Expression, Purification, Crystallization and Preliminary X-ray Analysis of PR-AMP Cyclohydrolase, a Zinc Binding Enzyme from *Methanobacterium Thermoautotrophicum*

Lorena Boju

A Thesis

in

The Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements

for the Degree of Master in Science at

Concordia University

Montreal, Quebec, Canada

Your file Votre référence
ISBN: 0-612-94665-7
Our file Notre référence
ISBN: 0-612-94665-7

Expression, Purification, Crystallization and Preliminary X-ray Analysis of PR-AMP Cyclohydrolase, a Zinc Binding Enzyme from *Methanobacterium Thermoautotrophicum*

Abstract

Histidine is an essential amino acid since humans and other higher eukaryotes lack the cellular biosynthetic machinery for its manufacture. Phosphoribosyl-AMP cyclohydrolase (PRA-CH, EC 3.5.4.19) is the third enzyme in the histidine biosynthetic pathway. PRA-CH catalyzes the hydrolysis of the N1-C6 bond of the purine substrate N1-(5`-phosphoribosyl)adenosine 5`-monophosphate (PR-AMP). The enzyme from *Methanobacterium thermoautotrophicum,* overexpressed in *E. coli*, was selected for crystallization trials as steps towards obtaining its three dimensional structure by X-ray diffraction.

Initial attempts of purifying the recombinant hexa histidine tag protein were not successful. A new construct encoding for PRA-CH without the His-tag was made, overexpressed in *E. coli* and purified to homogeneity using heat treatment followed by one chromatographic step. Yields of 50-60 mg were obtained per liter of culture. The enzyme was shown by Dynamic Light Scattering studies to form a homodimer of 31 500 Da.

Crystals of the protein were formed at an unusually low concentration of 2.1 mg/ml by the vapour-diffusion method using sodium acetate as precipitant. The presence of 50 mM $Cd(SO_4)_2$ in the reservoir solution was essential for the crystallization of the protein. Crystals appeared within two weeks at room temperature. They belong to the orthorhombic system, space group $P2_12_12_1$, with cell dimensions a=39.7Å, b=54.3Å, c=117.3Å, ↦= ↕=□=90°. These crystals diffracted to 1.76 Å at the synchrotron beamline. Several attempts to determine the 3-D structure of PR-AMP cyclohydrolase by multiwavelength anomalous dispersion were unsuccessful. Following completion of the thesis, however, the structure was eventually solved by single wavelength anomalous dispersion.

ACKNOWLEDGMENTS

I would like to express my gratitude to Dr. Joanne Turnbull for her guidance, invaluable help and support.

I would like to express my appreciation for all the good friends I have made at MSG, especially to Dr. Jayaraman Sivaraman, Dr. Jia Jia and Yunge Li.

I would like to acknowledge Dr. Miroslaw Cygler for directing this project at the Biotechnology Research Institute/NRCC. I would also like to thank him for the guidance, encouragement and support in the supervision of this thesis. Without his generous support, this work would have not been possible.

I am indebted to my family for their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1:

Chapter 2:

LIST OF TABLES

Chapter 3:

LIST OF ABBREVIATIONS

Amp	ampicillin
ATP	adenosine triphosphate
BME	↕-mercaptoethanol
BSA	bovine serum albumin
CG	circle grow medium
DEAE	diethylaminoethyl
DLS	dynamic light scattering
DTT	dithiothreitol
EDTA	ethylenediamine tetra-acetic acid
HEPES	*N*'-2-hydroxyethylpiperazine-*N*'-ethanesulfonic acid
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-1-thio-ß-D-galactopyranoside
Kan	kanamycin
LB	Luria-Bertani broth
MAD	multiwavelength anomalous dispersion
MIR	multiple isomorphous replacement
MW	molecular weight
NTA	nitrolotriacetic acid
PAGE	polyacrylamide gel electrophoresis
PRA-CH	N^1-(5`-phosphoribosyl)adenosine 5`-monophosphate cyclohydrolase
PR-AMP	N^1-(5`-phosphoribosyl)adenosine 5`-monophosphate

5'-ProFAR	*N*'[(5'-phosphoribosyl)-formimino]-5aminoimidazole-4carboxamide ribonucleotide
PRPP	5-phospho-D-ribosyl-1-pyrophosphate
RH	hydrodynamic radius
RPM	revolutions per minute
SAD	single wavelength dispersion
SDS	sodium dodecylsulfate
Tris	tris(hydroxymethyl)aminomethane

Chapter 1 Introduction

1.1 Archaea

The Archaea are a unique class of organisms that form a third distinct kingdom of life. Archaeal organisms include methanogens, sulfur-reducing thermophiles, sulfur-dependent thermophiles, and halophiles (1). Early interest in the Archaea stemmed from their ability to thrive in extreme environments, such as high temperatures, high salinity, or low pH, which are inhospitable to most other forms of life (2). Genomes from all four categories are being sequenced or have already been released. The methanogens (e.g. *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*) are strict anaerobes and use variations of methanogenesis to convert CO_2, methyl compounds, or acetate to methane. Methanogenesis serves as a form of anaerobic respiration. In the sulfur-reducing thermophiles (*Archaeoglobus fulgidus*), oxidized sulfur species act as electron acceptors for anaerobic respiration. Sulfur-dependent Archaea (*Sulfolobus solfataricus*) can grow autotrophically using elemental sulfur as an energy source. The halophiles (*Halobacterium salinarum*) require a high salt environment. Because their outward physical appearance resembles the Bacteria, it was originally assumed that these "extremophiles" were unusual members of the bacterial kingdom. This assumption was later challenged after a detailed analysis of ribosomal RNA genetic patterns (1). Archaeal proteins that participate in genetic information storage and expression, namely in DNA packaging and replication, in RNA synthesis and processing, and in protein synthesis are more closely related to their eukaryotic than bacterial counterparts (3).

1.2 The Importance of Methanogenic Organisms

Methane, also known as natural gas or biogas, is generated in the final step of the anaerobic degradation of industrial, urban and agricultural waste materials and, as such, represents a major, renewable energy source. However, methane is also a greenhouse gas and methane production contributes substantially to global warming (4). For this reason, studies are being conducted to understand the microbiology, biochemistry and molecular biology of methane production with the twin goals of maximizing the use of this biotechnology to convert waste materials into methane as an alternative to fossil fuels, and determining how methane generation and methane release to the atmosphere can be appropriately controlled as a key step in global carbon management (4, 5).

On a world-wide basis, microbial biodegradation of waste materials is almost certainly the largest biotechnology. It is used to reduce and detoxify municipal, commercial and agricultural wastes, in industrial and municipal sewage treatment facilities, in septic tanks and in landfills (6). The process works empirically, but opportunities may well exist to improve and control this biotechnology, to convert more waste to methane contributions to global warming (4).

1.3 *Methanobacterium thermoautotrophicum*

M. thermoautotrophicum is a methane producing microorganism that belongs to Archaea, a moderate thermophile positioned near the center of the methanogen evolutionary tree (Figure 1.1) (7); it has rod-shaped cells and grows optimally at 65°C (about 150°F) (8).

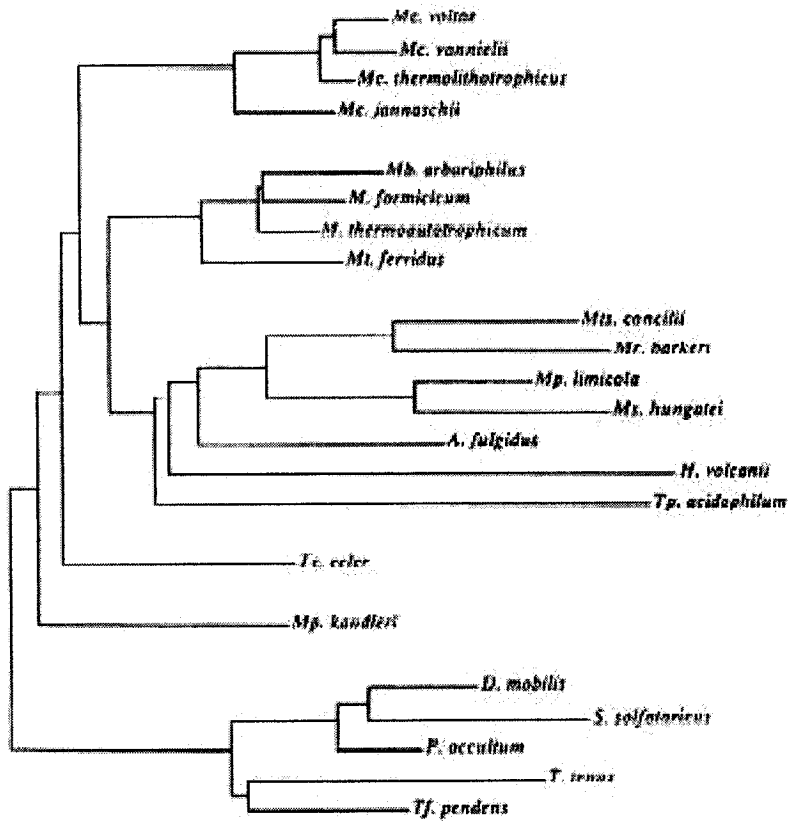


Figure 1.1. Phylogenetic tree of methanogens based on 16S ribosomal RNA sequences (7).

This anaerobic methanogenic archaeon was originally isolated from the municipal waste-treatment facility in Champaign, Il, in 1971 (9), and is it widely studied as a representative of the methanogens that inhabit all biodegradation facilities.

The whole genome *M. thermoautotrophicum* ΔH was sequenced in 1997 (10) and it consists of 1,751,377 base pairs. This Archaeon synthesizes all of its cellular components and derives energy from just CO_2, H_2, and salts. Nevertheless, it has a genome that is only approximately 40% the size of *Eschericia coli* (10).

1.4 Histidine Biosynthesis Pathway

The study of the histidine biosynthetic pathway in prokaryotes and lower eukaryotes began more than 40 years ago (11). The histidine biosynthesis pathway has been extensively studied in a number of microorganisms (12). In eubacteria such as *E. coli* and *Salmonella typhimurium*, the complete nucleotide sequences of the genes involved in the histidine biosynthesic pathway have been determined (Figure 1.2), and it was shown that 10 enzymatic activities are encoded by eight genes organized in a single operon (13). In eukaryotes such as *Saccharomyces cerevisiae* some of these genes are fused: *hisI* and *hisD* are coded by HIS4 and *hisF* and *hisH* by HIS7 (14). The *his* genes in Archaea are less well known than in eubacteria and a similar biosynthetic pathway is expected to operate throughout Archaea (15).

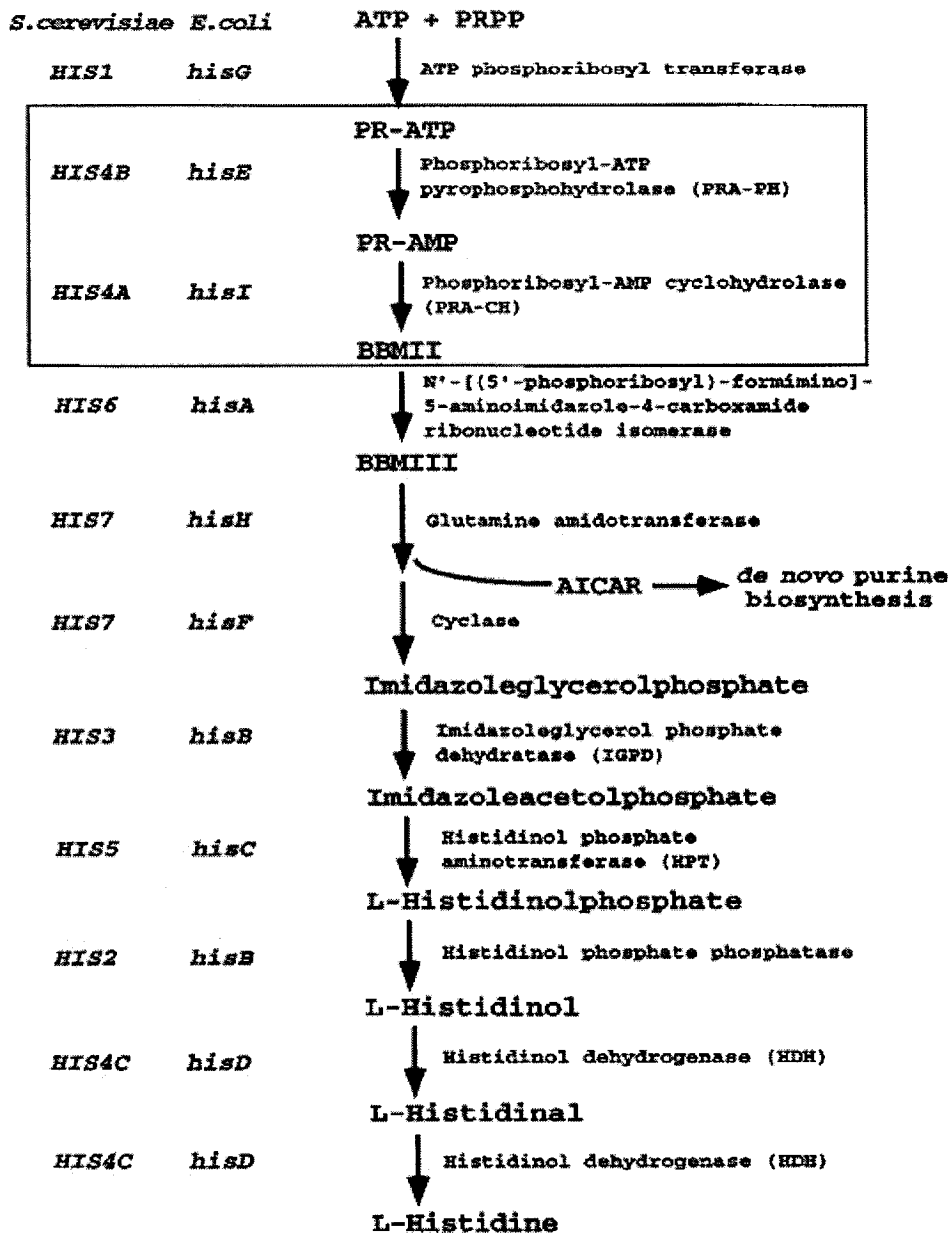


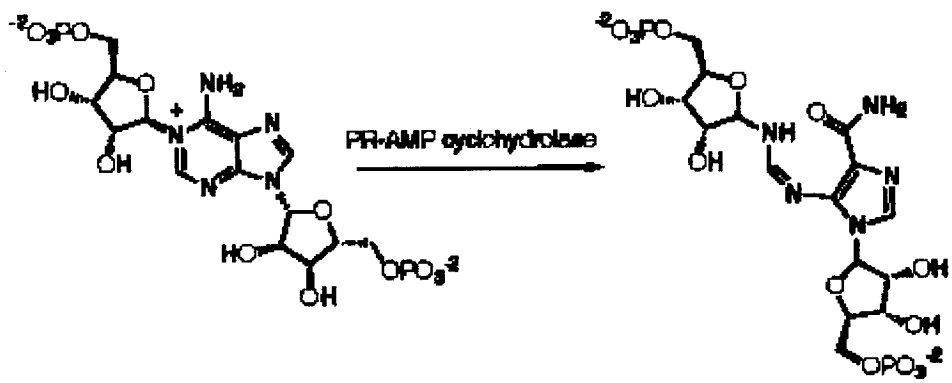
Figure 1.2. Histidine biosynthetic pathway. The pathway uses 5-phospho-D-ribosyl-1-pyrophosphate (PRPP) and ATP as the initial substrates. Boxed region refers to the reactions catalyzed by the gene products of HIS4 of *S. cerevisiae* and *hisIE* of *E. coli*.

Histidine is an essential amino acid since humans lack the biosynthetic apparatus for its manufacture. On the other hand, the genes encoding the enzymes of the histidine biosynthetic pathway are widely present in bacteria. Hence, these enzymes are attractive targets for the design of novel agricultural herbicides (16, 17). It follows that identification of the catalytic mechanism of enzymes of the histidine biosynthesis pathway is important.

Histidine production is accomplished at significant metabolic expense (18), highlighted by the precursors ATP and 5-phospho-D-ribosyl-1-pyrophosphate (PRPP), and requires integration with de novo purine metabolism. A key step for utilizing the purine heterocycle of ATP in histidine biosynthesis involves hydrolytic cleavage of the unusual nucleotide N^1-(5`-phosphoribosyl)adenosine 5`-monophosphate (PR-AMP) (Figure 1.3) (19, 20). The enzyme that catalyzes this reaction, N^1-(5'-phosphoribosyl)adenosine-5'-monophosphate cyclohydrolase (PRA-CH, EC 3.5.4.19) is the focus of our research.

