

The Role of the C-terminus
in *Bacillus subtilis*
Inorganic Pyrophosphatase
Activity

by

MONIKA ANNA KONOPKA

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Biosciences
The University of Birmingham
September 2002

SYNOPSIS

Inorganic pyrophosphatases hydrolyse inorganic pyrophosphate ensuring that the pool of phosphate is continuously available in a cell for subsequent metabolic reactions. *Bacillus subtilis* PPase is the first member of the Family II of PPases, unique not only in terms of physical/biochemical characteristics, but also enzyme primary structure.

Prompted by the plausible existence of two distinct evolutionary paths for development of this universal function in the cell and, consequently, distinct active sites, a research into the role of a signature sequence near the C-terminus of *B. subtilis* enzyme was commenced. Two truncated variants and a series of site specific mutants were produced and the effect of the mutations on enzyme activity, stability and multimerization was analysed.

Overall, there is no evidence for the C-terminus having a role in multimerization. However, both its conserved and non-conserved regions are absolutely vital for full enzyme activity. The signature sequence is found to be required for both the composition and correct conformation of the active site. Substrate specificity studies suggest the C-terminal domain is able to adopt a conformation permitting the binding of a larger substrate molecule. The work provides important insights into the nature of activity of Family II PPases that could facilitate design of drugs effective against a significant number of pathogenic microorganisms depending upon these enzymes for macromolecular syntheses.

Marcinowi i naszemu Maleństwu, które towarzyszyło mi
przez cały czas pisania tej pracy

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Tom W. Young for his help during the work on this project. I am grateful for his special attention and support during the writing stage, especially the final period of the fight against time before our baby's arrival. Also, I wish to acknowledge the financial support of the Adrian Brown Trust. Furthermore, I am grateful to Dr. Scott White for his assistance, especially with the structural part of the project, and Dr. Nicolas J. Kuhn for helpful discussions when setting up the enzymatic tests. I am also grateful to Peter Ashton from the School of Chemistry for the mass spectroscopy analysis.

My very special thanks go to Albert Wadeson, who guided me through the beginnings of the lab work, always ready to give help or advice, and inspired me with his patience, peace and openness to others. Many thanks to Cristina Sampayo for her friendship and stimulating energy and to other members of our small G3 community, John Osborne and Mohammad Ilias, for their help and discussions. I am also grateful to my colleagues from the School of Biosciences, for their friendship and support, especially to Lesley Griffiths and Tamara Belyaeva, for their smile and time to chat, all that made me feel more at home.

Finally, and most importantly, I thank my family for all their emotional support, encouragement, reminding me of the priorities and keeping things in the right perspective. It was my Mum in the first place who inspired me with her love to English language. Her faith in me never let me get disheartened. I am forever indebted to my husband for his support and encouragement in the most difficult moments. Marcin guided me to independence, helping to find the balance when everything seemed to go wrong and focus on what was the most important at the time. Above all, the final layout of this thesis would not be possible without his professional advice and help with \LaTeX and graphics programs.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 13 |
| 1.1 | Phosphates in Biological Systems | 13 |
| 1.1.1 | Phosphate Compounds | 13 |
| 1.1.2 | Role in Cell Metabolism | 14 |
| | Inorganic Pyrophosphate | 14 |
| | Polyphosphates | 19 |
| 1.1.3 | Role in Molecular Evolution | 22 |
| 1.1.4 | Enzymes of Phosphate Metabolism | 25 |
| | Active Site Metal Ions | 27 |
| 1.2 | Inorganic Pyrophosphatases | 28 |
| 1.2.1 | Classification | 28 |
| | Soluble Inorganic Pyrophosphatases | 28 |
| | Membrane-bound Inorganic Pyrophosphatases | 29 |
| 1.2.2 | Function in the Cell | 32 |
| | Soluble Inorganic Pyrophosphatases | 33 |
| | Membrane-bound Inorganic Pyrophosphatases | 33 |
| 1.3 | Family I Inorganic Pyrophosphatases | 35 |
| 1.3.1 | Classification | 35 |
| | Subfamily Ia, Prokaryotic PPases | 35 |
| | Subfamily Ib, Plant PPases | 36 |
| | Subfamily Ic, Animal/Fungal PPases | 37 |
| 1.3.2 | Conservation of Primary Structure | 37 |
| 1.3.3 | Catalytic Properties | 39 |
| 1.3.4 | Enzyme Structure | 40 |
| 1.3.5 | Catalysis | 41 |
| | Phosphate Binding Sites | 41 |
| | Metal Ions | 43 |
| | Nucleophilic Water | 44 |
| | General Acid | 46 |
| | The Mechanism | 46 |
| 1.4 | Family II Inorganic Pyrophosphatases – a New Line of Research | 48 |
| 1.4.1 | <i>Bacillus subtilis</i> PPase, the First Representative of a New Family | 48 |
| | <i>Bacillus subtilis</i> | 48 |
| | The PPase | 49 |
| 1.4.2 | Research Perspectives | 51 |
| | Aims of the Project | 54 |

