and Bases in Aqueous Solution*. Oxford, Pergamon Press**:** 136-138.

Quiocho, F. A. and F. M. Richards (1966). *Biochem.* **5**: 4062.

Rawlings, N. D. and A. J. Barrett (1995). *Methods in Enzymology*. **248**.

Rees, D. C., M. Lewis, et al. (1983). "Refined crystal-structure of carboxypeptidase A at 1.54 A resolution." *J. Mol. Biol.* **168**: 367-387.

Riechmann, L. and V. Kasche (1986). *Biochem. Biophys. Acta* **872**: 269-276.

Rosenberg, R. C. (1975). *J. Am. Chem. Soc.* **97**: 21.

Rowsell, S., R. A. Pauptit, et al. (1997). "Crystal Structure of Carboxypeptidase G_2, a Bacterial Enzyme with Applications in Cancer Therapy." *Structure* **5**: 337-347.

Schar, M., K. O. Bornsen, et al. (1991). "Monitoring of carboxypeptidase digestion by matrix-assisted laser desorption and ionization mass spectrometry." *Chimia* **45**: 123.

Schicho, R. N., K. Ma, et al. (1993). "Bioenergetics of sulfur reduction in the hyperthermophilic archaeon *Pyrococcus furiosus*." *J. Bacteriol.* **175**: 1823-1830.

Sebastian, J. F., R. F. Hinks, et al. (1987). *Biochem. Cell. Biol.* **65**: 717-725.

Sebastian, J. F. and W.-Y. Lo (1978). *Can. J. Biochem.* **56**: 329-333.

Segel, I. H. (1975). *Enzyme Kinetics*. New York, John Wiley & Sons.

Shi, J. Y., T. L. Blundell, et al. (2001). *J. Mol. Biol.* **310**: 243-257.

Shoham, G., D. W. Christianson, et al. (1988). *Proc. Nat. Acad. Sci. U.S.A.* **85**: 684-688.

Snowden, L. J., I. I. Blumentals, et al. (1992). "Regulation of proteolytic activity in the hyperthermophile *Pyrococcus furiosus*." *Appl. Environ. Microbiol.* **58**: 1134-1141.

Solc, M. (1995). *J. Theor. Biol.* **175**: 57-61.

Solc, M. and J. Hostomsky (1999). *React. Kinet. Catal. Lett.* **67**: 53-58.

Steiner, D. F. (1998). "The proprotein convertases." *Curr. Opin. Chem. Biol.* **2**: 31-39.

Stepanov, V. M. (1995). "Carboxypeptidase T." *Meth. Enzymol.* **248**: 675-683.

Sterner, R., G. R. Kleemann, et al. (1996). "Phosphoribosyl anthranilate isomerase from Thermotoga maritima is an extremely stable and active homodimer." *Prot. Sci.* **5**: 2000-2008.

Sterner, R. and W. Liebl (2001). "Thermophilic Adaptation of Proteins." *Crit. Rev. Biochem. Mol. Biol.* **36**: 39-106.

Thiede, B., J. Salnikow, et al. (1997). "C-terminal ladder sequencing by an approach combining chemical degradation with analysis by matrix-assisted laser desorption ionization mass spectrometry." *Eur. J. Biochem.* **244**: 750-754.

Thompson, J. D., D. G. Higgins, et al. (1994). *Nucleic Acids Res.* **22**: 4673-4680.

Trachtman, M., G. D. Markham, et al. (2001). **40**: 4230-4241.

Trachtman, M., G. D. Markham, et al. (1998). *Inorg. Chem.* **37**: 4421.

Tsunasawa, S., Y. Izu, et al. (1997). "Methionine aminopeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus* - molecular cloning and overexpression in Escherichia coli of the gene, and characteristics of the enzyme." *J. Biochem.* **122**: 843-850.

Tu, C. K. and D. N. Silverman (1985). *Biochem.* **24**: 5881-5887.

Tuan, R. S., Ed. (1997). *Recombinant Protein Protocols. Methods in Molecular Biology.* Totowa, N.J., Humana Press.

Umetsu, H., H. Matsuoka, et al. (1983). *J. Agric. Food. Chem.* **31**: 50-53.

Vallee, B. L., J. A. Rupley, et al. (1960). *J. Biol. Chem.* **235**: 64-70.

Voorhorst, W. G. B., R. I. L. Eggen, et al. (1996). "Isolation and characterization of the hyperthermostable serine protease, pyrolysin, and its gene from the hyperthermophilic archaeon *Pyrococcus furiosus*." *J. Biol. Chem.* **271**: 20426-20431.

Wayne, S. I. and J. S. Fruton (1983). *Proc. Nat. Acad. Sci. U.S.A.* **80**: 3241-3244.

Wingfield, P., P. Graber, et al. (1989). "Purification and characterization of a methionine-specific aminopeptidase from *Salmonella typhimurium*." *Eur. J. Biochem.* **180**: 23-32.

Woese, C. R., O. Kandler, et al. (1990). "Towards a natural system of organisms - proposal for the domains archaea, bacteria, and eucarya." *Proc. Natl. Acad. Sci. (USA)* **87**: 4576-4579.