1.5 Phosphoribosyl-AMP Cyclohydrolase (PRA-CH)

PRA-CH represents a link between a catabolic and anabolic process by catalyzing a step that initiates redistribution of carbon and nitrogen from ATP into the amino acid histidine (19). In *Azospirillum brasilense* (21) and in some Archaea (22), independent genes encode the pyrophosphohydrolase (*hisE*) and cyclohydrolase (*hisI*), while in *E. coli* and *S. typhimurium* (13) these functions are fused into a single gene *hisIE* (Figure 1.2).



PR-AMP **5'-ProFAR**

Figure 1.3. Conversion of PRA-CH to 5'-ProFAR. PRA-CH catalyzes the hydrolysis of the N1-C6 bond of the purine substrate to form *N*'[(5'-phosphoribosyl)-formimino]-5 aminoimidazole-4 carboxamide ribonucleotide (5'-ProFAR) (20).

Multifunctional enzymes with the activities corresponding to the HisIE and HisD (histidinol dehydrogenase) that catalyzes the final step of the histidine biosynthesis pathway are encoded by HIS4 in *S. cerevisiae* (14) and *Pichia pastoris* (23), and the HIS3 gene of *Neurospora crassa* (24).

The only PRA-CH from Archaea to be biochemically characterized is from *M. vannielii* (20). There are no structural studies reported for the enzyme from *M. thermoautotrophicum*, nor is there any structural information on the enzyme from any organism. R. L. D'Ordine et al. (20) reported the importance of metals in the mechanism of PRA-CH from *M. vannielii*. These authors showed that one equivalent of Zn^{2+} bound to the enzyme that could be removed only by extensive dialysis with 1,10-phenanthroline. Furthermore, the presence of Zn^{2+} in the lysis buffer was shown to enhance the specific activity and stability of the enzyme. The authors have proposed a unique motif, $C^{93}(X)_{15}C^{109}H^{110}(X)_5C^{116}$ (*M. vannielii* numbers), for the Zn^{2+} binding domain. In addition, the enzyme is reversibly inhibited by inclusion of EDTA in the assay mixture, demonstrating that Mg^{2+} is required for catalytic activity. In PRA-CH, the region $S^{67}R[ST]R(X)_2[LI]WXKG[EA]TSG^{81}$ (*M. vannielii* numbers) represents a consensus sequence that would constitute a 'P-loop type' region associated with Mg^{2+}-phosphate binding site.

Alignment of several PRA-CH domains from different organisms revealed a number of highly conserved residues including one His, three Cys and three Asp (Figure 1.4). These

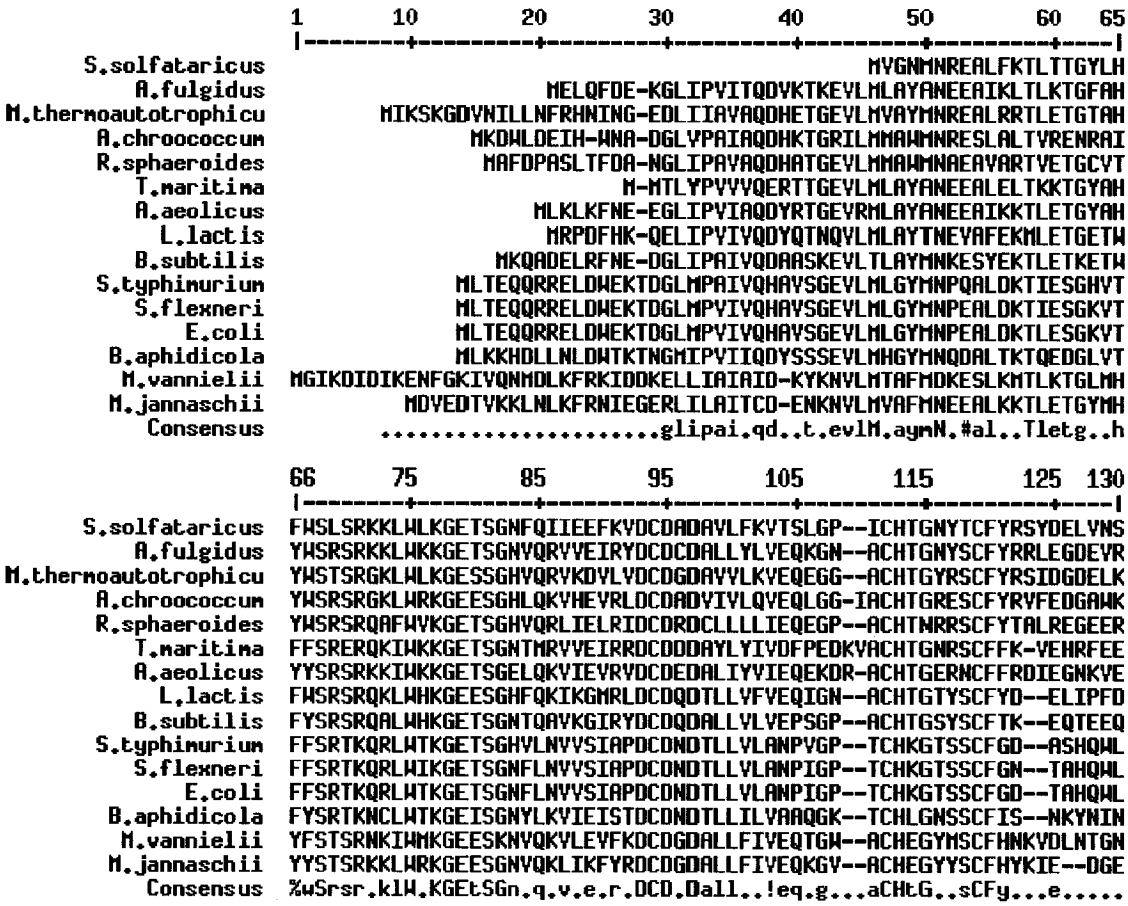


Figure 1.4. Sequence alignment of PRA-CH domain from different organisms. The sequence alignment was performed using Multalin (25). The sequences were obtained from different databases. The color red indicates the conserved amino acids mentioned in the text.

residues are often seen in Zn binding motifs (26, 27). Thus, the alignment suggests that zinc ion in *M. thermoautotrophicum hisI* gene product might be ligated by a combination of nitrogen (His 16), oxygen (Asp 85, Asp 87 and Asp 89), and sulphur (Cys 86, Cys 102 and Cys 109) atoms. The potential role for histidine residue is not unexpected since histidine is the most common ligand for Zn^{2+} among zinc metalloenzymes (26, 27); closely spaced ligands at structural zinc sites could be consistent with an overall stabilizing role for the protein structure.

1.6 Overview of Techniques And Technologies Used in the Thesis

1.6.1 Purification of Recombinant Protein Carrying a Hexa-His Affinity Tag

Recombinant protein that encodes a hexa-His affinity tag on either N or C terminus of the protein will bind to a Ni-containing matrix , such as Ni-NTA coupled to Sepharose CL-6B sold by Qiagen. Nitriloacetic acid (NTA) is a tetrahydratentate chelating adsorbent that occupies four of the six ligand binding sites in the coordination sphere of the Ni^{2+}, leaving two sites free to interact with the hexa-His tag (Figure 1.5).

The imidazole rings in the histidine residues of the hexa-His tag bind to the nickel ions immobilized by the NTA groups on the matrix; binding of tagged proteins to Ni-NTA resin is conformation-independent. The use of 5-20 mM imidazole in the chromatographic buffers allows efficient binding of the hexa-His tag proteins to the resin

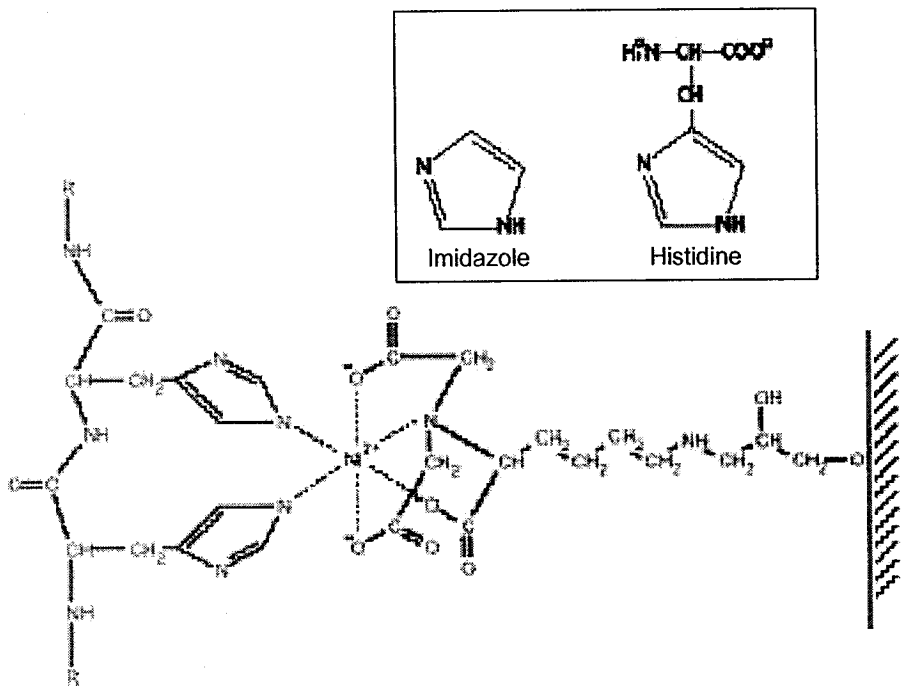


Figure 1.5. Ni-NTA resin illustrating its coordination to histidine group. The structure of imidazole is shown in inset. (Adapted from The QIAexpresionlist 5th edition, March 2001, user's manual.)

and prevents non-specific binding of proteins. The tagged protein is then eluted from the matrix at a high concentration of imidazole (200-300 mM) (Figure 1.5 inset).

Utilizing this technology generally allows the purification of proteins from less than 1% of the total protein preparation to more than 95% homogeneity in just one step (28).

1.6.2 Dynamic Light Scattering

Dynamic Light Scattering (DLS) is a non-invasive, non-destructive technique ideal for the investigation of a wide range of biological macromolecules such as proteins, DNA, polysaccharides and other macromolecules (29). Dynamic light scattering technique was used in our research to assess the homogeneity of protein samples after purification trials, since molecules that are monodisperse tend to readily yield crystals, while polydisperse and aggregating samples are less likely to yield crystals (30). This technique can be also used to rapidly scan for pH ranges, additives, and cofactors that tend to improve the solubility of the protein in solution.

The instrument used was the DynaPro-801, a molecular sizing device from Protein Solutions, Inc., Charlottesville, VA based on the principle of Dynamic Light Scattering. This instrument analyzes scattered light intensity fluctuations caused by the Brownian motion of the protein macromolecules in solution. This instrument can provide estimates of the hydrodynamic size, conformation, interaction, or extent of aggregation of the

biomolecule. We used this technique for its sensitivity to trace amounts of aggregation in protein sample that would go undetected using other characterization techniques.

The configuration of the Dyna Pro-801 DLS instrument is outlined in Figure 1.6 (29). A laser beam is directed through a sample of molecules in an aqueous solution. The light leaves the sample at a scattering angle of 90° and is guided to the Avalanche photo diode (APD), a single photon detector.

The following are some of the significant parameters used to estimate the quality of the protein sample using the instrument software:

(1) **(RH)** (nm), or hydrodynamic radius - represents the z-average size value of the molecules in solution.

(2) **polydispersity** (nm) - when its value is under 25% of the mean radius (polydispersity/RH) it shows that the sample is monodisperse.

(3) **scattering amplitude** – is a measure of signal to noise ratio. The highest theoretical limit for the amplitude is 1.0. A minimum of 0.2 should be obtained for the measurements to be considered reliable.

(4) **count rate** - total scattering signal for the measurement period. Contains contributions from interfering solvents (such as glycerol), and protein molecules : it correlates linearly with the sample concentration and exponentially with the molecular size.

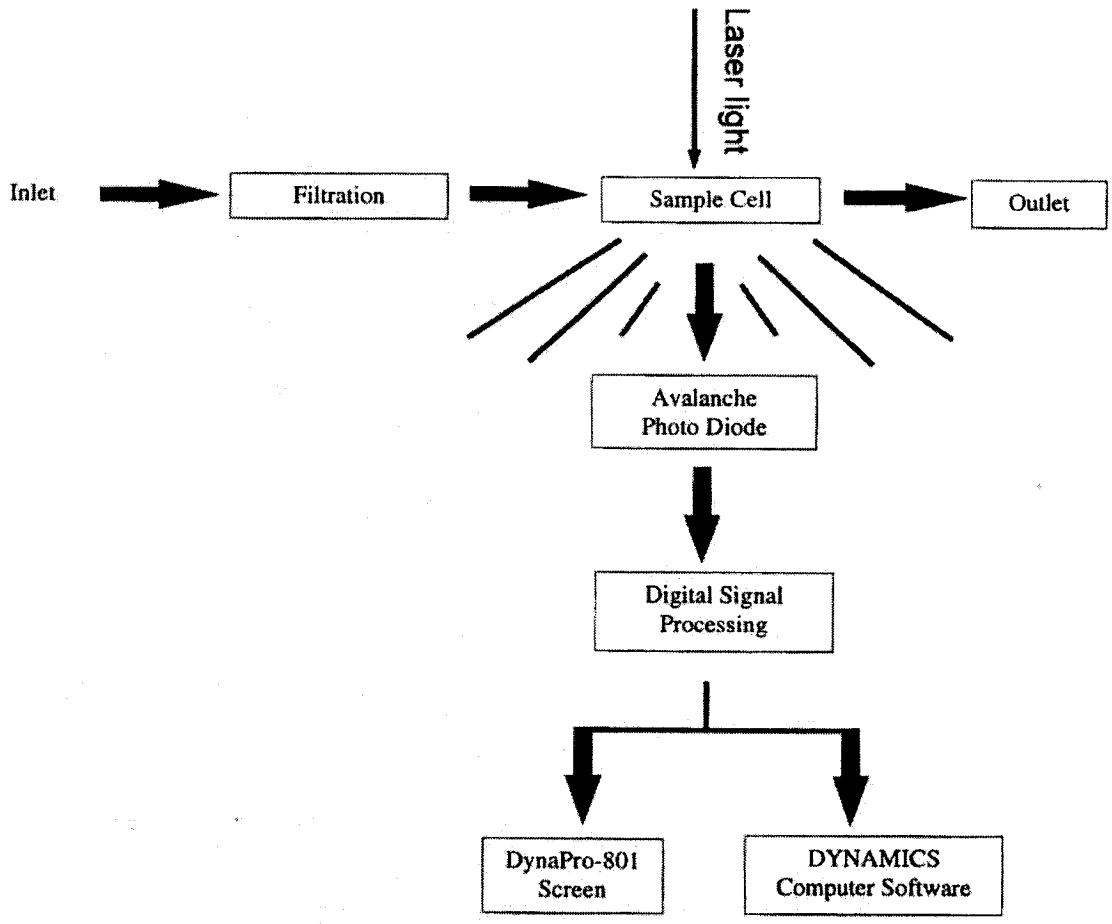


Figure 1.6. DynaPro Dynamic light scattering setup. (Adapted from the DynaPro-801™, operator's manual, version 4.)

(5) **baseline (B)** and **sum of squares (SOS)** - assess the quality of the analysis. B is derived from the completeness of the autocorrelation function and SOS represents the sum of squared residual errors between the raw data and the fitted curve.

1.6.3 General Overview of X-Ray Diffraction

The correlation of regulation, mechanism, and function of proteins with their detailed molecular structure has now become a primary aspiration of modern biochemistry and molecular biology. There are various physical-chemical approaches that yield information regarding macromolecular structure, however only two are widely used for the structural determination of macromolecules at atomic resolution: nuclear magnetic resonance (NMR) and X-ray diffraction of crystals (30). While NMR provides more detailed information on the dynamics of the molecule in question, it can be mainly used for proteins with a molecular weight of less than 20,000 Da (31). X-Ray crystallography can be applied to compounds with molecular weight up to at least 10^6 Da, and is a technique that provides a detailed and precise description of a macromolecule three dimensional structure (32).

1.6.3.1 Crystallization

The first requirement for protein structure determination by X-ray diffraction is to obtain suitable crystals. A crystal is built up from many small identical units; each may contain one or more protein molecules and solvent channels are formed between these individual

molecules (31). When the protein precipitates from a solution, its molecules attempt to reach the lowest free energy state (30). This is often accomplished by packing them in a regular way, in other words, a crystal grows. In this regular packing, three repeating vectors a, b, and c can be recognized with the angles a, ß, and ? between them. These six parameters define a unit cell in the crystal lattice.

In order to initiate crystallization, the protein solution has to be brought to a thermodynamically unstable state of supersaturation (33). Macromolecular crystallization implies, thus, a systematic search of ranges of the individual parameters (buffer, pH, salt) that impact upon crystal formation, finding a set of these factors that yield some kind of crystals, and then optimizing the variable sets to obtain the best possible crystals for X-ray analysis (30). This is done by establishing a large array of crystallization trials, evaluating the results, and using information obtained to improve the crystal quality in successive rounds of trials (30).