| | | |
|----------|---|-----------|
| 2 | Materials and Methods | 55 |
| 2.1 | Bacterial Strains and Plasmids | 55 |
| 2.2 | General Protocols and Procedures | 56 |
| 2.3 | Laboratory Equipment of General Use | 58 |
| 2.4 | Materials, Solutions and Buffers | 59 |
| 2.4.1 | Suppliers | 59 |
| 2.4.2 | Reagents for General Use | 59 |
| 2.4.3 | Media | 60 |
| 2.4.4 | Isolation and Purification of Plasmid DNA | 61 |
| 2.4.5 | Restriction and Modifying Enzymes | 62 |
| 2.4.6 | Bacterial Transformation | 63 |
| 2.4.7 | Agarose Gel Electrophoresis | 63 |
| 2.4.8 | Protein Overexpression | 65 |
| 2.4.9 | SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) | 65 |
| 2.4.10 | Electrophoretic Transfer of Proteins (Western Blotting) | 68 |
| 2.4.11 | Immunodetection of Blotted Proteins | 68 |
| 2.4.12 | Blot Preparation for Protein Sequence Analysis | 70 |
| 2.4.13 | FPLC | 70 |
| 2.4.14 | Inorganic Pyrophosphatase Activity Assay | 71 |
| 2.4.15 | Protein Assay | 73 |
| 2.5 | Glycerol Stocks of Bacterial Strains | 73 |
| 2.6 | Isolation of Plasmid DNA | 74 |
| 2.6.1 | Small Scale Isolation ('Miniprep') | 74 |
| 2.6.2 | Large Scale Isolation ('Maxiprep') | 77 |
| 2.7 | Standard DNA Manipulation Techniques | 78 |
| 2.7.1 | Digestion with Restriction Endonucleases | 78 |
| 2.7.2 | Dephosphorylation | 79 |
| 2.7.3 | Ligation | 80 |
| 2.7.4 | Bacterial Transformation | 81 |
| 2.8 | Purification of DNA (Clean-up) | 82 |
| 2.8.1 | Purification of DNA Fragments from Solution | 83 |
| 2.8.2 | Recovery of DNA Fragments from Gels | 84 |
| 2.9 | Agarose Gel Electrophoresis | 84 |
| 2.10 | Polymerase Chain Reaction (PCR) | 85 |
| 2.10.1 | Generation of Truncated Coding Sequences | 86 |
| 2.10.2 | Site-specific Modification of Coding Sequence | 87 |
| 2.11 | Oligodeoxynucleotides (Primers) | 88 |
| 2.11.1 | Mutagenic Primers | 88 |
| 2.11.2 | Sequencing Primers | 89 |
| 2.12 | Construction of Mutant Proteins | 89 |
| 2.12.1 | Truncated Mutants | 89 |
| 2.12.2 | Site-Specific Mutants | 94 |
| 2.13 | DNA Sequencing | 95 |
| 2.14 | Protein Overexpression and Isolation | 95 |
| 2.15 | SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) | 96 |
| 2.15.1 | Sample Preparation | 97 |