1.6.3.2 X-Ray Diffraction

When the X-ray beam strikes the crystal most of the X-rays travel through it, but some interact with electrons in each atom and cause them to oscillate. The oscillating electrons serve as a new source of X-rays, emitted in all directions (scattering) (34). In most cases, these scattered X-rays collide from different directions and cancel out; however, those from certain directions (X-rays with the same diffraction angle emerging from all atoms)

will add together to produce diffracted beams, which are recorded by an electronic area detector as reflection spots.

In order to solve the atomic structure of a crystal, it is only necessary to solve the structure of a single cell denoted the *unit cell* (Figure 1.7). The rest of the crystal structure is generated by periodic translations in the three dimensions of space. The unit cell content is described by the electron density _(x) with the dimension electrons per unit volume. Since this electron cloud has a radius comparable to the X-ray wavelength, the contribution falls off at a higher diffraction angles, *i.e.* at higher resolution. This is represented by the normal atomic scattering factor. Such a signal from the whole atom is isotropic and can be treated as a real number, f^0 (?) (34, 35).

The X-ray diffraction data obtained from a crystal contains the structure factor amplitude $|F(hkl)|$, because $I(hkl) \sim |F(hkl)|^2$, but information about the relative phase angles of the reflections is lost. To calculate an electron density map the phase angles for the collected amplitudes must be determined. It is possible to use the molecular replacement (MR) (24) method to obtain the initial phases if the structure of the target protein is sufficiently similar to a known structure. Otherwise derivatization must be made to make use of the phasing information that heavy metals produce by the multiple isomorphous replacement (MIR), a useful method for phase determination in macromolecular crystallography (36). The initial step in the MIR method requires attachment of heavy atoms and subsequently the determination of the coordinates of the heavy atoms in the unit cell. The next step is to refine heavy atom parameters in order to calculate the most

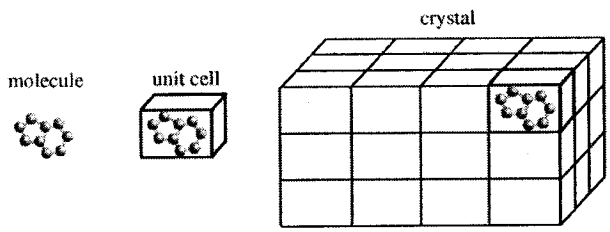


Figure 1.7. The crystal structure representation as periodic translations in the 3D space of the unit cell (31).

accurate phases, and to get statistics that give information about the quality of the derivative and the highest resolution at which it can be used.

To extract information about individual atoms from such a system, Fourier transformation is used. Each diffracted intensity I that is recorded as a reflection spot on the area detector can be associated to a reciprocal vector h: I = I(h) and is defined by three properties: amplitude that can be measured from the intensity of the spot, the wavelength of the X-rays which is known, and the phase that is lost during the X-ray experiments. The problem to overcome is to obtain the density as a function of the intensities (34).

1.6.3.3 Anomalous Scattering

Anomalous scattering (36) provides an invaluable source of phase information in macromolecular crystallographic structure determination; anomalous scattering is usually measured from heavy atoms which are incorporated into the crystal lattice to form derivative crystals. This is performed either by soaking in a suitable heavy-atom salt solution (30, 36) or by direct modification of amino acid residues or nucleic acid bases (37). Metalloenzymes possess a natural source of significant anomalous scattering. These heavy atoms are strong scatterers, that is, they contain more electrons than the light atoms such as H, N, C, O and S (38). They significantly enhance the diffraction pattern, however they should not change the structure of the molecule (crystals should be isomorphous). All diffracted beams would therefore increase in intensity after heavy-

metal substitution: some spots increase in intensity, some will decrease, and some will show no difference.

Precise control of the wavelength of synchrotron radiation allows both anomalous and dispersive differences to be varied. At the same time, phasing software improved dramatically. These developments contributed to the popularity of the multiwavelength anomalous dispersion (MAD) (36) method of phasing. The full atomic form factor is described by $f(?,\ell) = f^{0}(?) + f'(\ell) + i f''(\ell)$, where f' and f'' are the anomalous dispersion corrections that depend only on the wavelength ℓ of the X-rays used for the diffraction experiment. The observed anomalous signal is typically very small, therefore, maximizing the anomalous signal is crucial to the experiment. Both anomalous-scattering factors f'' and f' from X-ray fluorescence spectra have their respective local maxima and minima in the vicinity of characteristic absorption edges. Thus, data are commonly measured at X-ray wavelengths near these edges.

The most popular MAD vehicle is seleno-methionine (SeMet) (37), which is introduced into proteins in place of normally occurring methionine by genetic engineering. Although this approach is routinely used in the current practice, it has some drawbacks. A common problem with heavy-atom derivatization is the resulting nonisomorphism between the native and derivatized crystal. Sometimes, several different derivatives are required to achieve success.

A third approach of phasing macromolecular crystal structures is single wavelength anomalous dispersion (SAD) (39). With better data-collection facilities and cryogenic techniques, combined with powerful data processing, phasing and density-modification programs, the SAD approach may prove simpler than phasing from multiwavelength (MAD) measurements. It can be performed at any wavelength where anomalous scattering can be observed, in many cases using laboratory X-ray sources. However, there is still a need for accurate data, successful phase improvement and a certain amount of luck.

1.6.3.4 Model Building and Crystallographic Refinement

The amplitudes and the phases of the diffraction data from the protein crystals are used to calculate an electron density map (33, 35) of the repeating unit cell (Figure 1.8). The quality of this map depends on the resolution of the diffraction data. The initial model contains large errors. Structure factors calculated based on this model are generally in rather poor agreement with the observed structure factors. Crystallographic refinement (33) is the process of adjusting the model in order to find a close agreement between the calculated and observed structure factors. The refinement can significantly decrease the errors of the model and in most cases, it is necessary to rebuild the model according to the new electron density maps calculated from the refined coordinates. Crystallographic refinement and model rebuilding are iterated until the model is as complete as the data allows, and no more improvement can be obtained by further refinement.

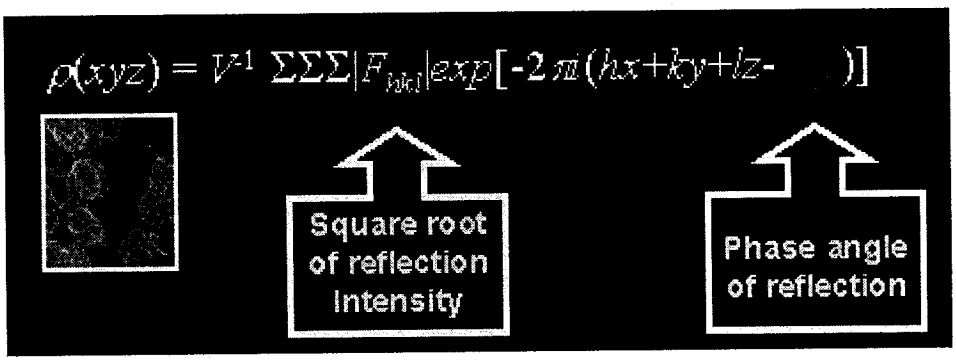


Figure 1.8. Electron density map as a function of the amplitudes and the phases of the diffraction data collected from the protein crystal (31).

1.7 Research Objectives

The general goal of the research described in this thesis was to obtain structural information for phosphoribosyl-AMP cyclohydrolase, a zinc metalloenzyme from the thermophile Archaea *M. thermoautotrophicum*.

Studies to date indicate that subtle structural differences between enzymes that possess a metal center and those that do not might help in the generation of diffractive protein crystals (30). Such information will lead to a better understanding of the catalytic mechanism of this important enzyme from Archaea and other organisms.

The specific objectives of my M. Sc. work were:

1) Purification of the enzyme to near homogeneity.

2) Crystallization of the protein.

3) Obtaining preliminary structural information by X-ray diffraction.

Chapter 2 Strategy for Purifying Recombinant Hexa-Histidine Tagged and Non-His Tagged PRA-CH

2.1 Experimental Procedures

2.1.1 Materials

The pET-15b vector was purchased from Novagen. Ampicillin (Amp) was obtained from Life Technologies and kanamycin (Kan) from Sigma. Isopropyl-1-thio-ß-D-galactopyranoside (IPTG) was from Sigma, CompleteC-EDTA free protease inhibitors from Roche Diagnostics and human-alpha thrombin from Haematologic Technologies Inc. MT 245 plasmid and BL21(DE3) Gold Magik *E. coli* strain were kindly donated by Dr. Aled M. Edwards, Ontario Center for Structural Proteomics, University of Toronto. MT245 plasmid contains the *hisI* gene cloned into a pET-15b vector, which adds a sequence encoding an N-terminal his tag followed by a thrombin recognition cleavage site. It also encodes a T7 RNA polymerase-dependent promoter for protein overexpression. The helper plasmid Magik encodes three tRNAs, Arg AGG, Arg AGA, and Iso ATA that are frequently used by archaeons and eukaryotes, but are rare in *E. coli* (40, 41). The MC1061 *E.coli* strain, chosen because it is a good source of DNA, was obtained from Stratagene, and pET-11a DNA was from Novagen. QIAprep spin miniprep kit, QIAex II agarose gel extraction kit and QIAquick PCR purification kit were purchased from Qiagen. *BamH I* (20,000 U/mL), *Nde I* (20,000 U/mL), *Ava I* (20,000 U/mL), T4 DNA ligase (400,000 U/mL), alkaline phosphatase (10,000 U/mL) and their

corresponding buffers were obtained from New England BioLabs. The components of the buffers are as follows: **NEBuffer for BamHI** (10x): 1.5 M NaCl, 0.1 M Tris-HCl (pH 7.9), 0.1 M $MgCl_2$, 0.1 M dithiothreitol. **NEBuffer 4 for NdeI** (10x): 0.5 M potassium acetate, 0.2 M Tris-acetate (pH 7.9), 0.1 M magnesium acetate, 10 mM dithiothreitol (DTT). **T4 DNA ligase buffer** (10x): 0.5 M Tris-HCl (pH 7.8), 0.1 M $MgCl_2$, 0.1 M DTT, 10 mM ATP, 0.25 mg/mL.

C/AP buffer (10x) (buffer for alkaline phosphate): 0.1 M Tris-HCl (pH 8), 0.5 M KCl, 10 mM $MgCl_2$, 1 mM $ZnCl_2$, 50 % glycerol. **Lysis buffer**: 50 mM Tris (pH 7.5), 0.5 M NaCl, 20 mM ↕-mercaptoethnol (BME), 5% glycerol (v/v), 0.5 % Triton X-100 (v/v), 20 mM BME and one tablet of CompleteC-EDTA free protease inhibitors. **Binding buffer**: 5 mM imidazole, 50 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 0.5% Triton X-100 (w/v), 5% glycerol (v/v), 20 mM BME and CompleteC-EDTA free protease inhibitors (one tablet/50 mL buffer solution). **Washing buffer**: 30 mM imidazole, 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 5% glycerol. **Elution buffer**: 250 mM imidazole, 50 mM Tris-HCl (pH 7.5), 500 mM NaCl. Circle Grow medium was purchased from Bio 101, Inc., Carlsbad, CA. Ni-NTA affinity resin was purchased from Qiagen and CHELEX 100 resin from BioRad. DEAE-Sepharose and Poros-50D resins were purchased from Pharmacia, Uppsala. Prepacked Superose 12 HR 10/30 column was also from Pharmacia, as well as the AKTA protein purifier system that was used as an FPLC. Hydroxyapatite resin (Bio-Gel HTP Hydroxyapatite) was obtained from Bio-Rad. Mono Q HR 5/5 (0.96 mL bed volume) and the prepacked Superdex 75 HR 10/30 (24 mL bed volume) were from Pharmacia, Piscataway, NJ. Centricon and Centriprep YM-30 ultrafiltration units were purchased from Millipore. All other chemicals were of the highest purity available.

2.1.2 Methods

2.1.2.1 Preparation of Glycerol Stocks

Ampicillin (Amp) and kanamycin (Kan) were used at concentrations of 0.1 mg/mL and 0.05 mg/mL, respectively. For the preparation of glycerol stocks of His tagged PRA-CH, *E. coli* BL21(DE3) Gold frozen cells containing the Magik plasmid were thawed on ice for 10 minutes and 30 ∓L of cells were transformed with 3 ∓L of MT245 plasmid. The DNA/cell mixture was kept for 5 min on ice and incubated for 45 s in a 42°C waterbath. The content of the vial was plated on Luria-Bertani (LB) (1% Bacto-tryptone, 0.5% Bacto-yeast extract, 1% NaCl pH 7.0) agar plate supplemented with the appropriate antibiotics (0.1 mg/mL Amp, 0.05 mg /mL Kan). The plate was then placed in a 37°C incubator for 14-16 h. A single colony was selected from the plate, inoculated into 5 mL LB medium supplemented with the appropriate antibiotics as previously outlined and allowed to grow for 14-16 h with gentle shaking at 37°C. Glycerol stocks were prepared by mixing well 1 mL of culture with 0.5 mL sterile 50% glycerol, and stored at -80°C. For preparing glycerol stocks of the non-His tagged PRA-CH, *E. coli* BL21(DE3) Gold Magik cells were freshly transformed with 0.5 ∓L pLB1_2 plasmid by selection on LB/Kan + Amp plates incubated at 37°C for 14-16 h, and the procedure continued as described above.

2.1.2.2 Expression of Recombinant Hexa-His Tagged PRA-CH

After testing the expression of recombinant PRA-CH on a small scale, the production was scaled to 1 L. A stab of glycerol stock was used to initiate an overnight culture of LB/Amp + Kan grown at 37°C. The 50 mL culture was added into 1 L of the same medium and grown with shaking (250 rpm) at 37°C for an additional 2 h. When the culture reached an OD_{600nm} of 0.9, protein production was induced by the addition of 0.1 mM IPTG. The culture was then incubated for 20 h at room temperature with shaking. Cells were harvested by centrifugation at 4000 g for 20 min and 4°C, then resuspended in 45 mL of lysis buffer (see 2.1.1 for buffer components). The cells were lysed by sonication (Heat Systems/Ultrasonics) using a short probe at 4°C, with 3 cycles at a setting of 70% intensity, 30 s on followed by 30 s off. The cell lysate was cleared by centrifugation at 150,000g for 30 min and 4°C, resulting in 45 mL of crude extract.

2.1.2.3 Purification of Recombinant Hexa-His Tagged PRA-CH

The crude extract was briefly incubated at room temperature with gentle shaking with 2 mL of DEAE-Sepharose anion exchange buffer in **binding buffer** (see 2.1.1 for buffer components). The supernatant was collected then incubated with 5 mL of Ni-NTA affinity resin in **binding buffer** and gently agitated as a batch for 90 min at 4°C. The protein/resin slurry was transferred into an empty Bio-Rad Econo-column (2.5 cm ID) and the flow-through was collected. The resin was washed thoroughly with 200 mL of **washing buffer** and the recombinant protein was eluted at room temperature with 20 mL

of **elution buffer**. CHELEX 100 resin was added to the eluted fractions in order to prevent protein precipitation by Ni^{2+} leaching from the Ni-NTA resin. A total of 60 mg of recombinant protein was isolated from 1 L of cell culture.

A small scale test was performed to determine the appropriate conditions for removal of the His tag from the full-length PRA-CH recombinant protein. The hexa-His tagged protein (1 mg/mL in Ni-NTA elution buffer) was incubated with five different dilutions of human-alpha thrombin (0.1 mg/mL in 50 mM Tris-HCl (pH 7.5), 500 mM NaCl). Thrombin to protein ratios were varied between 1/250 and 1/4000 (v/v). The incubation was performed at room temperature for 30 min and 2 h, respectively, and samples were then analyzed by SDS-PAGE. Optimal conditions were determined to be 1/4000 dilution (v/v thrombin to protein) for 2 h at room temperature.

Thrombin treated PRA-CH was further subjected to buffer exchange with **buffer A** (50 mM Tris-HCl (pH 7.5)) by using a Centricon-10 ultrafiltration unit (Amicon). The protein was purified further with a Poros50 D anion exchange column (1.66 mL) previously equilibrated with buffer A, attached to the AKTA protein separation system (Pharmacia). After sample application, the column was washed with 10 column volumes of buffer A at a flow rate of 4 mL/min. PRA-CH was eluted with a linear gradient of 0-800 mM NaCl in buffer A.

Fractions enriched with recombinant protein were identified by SDS-PAGE, pooled, concentrated and subjected to size exclusion chromatography on a prepacked Superose

12 HR 10/30 column. Protein was eluted with 1.5 column volumes of 50 mM Tris (pH 7.5), 200 mM NaCl. After different stages of purification, the protein was stored at 4°C until analyzed by dynamic light scattering (DLS) measurements.

2.1.2.4 Preparing the DNA Construct for Wild-Type (Non-His Tagged) PRA-CH

Amp and Kan were used at 0.1 mg/mL and 0.05 mg/mL, respectively. A QIAprep Spin Miniprep kit was used to purify up to 10 ŦL of MT245 plasmid DNA from a 5 mL overnight culture of MC1061 *E. coli* strain harboring MT245. The activity in U/mL of all the enzymes used in the molecular biology experiments are listed in section 2.1.1. To isolate the *hisI* gene from the pET-15b vector, 10 ŦL of MT245 plasmid DNA in 10 mM Tris-HCl (pH 8.5) were incubated for 90 min at 37°C with 2.5 ŦL of *BamH I*, 2.5 Ŧl of *Nde I*, 5 ŦL of NE buffer for *BamH I*, 5 ŦL of BSA (1 mg/mL), and 10 ŦL deionized distilled water (dH_2O). The restriction digest was run on a 2% agarose gel. The band corresponding to the gene was manually excised and purified using a QIAex II agarose gel extraction kit, then eluted into 20 ŦL of 10 mM Tris-HCl, pH 8.5.

Meanwhile, a pET vector not encoding a His tag (pET-11a from Novagen) was prepared for insertion of the *hisI* gene. One ŦL of pET-11a in 10 ŦL of 10 mM Tris-HCl buffer (pH 8.5) was digested with 0.5 ŦL *BamH I* and 0.5 ŦL Nde I for 180 min at 37°C, in the presence of 1.5 ŦL NE buffer for *BamH I*, 1.5 ŦL BSA (1 mg/mL), and 1 ŦL dH_2O. The mixture (15 ŦL) was treated with 0.5 ŦL alkaline phosphatase, 5 ŦL *C/AP* buffer and

29.5 ŦL dH_20, to remove the phosphate groups from the two ends of the cut pET-11a vector, and incubated at 37°C for 30 min. An additional 0.5 ŦL of alkaline phosphatase was added to the mixture followed by incubation for another 30 min. The DNA was purified using the QIAquick PCR purification kit, and the excised pET-11a was eluted with 100 ŦL 10 mM Tris-HCl (pH 8.5).

Next, the *hisI* gene was ligated into the dephosphorylated vector by mixing together 10 ng of pET-11a, 20 ng *hisI* gene (1.5 mL insert:1.5 mL dH_2O; 0.5 mL insert:2.5 mL dH_2O), 0.5 ŦL *T4 DNA ligase* and 0.5 ŦL buffer for *T4 DNA ligase* (10x), overnight at 16°C. *E. coli* strain MC1061 was transformed with the resulting plasmid construct by incubating together for 30 min at 4°C. The mixture was heated at 42°C for 45 s then placed on ice for 2 min. The transformation mixture was plated on LB/amp medium and incubated overnight at 37°C.

Recombinant plasmid DNA (*hisI* in pET-11a) denoted pLB1_2, chosen as a result of the screening, was purified with Qiagen PCR purification kit. The insert was verified by restriction analysis by using *NdeI* and *BamHI*. The DNA was then resolved on an agarose gel and yielded two bands, one corresponding to the excised inset and the other to the vector. As expected, two fragments of 419 bp and 5639 bp were identified. As a further confirmation digestion of MT245 DNA with *AvaI* yielded two fragments of 2568 bp and 3490 bp. The size of the fragments are as expected for the *AvaI* digest, which cleaves at two positions, one inside the *hisI* gene, and the other inside the vector.

2.1.2.5 Expression of Recombinant Wild-Type (Non-His Tagged) PRA-CH from pLB1_2

Amp and Kan were used at 0.1 mg/mL and 0.05 mg/mL, respectively. After a successful small scale expression test of recombinant PRA-CH, the enzyme was produced in 1 L Circle Grow (CG) culture, as described next. A 50 mL cell culture prepared from the glycerol stock, was incubated overnight at 37°C and then transferred to 1 L of CG/Amp + Kan. After 2 h of growth to an OD_{600nm} of 0.9, the production of the enzyme was induced by the addition of 1 mM IPTG. The culture was incubated for 20 h at room temperature with shaking (250 rpm). Cells were then harvested by centrifugation at 4000 g for 20 min.

A small scale test was performed to identify the appropriate components for the lysis buffer, notably the concentration of Zn^{2+} ions. A range of 0 to 5 mM $ZnCl_2$ was tried. After assessing protein expression of the cell lysate by SDS PAGE, the following lysis buffer was selected: 50 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 0.5 mM $ZnCl_2$, 0.1% (w/v) Triton X-100, 5% (w/v) glycerol, 20 mM BME and CompleteC-EDTA free protease inhibitors (one tablet per 50 mL). Cell lysis was performed by ultrasonication, as previously described. The cell lysate was clarified by centrifugation (150,000g, 30 min, 4°C) yielding 50 mL of crude extract.

2.1.2.6 Purification of Recombinant Wild-Type (Non-His Tagged) PRA-CH

The crude extract was incubated at room temperature with 2mL of DEAE Sepharose previously equilibrated in lysis buffer, with gentle mixing, to remove highly anionic material. The supernatant (26 mL) was diluted with 104 mL of **buffer B** (10 mM $NaPO_4$ pH 7.5, 5 mM BME) then loaded onto a hydroxyapatite column (24 mL bed volume). The column was washed with 2 volumes of buffer B, and the protein was eluted with a linear gradient from 10 mM to 400 mM $NaPO_4$ pH 7.5, 5 mM BME (15 volumes) at a flowrate of 1 mL/min. Fractions were analyzed for protein by SDS-PAGE (15% acrylamide).

Appropriate fractions were pooled together, and after the buffer was exchanged to **buffer C** (50 mM Tris-HCl (pH 7.5), 5 mM BME), the sample was loaded onto a Mono Q HR 5/5 column (0.96 mL bed volume) at a flow rate of 1 mL/min. The column was washed with 2 volumes of buffer C, and the protein was eluted with a linear gradient of 0.2 - 1 M NaCl in buffer C (15 volumes). Fractions were analyzed by SDS-PAGE (15% acrylamide).

Fractions enriched with the protein were pooled and loaded onto a Superdex 75 HR 10/30 (24 mL bed volume) previously equilibrated with buffer C containing 200 mM NaCl. The protein was eluted in this buffer. Native and SDS-PAGE were performed on the enzyme at various stages of purification as described previously.

2.1.2.7 PAGE and Protein Determination

Protein samples were subjected to SDS-PAGE analysis on discontinous Laemmli gels (42) with a 15% (w/v) resolving gel. Phosphorylase b (94 kDa), albumin (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20 kDa) and ↦-lactalbumin (14 kDa) (Amersham Pharmacia) served as references for molecular weight estimation. A nondenaturing gel of 10 % acrylamide was used to analyse the quaternary structure of the PRA-CH. Gel electrophoresis was performed using a BioRad electrophoresis unit. Proteins were visualizedd with Coomassie brilliant blue G-250.

Protein concentration was determined using the BioRad protein assay kit using a Cary 3E UV-visible spectrophotometer from Varian instruments. Bovine serum albumin (BSA) served as reference standard.

2.1.2.8 Dynamic Light Scattering

DLS experiments were performed using a DynaPro-801 DLS (Protein Solutions, Charlottesville, VA) instrument. The laser wavelength used was 790 nm and measurements were taken at a scattering angle of 90 degree. The sample solution was centrifuged and filtered through an inorganic alumina matrix membrane filter (0.01 ∓m) (Whatman) before recording measurements.

DLS measurements were taken after every chromatography step using a protein sample of 1 mg/mL and a 12 ŦL quartz cuvette. Data were analyzed with Dynamics 3.3 software (29) supplied by Protein Solutions.

2.1.2.9 Mass Spectrometry

MS was performed using a triple-quadrupole SCIEX electrospray mass spectrometer (PE SCIEX, Thornhill, ON) with a needle potential of 80 V and a flow rate of 2 ŦL/min. Positive ion detection was used in all cases, scanning over a m/z range of 500-2000. After acquisition, the molecular weights of the protein were obtained by using the deconvolution software, MacSpec 3.3. The protein sample was prepared by dialysis against 5% acetic acid and 20% methanol and concentrated using a Centricon YM-30 ultrafiltration unit to 1 mg/mL.

2.2 Results and Discussion

2.2.1 His Tagged PRA-CH

PRA-CH was expressed with a removable His tagged in order to facilitate its purification from the cell extract. The expression of the protein in the soluble fraction was high. SDS-PAGE of the crude cell extract (Figure 2.1 A, lane 1) showed a major band at about 17.5 kDa, which is in agreement with the predicted mass of the hexa-His tagged monomeric enzyme (17.46 kDa). The next step, DEAE Sepharose chromatography, allowed the

removal of highly anionic material which often binds to proteins and impedes their crystallization (D. Christendat, personal communication). As expected, the protein did not bind to the resin (see Figure 2.1 A, lane 2) since a buffer containing 500 mM NaCl was used. Ni^{2+} affinity chromatography (250 mM imidazole wash) yielded two protein bands of molecular weights 17.5 and 34 kDa (Figure 2.1 A, lane 3). The total amount of the His tagged PRA-CH that eluted from Ni-NTA column with 250 mM imidazole was of about 50 mg protein as determined from a Bio-Rad protein assay. The protein was then treated with thrombin as previously described in 2.1.2.3, and then analyzed on gel electrophoresis. Denaturing PAGE (15% acrylamide) resolved two major protein bands at 15.5 kDa and 31 kDa (Figure 2.1 A, lane 7). The mobility of both bands increased after thrombin treatment of the protein (compare Figure 2.1 A, lanes 6 and 7), consistent with the removal of the His tag. Thrombin treated protein resolved in lane 7 was transferred to a PVDF membrane for amino acid sequencing.

N-terminal sequencing (BRI, DNA Sequencing Facility) indicated that the proteins associated with each of these two major bands commenced with **NH_2-GSH-MIKSKGD**. GSH correspond to the residues from the engineered thrombin site, and indicate that the His tag was removed by cleavage. The results are consistent with the presence of monomeric and dimeric forms of the same protein resolved on the denaturing gel (Figure 2.1 A, lane 7).

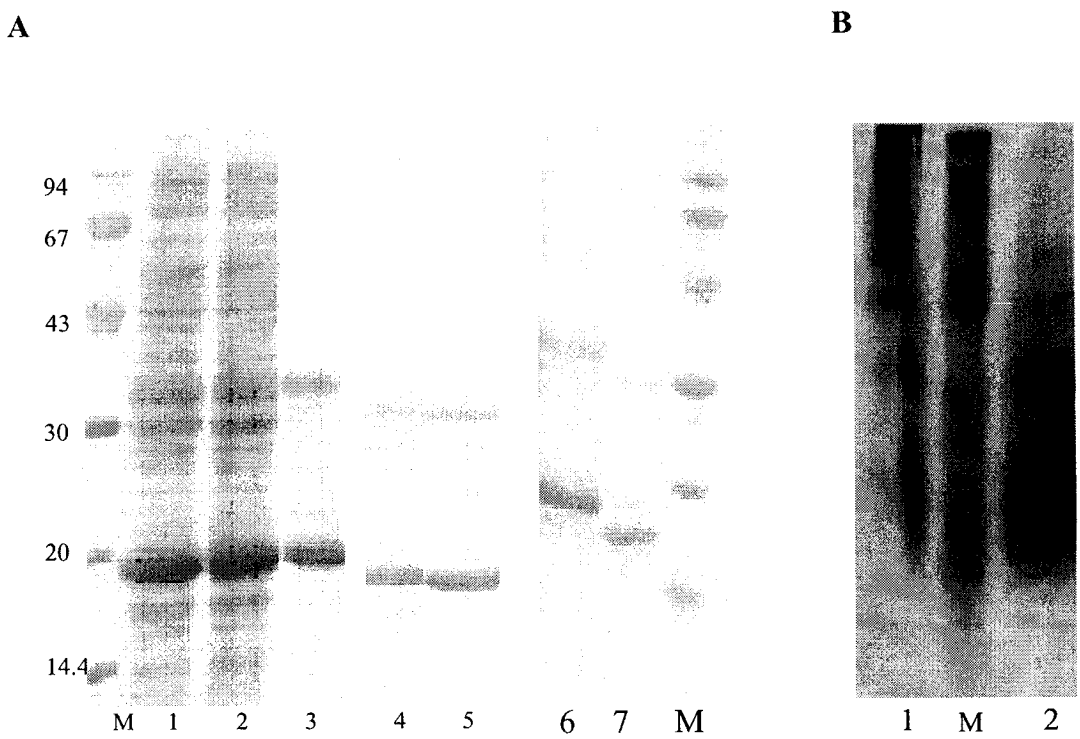


Figure 2.1. PAGE of selected fractions of PRA-CH during various stages of purification. **A SDS PAGE**: Lane M, protein molecular weight markers; lane 1, crude extract; lane 2, DEAE-Sepharose; lane 3, Ni-NTA, 250 mM imidazole; lane 4, Poros 50D after thrombin treatment; lane 5, Superose 12 HR 10/30 after Poros 50D; lane 6, Ni-NTA 250 mM imidazole, lane 7, thrombin-treated PRA-CH. **B Native PAGE**: lane 1, thrombin-cleaved PRA-CH (fraction 18 from Superose 12 HR 10/30); lane 2, thrombin-cleaved PRA-CH after incubation with 0.5 mM DTT; lane M, protein molecular weight markers.

DLS (see section 1.6.2) was performed at different stages of purification in order to assess the homogeneity of the protein.

In order to eliminate the apparent aggregation forms, the protein was further subjected to anion exchange and gel exclusion chromatography. The 250 mM imidazole wash from Ni-NTA resin was thrombin-treated as previously described. The sample was loaded onto a Poros 50D anion exchange column (see section 2.1.2.3 for experimental details). The protein eluted in a single peak at 380 mM NaCl (data not shown) and was heterogenous as determined by DLS, and denaturing PAGE (Figure 2.1 A, lane 4 shows two bands). The protein was then loaded on a prepacked Superose 12 HR 10/30 gel filtration column. The elution pattern monitored at OD_{280nm}, revealed a large peak followed by a minor peak with two shoulders (Figure 2.2). However, denaturing PAGE showed once more the presence of the two bands in both peaks (data shown for fraction 18 only) (Figure 2.1 A, lane 5).

The dynamic light scattering experiments performed on fractions 12 and 18 from Superose 12 HR (data shown for fraction 18 only) (Figure 2.3), also confirmed the presence of multiple forms of the enzyme. A polydispersity coefficient of 12.1 nm and an average radius of 17.5 nm were obtained. The relative polydispersity (polydispersity/RH) value was calculated to be 69% and indicated the presence of multiple species (the upper limit for monodisperse samples is 25% (29)). The estimated molecular weight for the species present in the solution was 3005 kDa. Thus, the experiment confirmed the presence of a broad distribution of large aggregates (see 1.6.2).

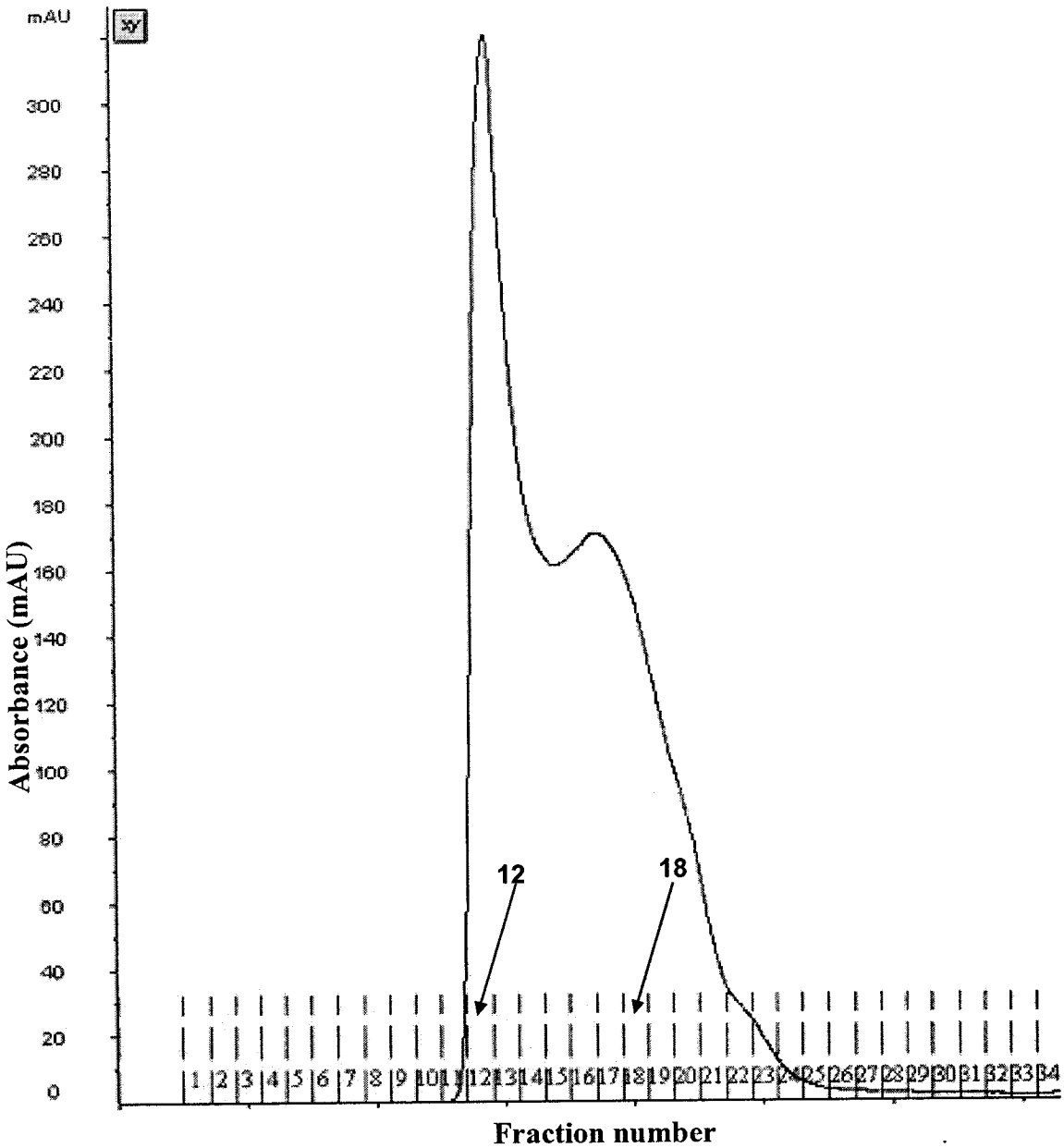


Figure 2.2. Size exclusion chromatography elution profile of PRA-CH from a Superose 12 HR 10/30 column. See section 2.1.2.3 for experimental details.