| | | |
|--|--|------------|
| 2.15.2 | SDS-PAGE | 97 |
| 2.15.3 | Staining and Destaining SDS-Polyacrylamide Gels | 98 |
| 2.15.4 | Drying SDS-Polyacrylamide Gels | 99 |
| 2.16 | Electrophoretic Transfer of Proteins (Western Blotting) | 99 |
| 2.17 | Immunodetection of Blotted Antigens | 100 |
| 2.18 | Protein Sequence Analysis | 101 |
| 2.19 | Mass Spectrometry | 102 |
| 2.20 | Fast Protein Liquid Chromatography (FPLC) | 102 |
| 2.20.1 | Protein Purification. Ion Exchange Chromatography (IEC) | 102 |
| 2.20.2 | Native Molecular Mass Analysis. Size Exclusion | 104 |
| 2.20.3 | Cleaning of MonoQ Column | 105 |
| 2.21 | Inorganic Pyrophosphatase Activity Assay | 106 |
| 2.22 | Protein Assay | 107 |
| 2.22.1 | Bradford Method | 107 |
| 2.22.2 | BCA Method | 109 |
| 3 | The C-terminus is Essential for Enzyme Activity but Not Subunit Composition | 113 |
| 3.1 | Introduction | 113 |
| 3.2 | Generation of Truncated Mutants | 114 |
| 3.2.1 | PCR and Cloning | 114 |
| 3.2.2 | Overexpression, Protein Isolation and Immunostaining | 115 |
| 3.2.3 | Relative Subunit Molecular Mass | 116 |
| 3.2.4 | Purification | 120 |
| Full-length PPase and 5AATR | 120 | |
| 17AATR and <i>E. coli</i> Control | 122 | |
| 3.3 | Enzyme Activity | 127 |
| 3.3.1 | The Activity of 5AATR is Significantly Impaired | 128 |
| 3.3.2 | 17AATR is Not Active | 128 |
| 3.4 | Native Molecular Mass | 129 |
| 3.5 | Conclusions | 131 |
| 4 | Five Non-Conserved C-terminal Residues Are Crucial for Overall Enzyme Stability | 135 |
| 4.1 | Introduction | 135 |
| 4.2 | The 5AATR Mutant Retains the Main Characteristics of the Untruncated Enzyme | 136 |
| 4.2.1 | Activation by Mn ²⁺ | 136 |
| Preliminary Assays | 136 | |
| Mn ²⁺ Concentration Dependence of Enzyme Activation | 138 | |
| Protein Concentration-Dependent Mode of Activation | 141 | |
| Time-course of Activation | 141 | |
| 4.2.2 | Insensitivity to Fluoride and Inhibition by EDTA | 144 |
| 4.3 | The 5AATR Mutant is Unstable | 145 |
| 4.3.1 | Protein Degradation | 145 |
| 4.3.2 | Loss of Activity with Time | 145 |
| 4.3.3 | Loss of Activity with Increasing Concentration of MnCl ₂ | 147 |

| | | |
|----------|--|------------|
| 4.3.4 | Time-course of EDTA Inhibition | 149 |
| 4.4 | Computer Predictions | 149 |
| 4.5 | Activity on Tri-Polyphosphate | 151 |
| 4.6 | Conclusions | 154 |
| 5 | The Signature Sequence Preserves the Functionality of the Active Site | 156 |
| 5.1 | Introduction | 156 |
| 5.2 | Generation of Site-Specific Mutants | 159 |
| 5.2.1 | PCR and Transformation into Epicurian Coli Cells | 159 |
| 5.2.2 | Re-cloning | 160 |
| 5.2.3 | Overexpression and Protein Isolation | 160 |
| 5.2.4 | Relative Subunit Molecular Mass | 161 |
| 5.3 | Arg295 and Lys296. Direct Involvement in Catalysis | 161 |
| 5.4 | Ser294, Val299 and Val 300. Preservation of the Correct Folding Pattern . | 162 |
| 5.4.1 | Ser294 | 162 |
| 5.4.2 | Val299 and Val300 | 163 |
| 5.5 | Conclusions | 164 |
| 6 | Discussion | 172 |
| 7 | Future Perspectives | 185 |
| 7.1 | Spectroscopic Techniques | 185 |
| 7.1.1 | Circular Dichroism (CD) | 185 |
| 7.1.2 | Fourier Transform Infrared Spectroscopy (FTIR) | 186 |
| 7.1.3 | Fluorescence Spectroscopy | 187 |
| 7.2 | Isothermal Titration Calorimetry (ITC) | 189 |
| 7.3 | Kinetic Studies | 190 |
| | List of References | 191 |
| 8 | Publications | 205 |