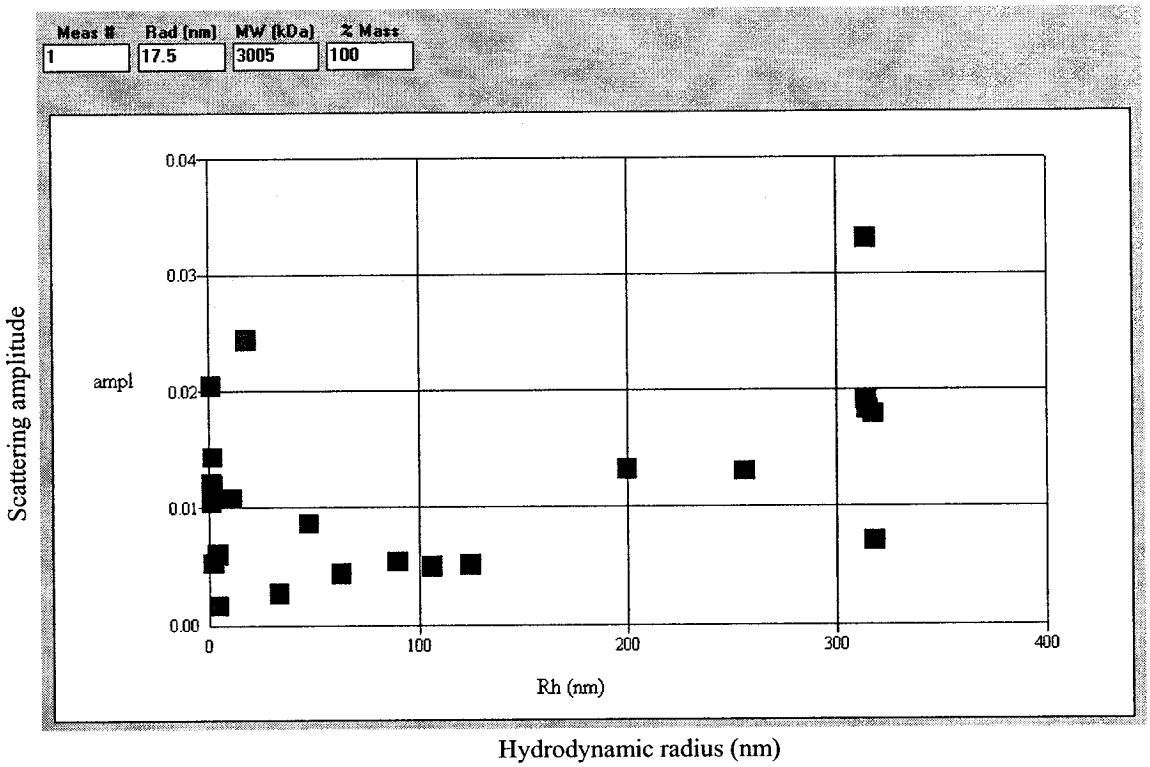


Figure 2.3. DLS measurements performed on fraction 18 collected from Superose 12 HR 10/30 column. RH represents the hydrodynamic radius, that is the z-average size value for the molecules in motion, and the Y-axis represents the scattering amplitude. Each data point represents a specific species of the protein in solution.

Interestingly, after incubating fraction 18 for 30 min at 4°C with 0.5 mM DTT, a change in the electrophoretic mobility of the protein on a nondenaturing gel was noted (Figure 2.1 B, lanes 1 and 2). There was either a decrease in the molecular weight of the species/oligomers of PRA-CH, a change in the isoelectric points of these species, or a combination of these two factors that occured upon treating the enzyme with the reducing agent. The DTT treated protein was analyzed by DLS and was shown to be heterogenous (data not shown).

In spite of the fact that the protein in the preparation was not homogeneous, a crystal screen was performed (see later discussion). No crystals were obtained.

2.2.2 Non-His Tagged PRA-CH

As previously mentioned, PRA-CH from *M. thermoautotrophicum* was initially expressed as a His tagged protein to take advantage of the one-step Ni^{2+} affinity purification. However, since the protein did not crystallize, and due to the difficulty of obtaining homogeneous PRA-CH, a new construct encoding *M. thermoautotrophicum* PRA-CH was designed. This new strategy was based on the findings of D'Ordine et al (20).

D'Ordine et al developed a purification protocol for non His-tagged PRA-CH from *M. vannielii*, which yielded an enzyme preparation suitable for functional characterization. Their protocol included chromatography on a DEAE Sepharose fast flow (FF) and

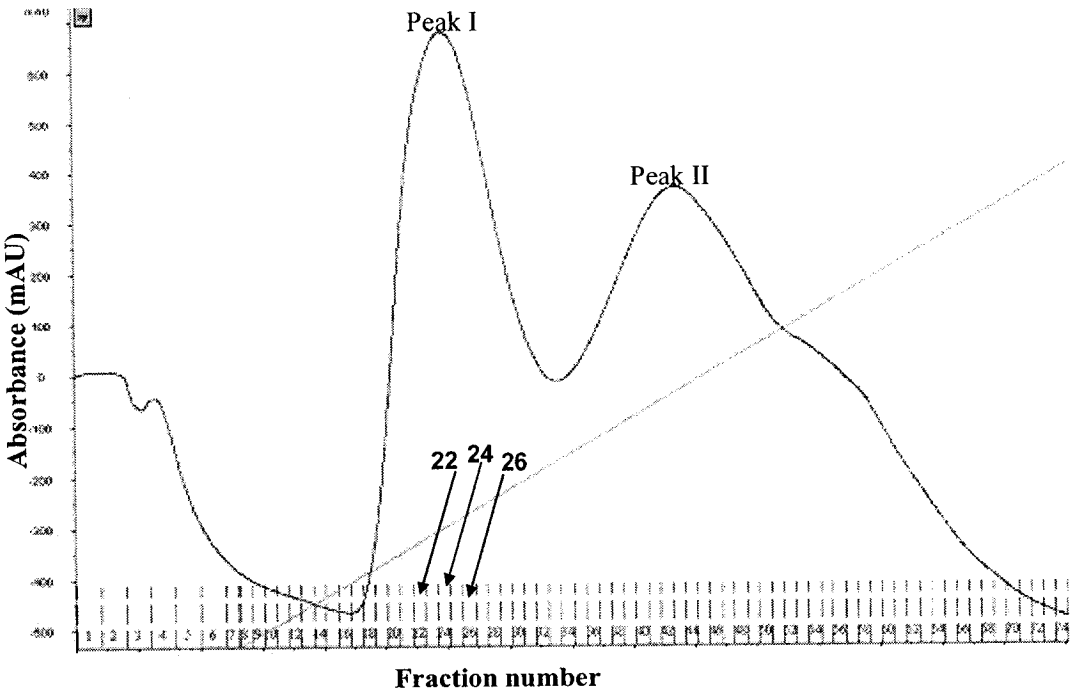
sulfopropyl (SP) Sepharose FF columns (linked in series), followed by a hydroxylapatite column. At this stage in the purification, analytical ultracentrifugation revealed the presence of two species of 31,200 (86%) and 75,800 (14%). The protein in the final preparation, however, showed two minor and one major species that together constituted more than 95% of the isolated protein as determined by scanning densiometry analysis of the denaturing polyacrylamide gel. Protein from all these three bands possessed the same N-terminal sequence and lacked an N terminal methionine. *M. vannielii* PRA-CH was also analyzed by gel filtration chromatography on a Superdex 200 HR 16/60 column. A major peak eluted at a volume that was consistent with a dimer (34 kDa) and exhibited PRA-CH activity, and a smaller amount eluted at a volume corresponding to a pentamer (79 kDa). In addition, atomic absorption technique indicated that the enzyme from *M. vannielii* contained a tightly bound Zn^{2+} per subunit (20). Sequence alignment of PRA-CH from *M. vannielii* and *M. thermoautotrophicum* suggests the existence of a metal cofactor in both enzymes (Figure 1.4, section 1.6.1).

Since D'Ordine et al found a tightly bound Zn^{2+} per PRA-CH subunit, our protein purification strategy was further modified by the addition of 0.5 mM $ZnCl_2$ to the lysis buffer; the protein precipitated at high concentrations (5 mM) of the metal. Initial purification trials of wild-type PRA-CH included chromatography using DEAE, hydroxyapatite, Mono Q HR ion exchange and Superdex 75 HR size exclusion. The crude extract was passed through DEAE column and then loaded onto a hydroxyapatite. The chromatogram from hydroxyapatite (Figure 2.4 A) showed a large peak (peak I, fractions 20-30), followed by a broad minor peak with a shoulder (peak II, fractions 32-

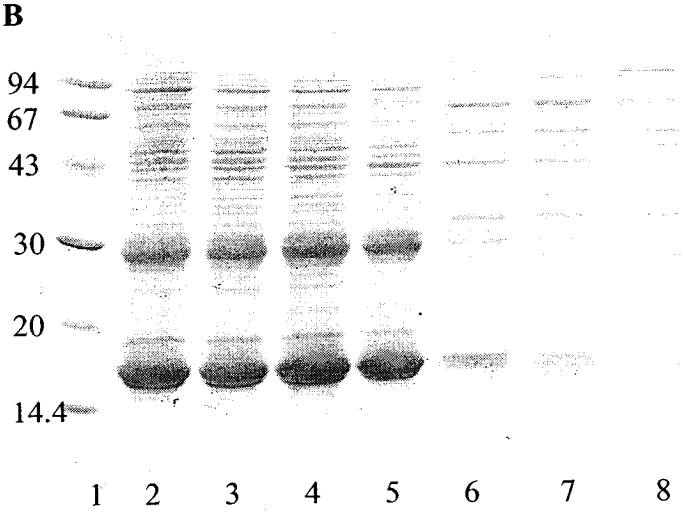
50). The SDS-PAGE of selected fractions from hydroxyapatite chromatography revealed the presence of low-purity PRA-CH mainly in peak I, in the form of two bands: 15.5 kDa and 31 kDa (Figure 2.4 B lanes 2-5). Note that peak II from the hydroxyapatite column contained only small amounts of PRA-CH (Figure 2.4 B, lanes 6-8). The protein from the major peak was further loaded onto Mono Q HR anion exchange resin. The elution pattern showed two major peaks (Figure 2.5 A), the first one at about 0.59 M NaCl (fractions 24-26), and the second one at 0.63 M NaCl (fractions 27-29). SDS-PAGE (Figure 2.5 B, lanes 10-13) confirmed the presence of two forms of PRA-CH (monomeric at 15.5 kDa and dimeric at 31 kDa) in both peaks from Mono Q chromatography (fractions 25-28 as resolved in lanes 10-13).

Figure 2.4. A Elution profile of PRA-CH from a Hydroxyapatite column (see 2.1.2.7 for details). Line represents the progression of linear gradient from 10 to 400 mM sodium phosphate during protein elution. **B SDS-PAGE** (15% acrylamide) analysis of the enzyme within peak I. Lane 1, molecular markers; lane 2, fraction 20; lane 3, fraction 22; lane 4, fraction 24; lane 5, fraction 26; lanes 6-8, fractions from the minor peak.

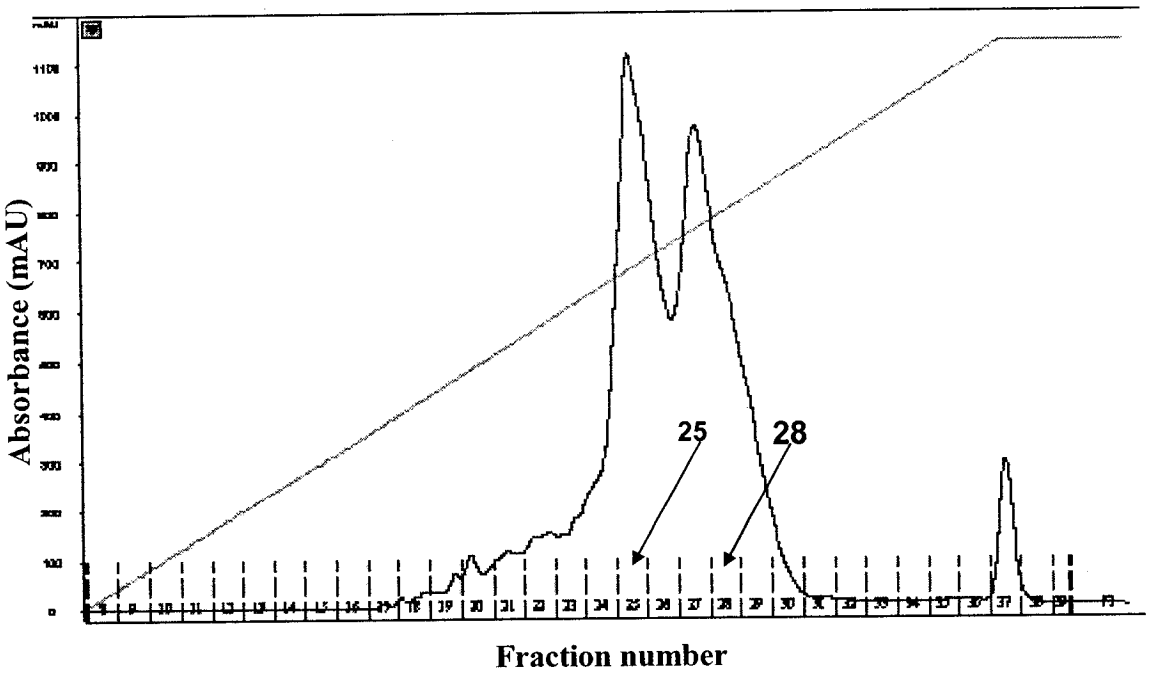
A


Peak I
Peak II
22
24
26
Absorbance (mAU)
Fraction number

B


94
67
43
30
20
14.4
1
2
3
4
5
6
7
8

A



B

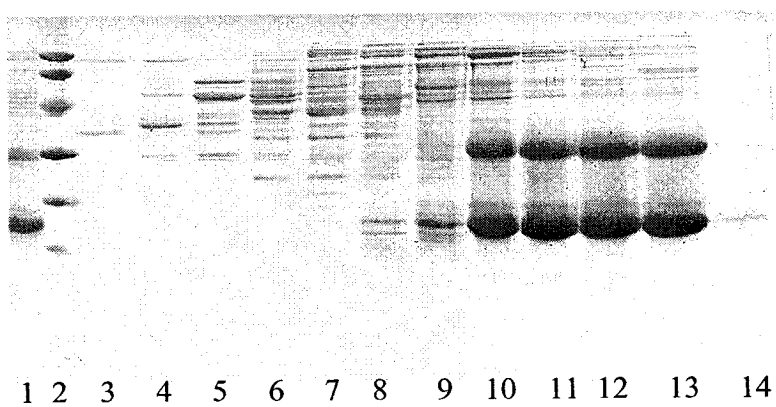


Figure 2.5. A Elution profile of PRA-CH on Mono Q anion exchange chromatography. Protein was eluted with a salt gradient from 0-1M NaCl. **B SDS-PAGE** analysis of the eluted protein from Mono Q. Lane 1, fractions 20-28 from Hydroxyapatite loaded on Mono Q; lane 2, protein molecular weight markers; lanes 3-9, fractions 18-24; lane 10-13, fraction 25-28; lane 14 is fraction 30.

Fractions 25 to 28 from Mono Q HR were pooled and subjected to Superdex 75 HR. The recorded chromatogram showed two peaks, a minor peak with a shoulder (peak I), followed by a major one (peak II) (Figure 2.6 A). Denaturing PAGE of fractions 14-18 (Figure 2.6 B, lanes 3-6) revealed the presence of a monomer (15.5 kDa) as well as a faint dimer (31 kDa). However, the non-denaturing gel electrophoresis performed on protein from Superdex 75 HR revealed the presence of two main bands that were partially resolved (Figure 2.6 C); the analysis of fraction 16 (lane 4), showed the presence of three bands, the upper one disappearing in lane 6 (fraction 18), where the presence of the lower band became more prominent. Calibration of the Superdex 75 HR with aldolase (158 kDa), ovalbumin (43 kDa), myoglobin (17 kDa) and vitamin B_{12} (1.35 kDa) (Pharmacia) served as reference for molecular weight estimation. The molecular weights of the two species were calculated using the relative migration of PRA-CH on Superdex 75 (Figure 2.6 A). Protein that eluted in fractions 14 to 18 (corresponding to peaks I and II, respectively) were analyzed using the calibration curve from Superdex 75. A clear separation was noted between the tetramer (62.64 kDa), corresponding to peak I, and dimer (35.68 kDa), corresponding to peak II.

DLS performed on the dimeric form of PRA-CH resolved with Mono Q HR (data not shown) and with Superdex 75 HR (Figure 2.7) still showed protein heterogeneity.

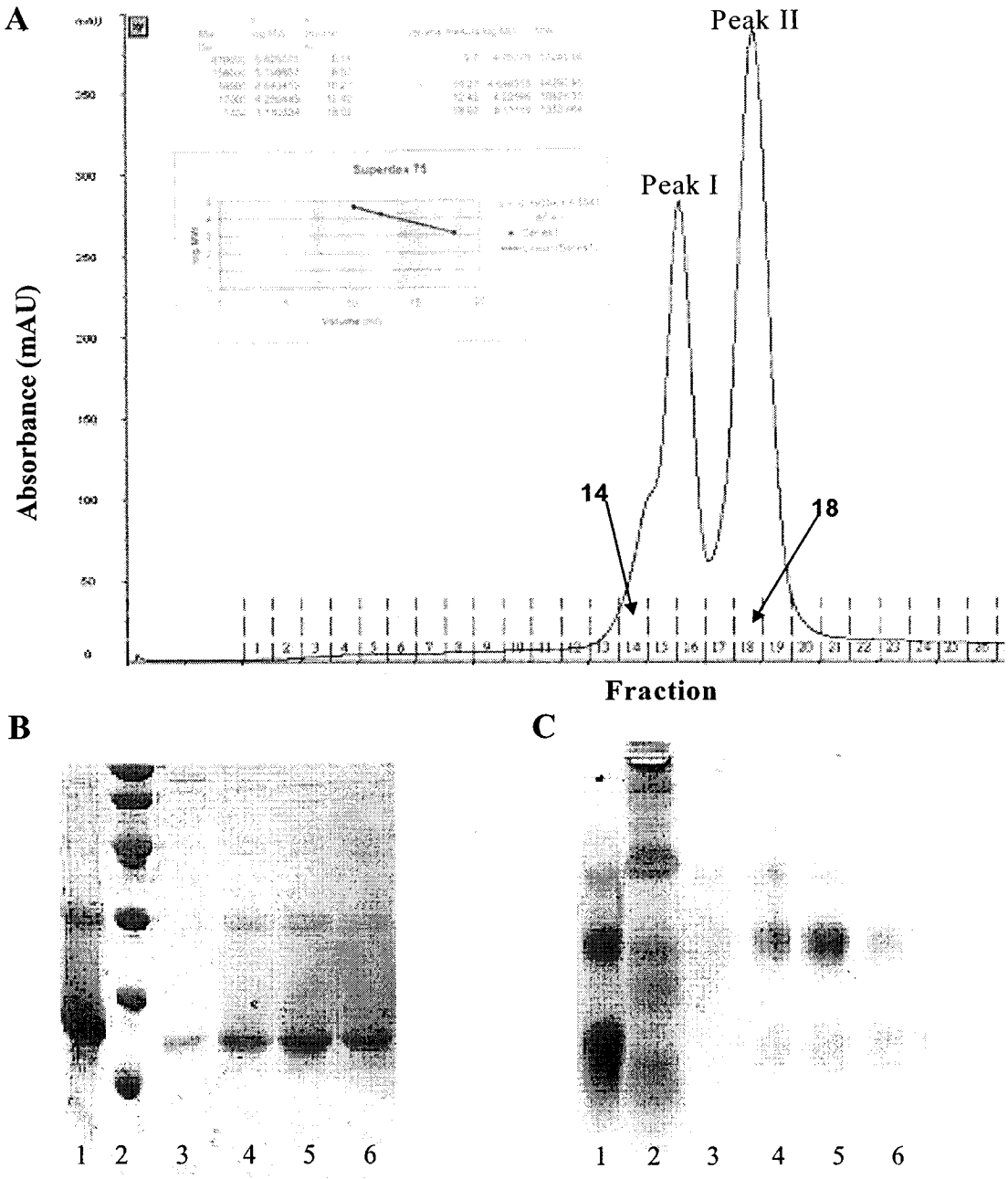


Figure 2.6. Elution profile of PRA-CH from Superdex 75 gel filtration **(A).** Chromatogram. **Inset**: Migration of protein standards as a function of molecular weight (MW) along with relative migration of the protein. **SDS PAGE (B),** and **native PAGE (C)** of selected fractions from the column. **A:** Lane 1, fractions 25-28 from Mono Q; lane 2, protein molecular weight markers; lane 3, fraction 14; lane 4-6, fractions 16-18.

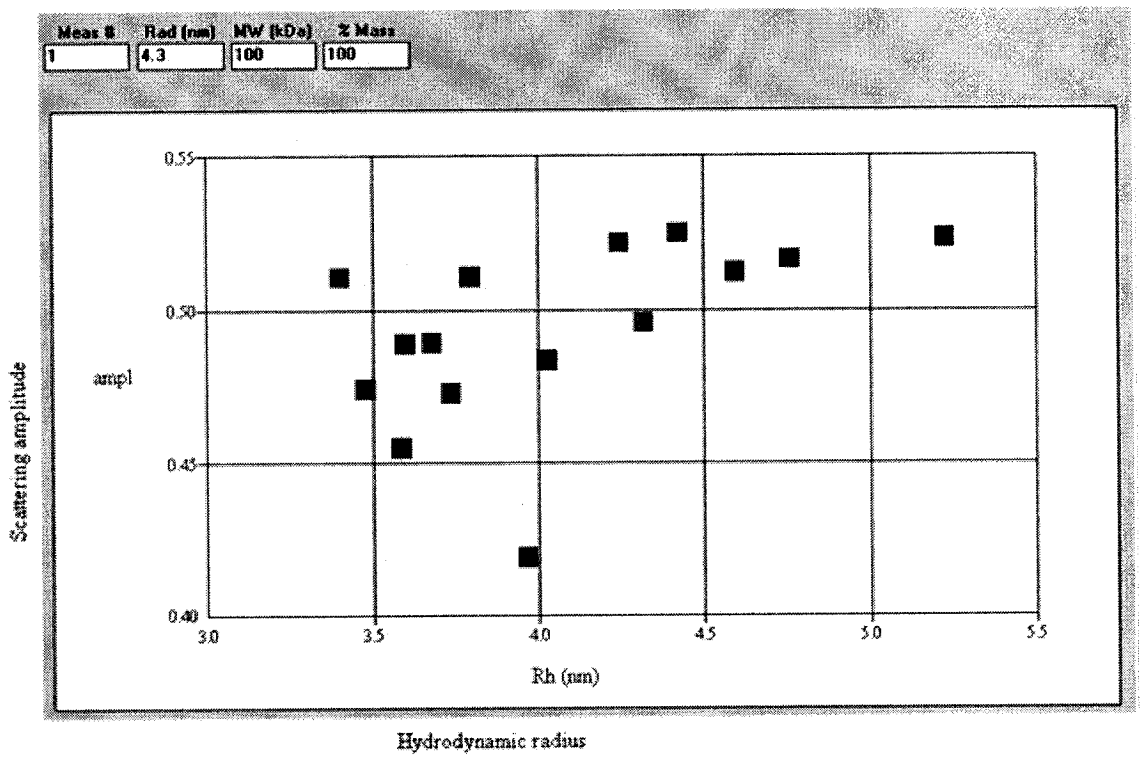


Figure 2.7. DLS measurements performed on PRA-CH collected from Superdex 75 HR 10/30 column. RH represents the hydrodynamic radius, and the Y-axis represents the scattering amplitude.

2.2.3 Thermal Stability of PRA-CH

Since recombinant PRA-CH was cloned from a thermophilic organism (7, 8), the use of a heat treatment step in the purification protocol was explored. Aliquots of crude extracts (0.5 mL in small eppendorf tubes) were incubated at different temperatures: 60, 80 and 90°C for 30 min and 60 min while being gently agitated from time to time. Heat-denatured proteins were removed by centrifugation at 40,000 g for 25 min and 4°C. The recovered supernatant was then analyzed on both denaturing and native-PAGE (Figure 2.8 A, B). Considerable purification of wild-type PRA-CH could be obtained by heating the crude extract (compare lane 1 with lanes 3-6). Native PAGE analysis of the supernatant derived from extract heated at 60, 80 and 90 °C for 30 min and 60 min showed the presence of one form of the protein (Figure 2.8 B, lanes 3-6). This form was a dimer as identified on Superdex 75 HR gel filtration column (Figure 2.6 C). SDS-PAGE revealed a major band corresponding to the monomer, and a faint band corresponding to the dimer (Figure 2.8 A, lanes 3-6).

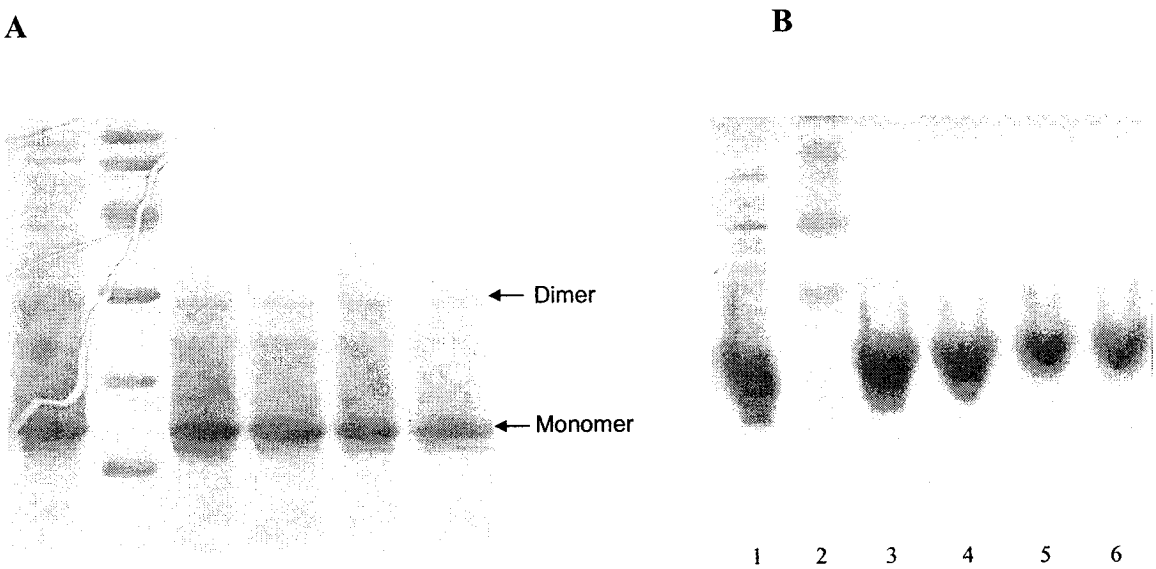


Figure 2.8. Heat-treated PRA-CH analyzed by **SDS PAGE (A)** and **native PAGE (B)**. **A** and **B**: lane 1, non heat treated crude extract; lane 2, protein molecular weight markers; lane 3, crude extract heated at 60°C, for 30 min; lane 4, heated at 60°C for 60 min; lane 5, heated at 80°C, for 30 min; lane 6, heated at 90°C for 30 min.

2.2.4 Final Purification Protocol

The final purification protocol for the native PRA-CH consists of heating the crude extract at 85°C for 15 min, then loading the supernatant onto a Mono Q HR 5/5 (0.96 mL bed volume) (see 2.1.2.7 for details). The protein eluted at a salt concentration of 0.3 M NaCl in one major peak (data not shown). A yellow pigment was removed at 0.6 M NaCl. The protein in this preparation revealed two species by SDS-PAGE (15% acrylamide) (Figure 2.9 A, lane 5) corresponding to a monomer (15.5 kDa) and dimer (31 kDa). Furthermore, only one band was visualizedd by native-PAGE (Figure 2.9 B, lane 6). Light scattering studies performed on a protein sample from Mono Q HR (fraction 25) verified the presence of a single species with a molecular weight of 43 kDa (Figure 2.10), identified as a dimer. The calculated relative dispersity for this sample was of 17 %, which is in accordance with a monodisperse protein (see 1.6.2). Remarkably, SDS-PAGE still showed monomeric and dimeric forms from the fraction that is near homogenous as determined by DLS and native PAGE; however, the band corresponding to the dimer was very faint. Protein was characterized by electrospray ionization mass spectrometry (ESI-MS) and a major peak corresponding to 15,445 Da was identified (Figure 2.11 inset), which was in reasonable agreement with a value of 15,458 Da predicted from the protein's primary sequence (Figure 1.4). Two other peaks corresponding to mass values of 15,317 Da and 15,630 Da were identified (Figure 2.11 inset). However, the identity of these other peaks was not persued. The differences in mass between the three species were not due to an amino acid replacement. This was confirmed from the crystal structure of the protein (see section 3.3).

Wild-type PRA-CH obtained from heat treatment followed by separation on Mono Q HR anion exchange chromatography was dialyzed against 50 mM Tri-HCl (pH 7.5), 200 mM NaCl, 0.5 mM $ZnCl_2$, 5 mM BME then concentrated using a Centricon YM-30 ultrafiltration unit to 10 mg/mL. The obtained protein sample was stored at 4°C. It was this PRA-CH preparation that was further subjected to crystallization trials.

Although a PRA-CH activity assay was developed by d'Ordine et al (20), an enzymatic assay for the activity of the protein in this study was not performed. The substrate for PRA-CH, N^1-(5`-phosphoribosyl)adenosine 5`-monophosphate (PR-AMP), is not commercially available.

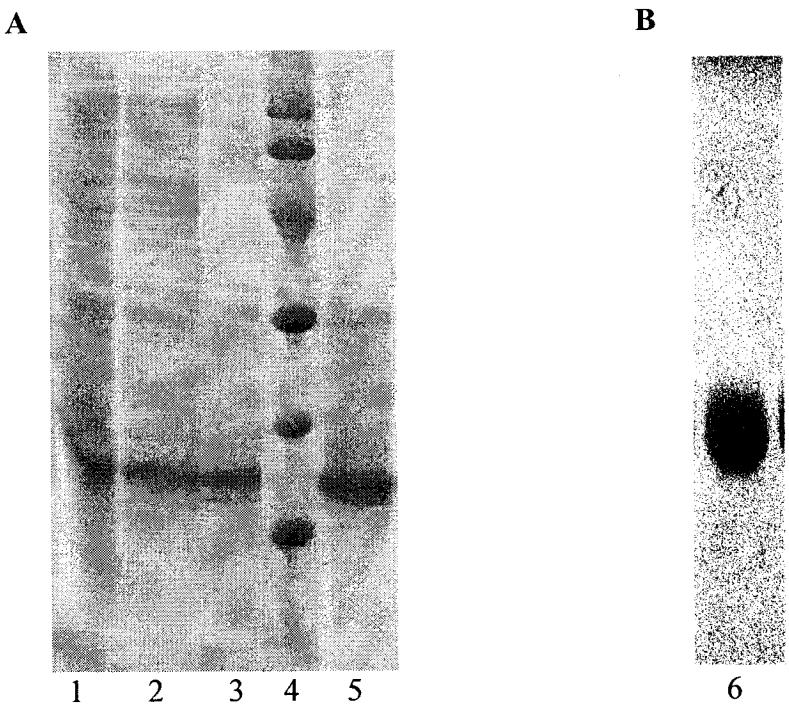


Figure 2.9. Final purification protocol of PRA-CH. **A SDS PAGE**: lane 1, crude extract; lane 2, DEAE Sepharose fraction; lane 3, heat-treated protein; lane 4, protein molecular weight markers; lane 5, fraction 25 from Mono Q HR. **B Native PAGE**: lane 6, fraction 25 eluted from Mono Q HR.

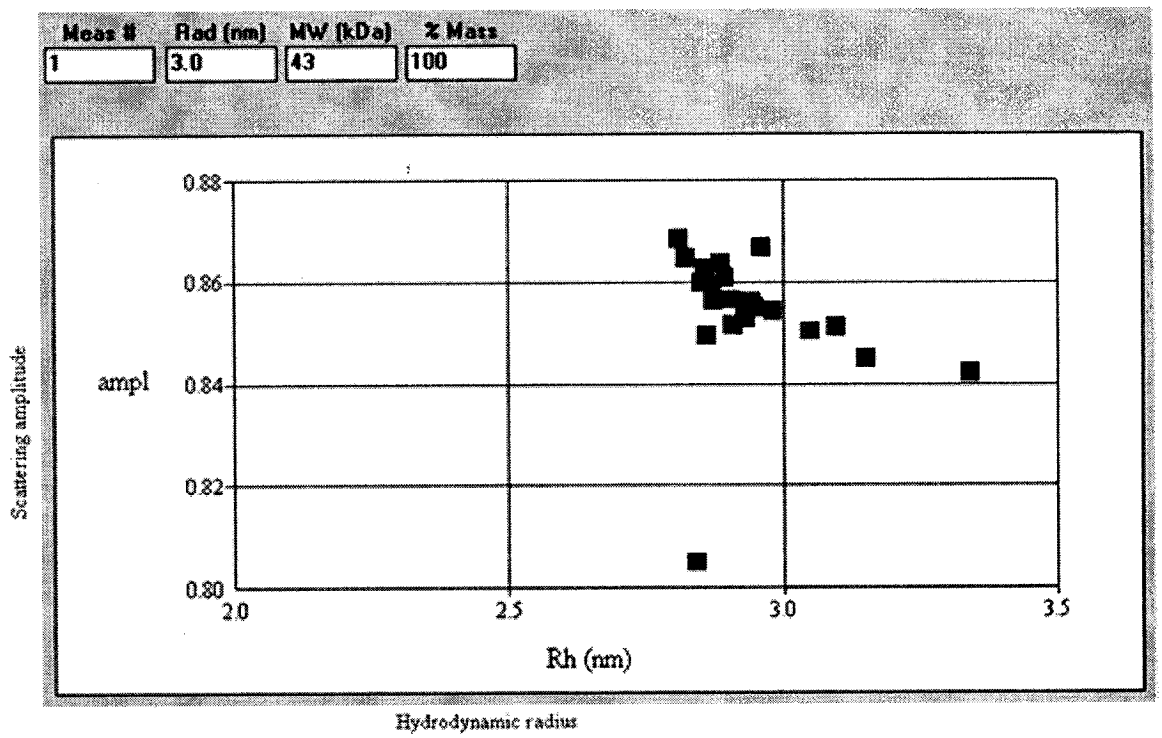


Figure 2.10. DLS measurements performed on fraction 29 collected from Mono Q anion exchange. The data suggest the presence of one form of the enzyme. RH represents the hydrodynamic radius, and the Y-axis represents the scattering amplitude.

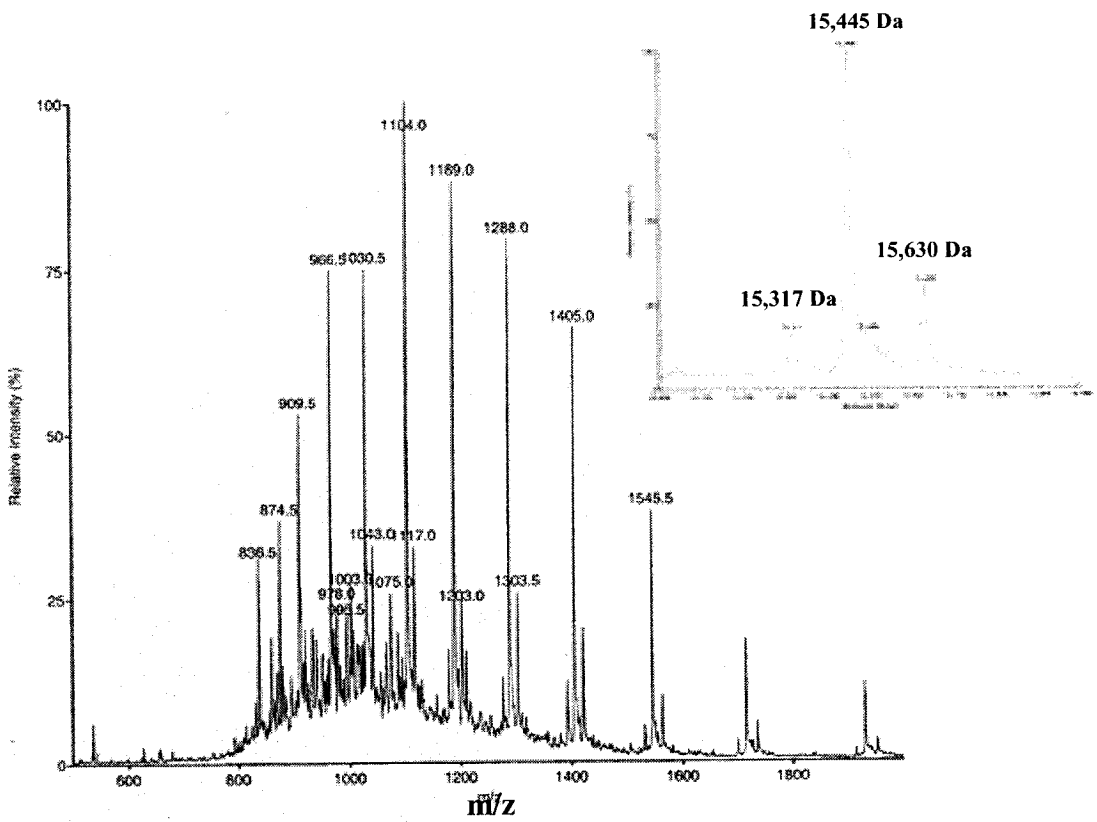


Figure 2.11. ESI mass spectrum of non-His tagged PRA-CH from Mono Q anion exchange. Shown is the charge envelope of the protein from m/z 800 to 1800. Deconvolution of this spectrum (see inset) yielded three peaks corresponding to mass values of 15,317, 15,445, and 15,630 Da.

Chapter 3 Crystallization and Preliminary X-ray Analysis of PRA-CH

3.1 Experimental Procedures

3.1.1 Materials

Sparse-matrix Hampton research screens I and II were purchased from Hampton Research, CA. Centricon and Centriprep YM-30 ultrafiltration units were obtained from Millipore, and L-SeMet from Sigma. All ingredients for preparing LeMaster medium (Appendix 1) (37) were purchased from Sigma.

3.1.2 Methods

3.1.2.1 Cloning, Expression and Purification of the Seleno-Methionine PRA-CH

E. coli strain DL41T7 yielding methionine auxotrophy was transformed with pLB1_2 DNA by incubating together for 30 min at 4°C. The mixture was heated at 42°C for 45 s then placed on ice for 2 min. The transformation mixture was plated on LB/amp medium and incubated overnight at 37°C in LeMaster's medium, in which the methionine was replaced by 30 mg/L L-SeMet. The seleno-methionine PRA-CH (SeMet PRA-CH) was prepared by transforming an *E. coli* strain carrying methionine auxotrophy (DL41T7) with pLB1_2 plasmid. The resulting strain was grown in LeMaster's medium in which

the methionine was replaced by 30 mg/L L-SeMet. Expression and purification of SeMet PRA-CH were performed as outlined for native PRA-CH (see Chapter 2).

3.1.2.2 Crystallization and X-Ray Diffraction of Wild-Type PRA-CH

Crystallization experiments were carried out at room temperature using the hanging-drop vapour-diffusion method (30). Initial crystallization screening was performed using the sparse-matrix Hampton research screen I and II. Drops were prepared by mixing 1.5 ∓L of the protein solution with 3 ∓l of the reservoir solution, were allowed to equilibrate against 500 ∓L of reservoir solution. Microcrystals appeared within two days when equilibrating against 0.5 mL of buffer #34 from screen II (100 mM HEPES-HCl (pH 7.5), 50 mM cadmium sulfate, and 1 M sodium acetate). Further optimization of parameters such as pH (5.6 to 8.5) and different concentrations of buffer (50 to 100 mM), precipitant (0.1 to 2 M), and $CdSO_4$ (2 to 50 mM) led to the final crystallization conditions: 4 ∓L of protein solution of approximately 2 mg/mL were equilibrated against 4 ∓L of 100 mM HEPES-HCl (pH 7.5), 50 mM $CdSO_4$, 1.6 M CH_3CO_2Na, 10 % (v/v) glycerol. Of particular note was the fact that crystallization was achieved with an unusually low protein concentration.

A single crystal was picked up directly from the drop in a nylon loop, and frozen at 100K in a N_2 (gas) cold stream (Oxford Cryosystems, Oxford, UK). In-house testing of the crystals was done on an R-AXIS IIC area detector mounted on a RU-300 rotating anode generator (Rigaku, Tokyo), operating at 50 kV and 260 mA.

Synchrotron data were measured at beamline X8C, NSLS, Brookhaven National Laboratory at a wavelength of 1.0091 Å. A total of 180 frames were collected with an oscillation angle of 1° per frame. The unit-cell parameters were determined by the autoindexing and parameter-refinement procedures of DENZO. All intensity data were integrated and reduced with DENZO-SCALEPACK (43). The space group was determined by examining the intensity distribution of the X-ray data.

3.1.2.3 Screening for the Derivative crystal

Heavy atom derivatives were screened by soaking crystals in a solution of mother liquor and the heavy atom to be screened (Table 3.1). Stock solutions of 0.1 M in 50 mM HEPES-HCl pH 7.5 were prepared, then PRA-CH crystals were soaked in 100 ∓L reservoir solution containing 0.5-5 mM of the heavy atom solution. The concentration of the heavy atom and the duration of the soak were varied in an attempt to optimize conditions to produce a derivative (30). To test whether a crystal soaked in a heavy atom solution was a successful derivative a few X-ray images were collected from the crystal soaked in a particular condition.

Crystallization experiments for Se-Met PRA-CH were carried out in a similar way to those performed for the native PRA-CH. Initial crystallization screening was conducted using the sparse-matrix Hampton research screen I and II. Drops were prepared by mixing 1.5 ∓L of the protein solution with 3 ∓L of the reservoir solution, and then they were allowed to equilibrate against 500 ∓L of reservoir solution.

Transition metals	**Lanthanides**
Platinum tetrachloride	Samarium acetate
Thimerosal	Terbium (III) acetate
Methyl mercuric acetate	Ytterbium (III) chloride hexahydrate
Mersalyl acid	Europium acetate
Potassium hexachloro-osmate (IV)	
Potassium hexachloro-iridate	
Tungstic acid	

Table 3.1. Test of heavy atoms used in order to obtain derivative PRA-CH crystals suitable for MAD experiments.

3.2 Results and Discussion

Using the sparse-matrix sampling kit II, crystalline precipitates were obtained with sodium acetate as precipitating agent. Higher quality single crystals were obtained when the protein concentration was dramatically decreased and the precipitant concentration increased. The presence of both cadmium and sulfate ions appeared to be crucial and even reduction of the Cd^{2+} concentration destabilized the crystals. Best results were obtained when adding 5-10 % glycerol. Under these conditions, well shaped crystals grew at room temperature within two weeks (Figure 3.1).

A complete X-ray data was collected up to 1.76 Å resolution from a native crystal. The crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 39.777, b = 54.360, c = 117.365 Å, a = ß = ? = 90°. Assuming the presence of two molecules of wild-type PRA-CH in the asymmetric unit the calculated volume (*Vm*) of 2.04 $Å^3Da^{-1}$ is well within the normal range of 1.6-3.6 $Å^3Da^{-1}$ typical for globular proteins (17). This corresponds to a solvent content of approximately 40%. From a total of 149 679 reflections, there were 23 827 unique reflections with I > _(I) in the resolution range of 50.0-1.76 Å. The final R_{merge} on all data was 0.043 and the completeness of the data was 91.3 %.

Since Zn^{2+} was shown to be required for enzymatic activity of PRA-CH from *M. vannielii* (20), a close homolog of *M. thermoautotrophicum*, 0.5 mM $ZnCl_2$ was added to

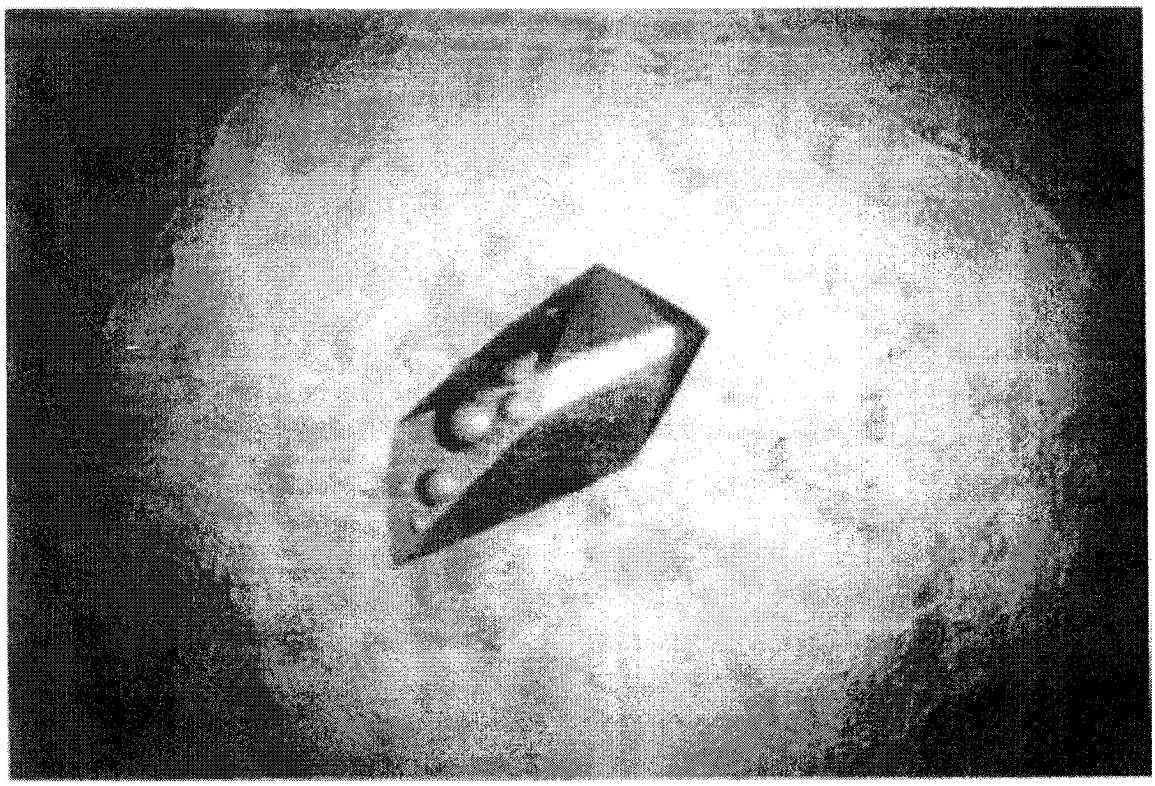


Figure 3.1. Photograph of PRA-CH crystal taken under the microscope.

some protein solutions used in the crystallization trials in attempts to stabilize the enzyme. However, the presence or absence of Zn^{2+} had no effect on the crystallization of the protein. X-ray fluorescence measurements from crystals grown from these solutions around the Zn^{2+} absorption edge (ℓ = 1.282312 nm) indicated the presence of Zn^{2+} in the crystal. Attempts to obtain MAD phases using data collected at the absorption maximum, the inflection point, and high energy remote wavelengths around this edge failed despite the readily apparent anomalous signal in the diffraction data.

3.3 Recent Progress

Using the enzyme purification protocols and the crystallization conditions developed in this thesis, the structure of PRA-CH was then solved by Drs. Joe Schrag and Jayaraman Sivaraman at the Biotechnology Research Institute, Montreal, by single-wavelength anomalous dispersion method (SAD) (44). Unpublished data of the structure are courtesy of Drs. Schrag and Sivaraman.

The asymmetric unit of the crystal comprises a dimer. Each monomer is composed of four beta hairpins (Figure 3.2). On one side of the monomer, hairpins 1 (residues 23-42) and 3 (residues 76-98) pack in an antiparallel manner, forming a four stranded mixed ↕-sheet. Packed against one face of this ↕-sheet are hairpin 2 (residues 54-69) and the helix that is inserted between hairpins 1 and 2. The dimer interface is formed by interactions of hairpins 2, 3 and 4 of each monomer (Figure 3.2). The ↕-sheets formed by hairpins 2 and

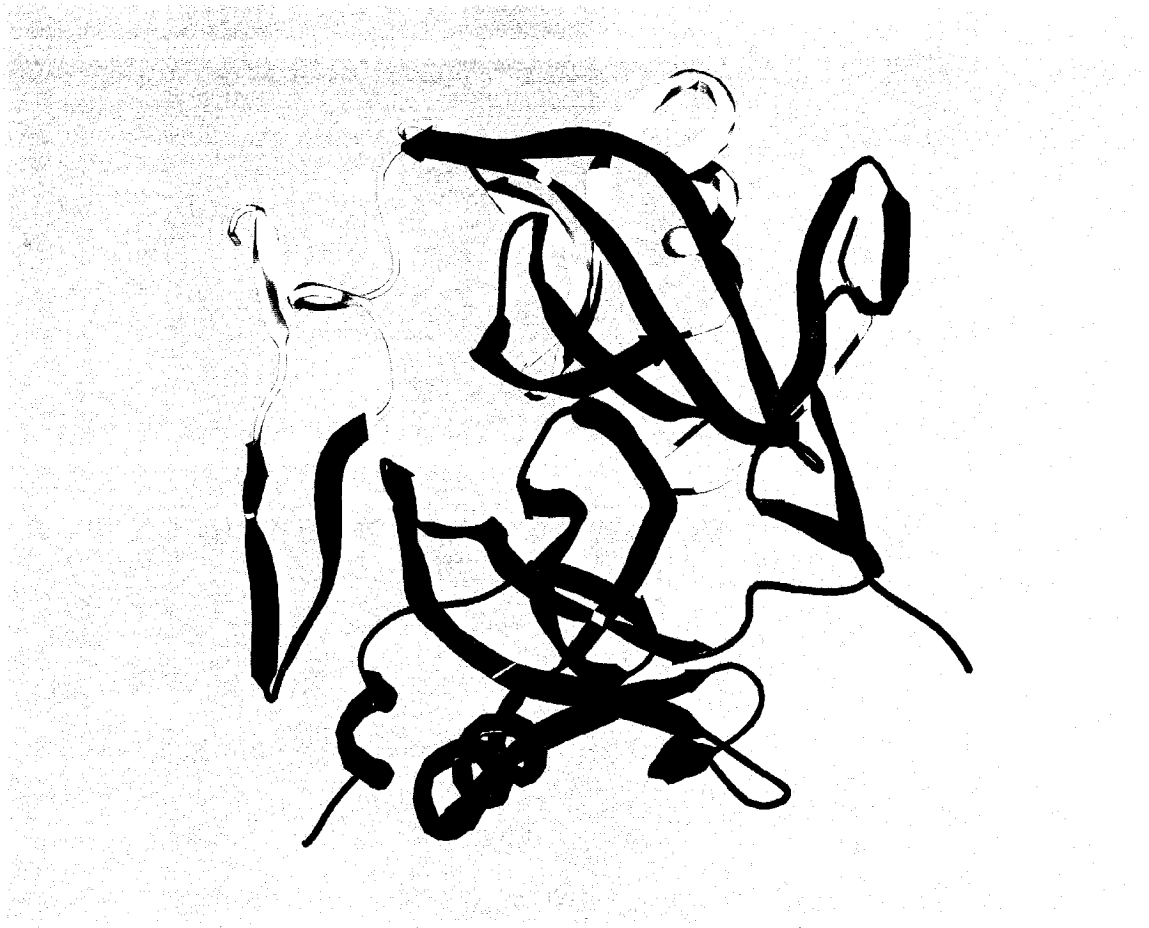


Figure 3.2. Ribbon diagram of the PRA-CH homodimer from *M. thermoautotrophicum.* Both monomers are formed of 4 beta-hairpins. One monomer is colored in magenta, the other is colored as follows: hairpin 1 is red, hairpin 2 is yellow, hairpin 3 is green, hairpin 4 is blue.

3 pack face to face, primarily by interactions of hydrophobic residues, with a large percentage being the branched residues valine, isoleucine and leucine. Hairpin 4 (residues 111-122) packs against hairpin 3 of the other monomer of the dimer via hydrogen bonding characteristic of parallel beta structure. This interaction extends the mixed ↕-sheet from four to six strands. The total surface area buried in the dimer interface is of about 3100 $Å^2$, more that 40% of the total molecular surface area of each monomer.

Two surface clefts are formed at the dimer interface (Figure 3.3), with four Cd^{2+} bound in each of these surface clefts. A negative potential results from a triad of aspartic acid residues (residues 85, 87 and 89). The carboxylate oxygen atoms of these residues coordinate one of the Cd^{2+} ions bound in the cleft. Adjacent to the acidic cluster is a cluster of three cysteine residues, Cys102 and Cys109 from one monomer and Cys86 from the other monomer. The sulphur atoms coordinate a second Cd^{2+} ion. A third Cd^{2+} ion is bound between the acidic and cysteine clusters and is coordinated by Asp85, Cys109 and Cys86. The fourth Cd^{2+} ion is bound to residues His16 and Asp89. The presence of the Cd^{2+} ions in the crystal contacts explain the necessity of having no less than 50 mM $Cd(SO_4)_2$ in the reservoir solution solution during protein crystallization. Two Cd^{2+} ions bind between symmetry related molecules. The first is coordinated by His A103 and Cys B86 from the symmetry related molecule. The second is bound by Cys A86 and His B103 from one molecule and by Glu B123 from the symmetry related molecule.

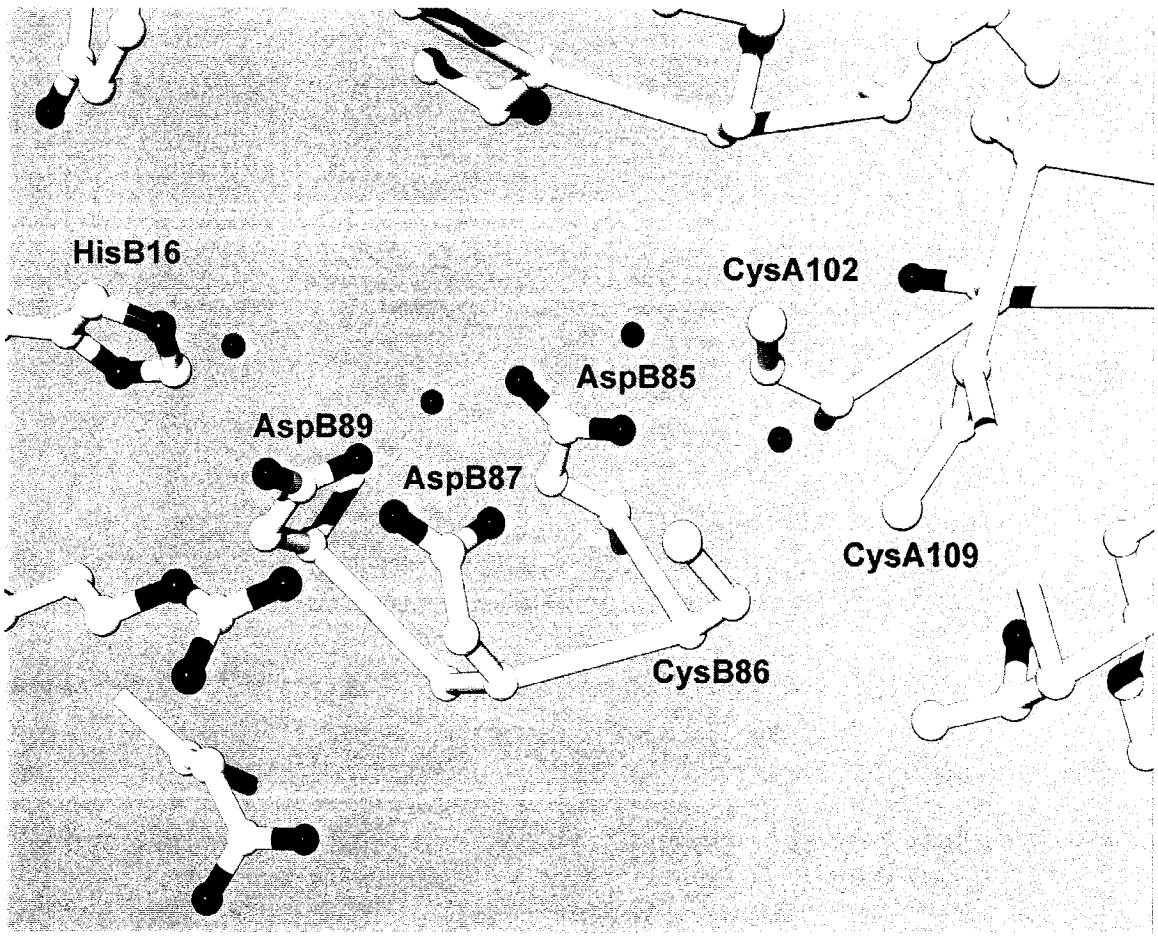


Figure 3.3. Cd^{2+} ions binding in the surface cleft of PRA-CH structure. The cysteine residues from both monomers are required to form the proposed Zn^{2+} binding site.

The three dimensional structure of the *M. thermoautotrophicum* PRA-CH showed that the three cysteine residues, Cys 86, Cys 102, and Cys 109 form a metal binding site. The equivalent Cys residues in the sequence motif $C^{93}(X)_{15}C^{109}H^{110}(X)_5C^{116}$ from *M. vannielii* were suggested to be the Zn^{2+} ligands. Mutations of these three cysteine residues in *M. vannielii* PRA-CH confirmed that they bind the Zn^{2+} ion and that are required for activity (20). Therefore, the enzyme requires dimerization to form the active site.

Although the crystal structure for PRA-CH was solved in the absence of the substrate, N^1-(5`-phosphoribosyl)adenosine 5`-monophosphate (PR-AMP), our results establish a venue for investigating the active site by molecular modelling studies and further crystallization trials with a substrate analog or with product. Such studies will aid in elucidating the catalytic mechanism of this enzyme.

REFERENCES

(1) Gaasterland, T (1999). *Curr. Op. Microbiol. 2*, 542-547.

(2) Thauer, R. K. (1998). *Microbiology 144*, 2377-2406.

(3) Edmonds, C.G., Crain, P.F., Gupta, R., Hashizume, T., Hocart, C.H., Kowalak, J.A., Pomerantz, S.C., Stetter, K.O., McCloskey, J.A. (1991). *J. of Bacteriology 173*, 3138-3148.

(4) Woese, C.R., Kandler, O., Wheelis, M.L. (1990). *Proc. Natl. Acad. Sci. USA 87*, 4576–4579.

(5) Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., Wolfe, R. (1979). *Microbiol. Rev. 43,* 260–296.

(6) Hoehler, T. M., Alperin, M. J., Albert, D. B., Martens, C. S. (1994). *Glob. Biogeochem. Cycles 8*, 451.

(7) Lange, M., Ahring, B. K. (2001). *Microbiol. Rev. 25*, 553-571.

(8) Zeikus J.G., Wolfe, R.S. (1972). *J. of Bacteriology 109*, 707-713.

(9) Heinen, W. (1971). *Arch. Mikrobiol. 76*, 2-17.

(10) Smith, D.R., Doucette-Stamm, L.A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Reeve, J.N. (1997). *J. Bacteriol. 179*, 7135–7155.

(11) Haas, F., Mitchell, M.B., Ames, B.N., Mitchell, H.K. (1952). *Genetics 37*, 217-226.

(12) Winkler, M.E. (1987). *Am. Soc. Microbiol 1*, 395-411.

(13) Carlomagno, M.S., Chiariotti, L., Alifano, P., Nappo, A.G., Bruni, C.B. (1988). *J. Mol. Biol. 203*, 585-606.

(14) Donahue, T.F., Farabaugh, P.J., Fink, G.R. (1982). *Gene 18*, 47-59.

(15) Alifano, P., Fani, R., Lio, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M.S., Bruni, C.B. (1996). *Microbiol. Rev. 60*, 44-69.

(16) Kishore, G.M., Shah, D.M. (1988). *Annu. Rev. Biochem. 57*, 627-663.

(17) Mori, I., Pfister-Fonne, R., Matsugana, S., Tada, S., Kimura, Y., Hatano, M., Koizumi, S. I., Scheidegger, A., Hayakawa, K., Ohta, D. (1995). *Plant Phisiol. 107*, 719-723.

(18) Fani, R., Pietro, L., Lazcano, A. (1995). *J. Mol. Evol. 41*, 760-774.

(19) Smith, D. W. E., Ames, B. N. (1965). *J. Biol. Chem. 240*, 3056-3063.

(20) D'Ordine, R. L., Klem, T. J., Davisson, V. J. (1999). *Biochem. 38*, 1537-1546.

(21) Fani, R., Alifano, P., Allotta, G., Bazzicalupo, M., Carlomagno, M. S., Gallori, E., Rivellini, F., Polsinelli, M. (1993). *Res. Microbiol. 144*, 187-200.

(22) Beckler, G. S., Reeve, J. N. (1986). *Mol. Gen. Genet. 204*, 133-140.

(23) Crane, D. I., Gould, S. J. (1994). *Curr. Genet. 26*, 443-450.

(24) Legerton, T. L., Yanofsky, C. (1985). *Gene 39*, 129-140.

(25) Corpet, F. (1988). *Nucl. Acids Res. 16 (22),* 10881-10890.

(26) Vallee, B. L., Auld, D. S. (1990). *Proc. Natl. Sci. U.S.A. 87*, 220-224.

(27) Vallee, B. L., Auld, D. S. (1990). *Biochem 29*, 5647-5659.

(28) Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. A., Nordheim, A., Stunnenberg, H. G. (1991). *Proc. Natl. Sci. U.S.A. 88*, 8972-8976.

(29) DLS: Protein Solutions. (1998). *Software Manual Dynamics, Version 4.0*, Charlottesville, VA.

(30) McPherson, A. (1990). *Eur. J. Biochem, 189*, 1-23.

(31) Branden, C., Tooze, J. (1999). *Garland Science Publishing*, 2nd edition.

(32) Branden, C., Tooze, J. (1990). *Nature 343*, 687-690.

(33) McPherson, A, Kuznetov, Y. G. (1995). *Structure 3*, 759-768.

(34) Blundell, T. L., Johnson, L. N. (1976). *Protein Crystallography*, London.

(35) Blow, D. M., Rossmann, M. G. (1961). *Acta Cryst. 14*, 1195-1202.

(36) Karle, J. (1980). *Int. J. Quantum Chem. Symp. 7*, 357-367.

(37) Hendrickson, W. A., Horton, J. R., LeMaster, D. M. (1990) *EMBO J.* 9, 1665–1672.

(38) Gonzales, A., Pedelacq, J. D., Sola, M., Benini, S. (1999). *Acta Cryst. D55*, 1449-1458.

(39) Dauter, Z. (2002). *Acta Cryst. D58*, 1958-1967.

(40) Zhang, S., Zubay, G., Goldman, E. (1991). *Gene 105*, 61-72.

(41) Kim, R., Sandler, J., Goldman, S., Yokota, H., Clarke, A.J. and Kim, S.H. (1998). *Biochem. J. 330*, 295-302.

(42) Laemmli UK. (1970). *Nature 227*, 680–685.

(43) Otwinowski, Z., Minor, W. (1997). Processing of X-ray diffraction data, *Academic Press*, New York.

(44) Terwillier, T. C. (2002). *Acta Cryst. D58,* 1937-1940.

APPENDIX 1

Recipe for Defined LeMaster Medium (11).

i) Autoclavable Portion

The following chemicals were added to one container. All amino acids are the L-enantiomer.

	Chemical	for 2 L (g)		Chemical	for 2 L (g)
1	alanine	1.00	14	serine	4.16
2	arginine HCL	1.16	15	threonine	0.46
3	aspartic acid	0.80	16	tyrosine	0.34
4	cystine	0.06	17	valine	0.46
5	glutamic acid	1.34	18	adenine	1.00
6	glutamine	0.66	19	guanosine	1.34
7	glycine	1.08	20	thymine	0.34
8	histidine	0.12	21	uracil	1.00
9	isoleucine	0.46	22	sodium acetate	3.00
10	leucine	0.46	23	succinic acid	3.00
11	lysine HCl	0.84	24	ammonium Cl	1.50
12	phenylalanine	0.26	25	NaOH	2.16
13	proline	0.20	26	K_2HPO_4	21.00

1. After measurement, the autoclavable portion of LeMaster medium should be stored dry in a freezer.
2. When LeMaster medium is needed, add 2 (6 or 10) L of dH2O. The solids will not all be dissolved.
3. Autoclave this "pre-medium".
4. Cool. Everything should now be in solution. Check that the medium is at ~pH 7.5. Adjust if necessary.
5. Distribute pre-medium to baffled flasks in 1 L amounts.
6. Autoclave again.
7. Cool. Make the non-autoclavable portion below.

ii) Non-Autoclavable Portion

The following chemicals were added to 200, 600 or 1000 mL deionized distilled H_2O.

Chemical	**200 mL dH_2O**
1) glucose	20 g
2) $MgSO_4$-$7H_2O$	0.5 g
3) $FeSO_4$-$7H_2O$	8.35 mg

1. Filter using a 0.22 μm filter system in 100 mL amounts.

2. Add 100 mL to each of the baffled flasks above.

3. Before inoculation add 50 mg DL SeMet to complete the Le Master medium to each 1 L baffled flask.