List of Figures

| | | |
|------|---|-----|
| 1.1 | Inorganic pyrophosphate (PP _i) | 15 |
| 1.2 | Polyphosphates (Poly P _(n)); <i>n</i> is the number of phosphoryl groups in the polyP molecule | 15 |
| 1.3 | Overall Fold of Yeast and <i>E. coli</i> PPases Monomers | 42 |
| 1.4 | Structure of the Yeast PPase (PPase:Pi ₂) Active Site | 45 |
| 1.5 | The Proposed Catalytic Mechanism for Family I PPases | 47 |
| 1.6 | Multiple Sequence Alignment of the First Family II Members | 52 |
| 2.1 | The Map of pET11c Vector | 57 |
| 2.2 | Scheme of the <i>B. subtilis</i> Full-length PPase, 5AATR and 17AATR Clones with Their Restriction Map | 93 |
| 2.3 | P _i Standard Curve | 108 |
| 2.4 | Bradford Protein Standard Curve | 110 |
| 2.5 | BCA Assay Protein Standard Curve | 112 |
| 3.1 | Truncated PCR Products | 117 |
| 3.2 | The C-terminus of Truncated <i>B. subtilis</i> PPase Mutants | 117 |
| 3.3 | Overexpression of Truncated <i>B. subtilis</i> PPase Mutants | 118 |
| 3.4 | SDS-PAGE of Supernatant and Pellet Material After Sonication Following Overexpression | 119 |
| 3.5 | Immunostaining of <i>B. subtilis</i> PPase and Truncated Mutants | 119 |
| 3.6 | TOF ES Mass Spectrum of 5AATR | 121 |
| 3.7 | Purification of <i>B. subtilis</i> Untruncated PPase and 5AATR Mutant | 124 |
| 3.8 | Elution Profile of 17AATR on MonoQ Ion-exchange Column | 126 |
| 3.9 | Purification of 17AATR | 126 |
| 3.10 | Elution Profiles of <i>E. coli</i> and 17AATR on MonoQ Ion-exchange Column | 134 |
| 4.1 | Effect of Mn ²⁺ Concentration on Activation of <i>B. subtilis</i> PPase and 5AATR Mutant at 57 nM Protein Concentration | 139 |
| 4.2 | Effect of Mn ²⁺ Concentration on Activation of <i>B. subtilis</i> PPase and 5AATR Mutant at 7 nM and 12 nM Protein Concentration | 140 |
| 4.3 | Time Course of Activation by Mn ²⁺ | 143 |
| 4.4 | Degradation of 5AATR Mutant Protein upon Incubation at 4°C | 146 |
| 4.5 | Inactivation of 5AATR Untreated with Manganese upon Incubation at 4°C | 148 |
| 4.6 | Time Course of Inactivation by EDTA | 150 |
| 4.7 | Computer Predictions of the Protein Secondary Structure at the C-termini (22 residues) of Family II PPases | 152 |
| 4.8 | Effect of Enzyme Dilution on Activity Against Tri-Polyphosphate | 153 |

| | | |
|-----|---|-----|
| 5.1 | Sequence Alignment of the C-termini of Family II PPases | 157 |
| 5.2 | PCR of Site Specific <i>B. subtilis</i> PPase Mutants | 167 |
| 5.3 | Re-cloning of Site-Specific Mutant Fragments: Digestion with <i>SacI</i> and <i>BamHI</i> of a Vector and Mutant PCR Products | 168 |
| 5.4 | Sequence Data Obtained from an Arg295Ala Clone Using T7 Terminator Primer | 169 |
| 5.5 | Overexpression of Site-Specific <i>B. subtilis</i> PPase Mutants | 170 |
| 5.6 | Activity of Site-Specific <i>B. subtilis</i> PPase Mutants | 171 |
| 6.1 | Overall Fold of <i>S. gordonii</i> and <i>B. subtilis</i> PPase Molecules | 175 |
| 6.2 | Overall Fold of <i>S. gordonii</i> PPase Monomer | 176 |
| 6.3 | Detailed View of the Mutated Amino Acids and Other Residues of the Pu- tative <i>B. subtilis</i> Active Site at the Domain Interface | 177 |
| 6.4 | Structure of the <i>B. subtilis</i> PPase Putative Active Site | 179 |
| 6.5 | The Proposed Catalytic Mechanism for Family II PPases | 182 |

List of Tables

| | | |
|-----|--|-----|
| 1.1 | Cloned Subfamily Ia, Prokaryotic PPases | 36 |
| 1.2 | Comparison of Family I and <i>B. subtilis</i> Inorganic Pyrophosphatases Properties | 50 |
| 2.1 | Sequences of Primers Used in Construction of <i>Bacillus subtilis</i> PPase Truncated Mutants | 89 |
| 2.2 | Sequences of Primers Used for Site-Directed Mutagenesis of <i>Bacillus subtilis</i> PPase | 90 |
| 2.3 | Primers Used for Sequencing | 91 |
| 3.1 | Molecular Mass of <i>B. subtilis</i> PPase and Its Truncated Mutants by TOF MS ES | 120 |
| 3.2 | Purification Summary of <i>B. subtilis</i> PPase and 5AATR Mutant | 123 |
| 3.3 | Effect of Manganese, EDTA and Fluoride on <i>E. coli</i> PPase and 17AATR Mutant | 130 |
| 3.4 | Calculated and Experimental Relative Subunit Molecular Masses of the Full-length PPase and Truncated Mutants | 131 |
| 4.1 | The Activity of Full-length <i>B. subtilis</i> PPase and Truncated Mutants . . . | 142 |
| 4.2 | Effect of EDTA and Fluoride on the Full-length PPase, Truncated Mutants and <i>E. coli</i> Control | 144 |
| 5.1 | Molecular Mass of <i>B. subtilis</i> PPase Mutants by TOF MS ES | 161 |
| 5.2 | Effect of EDTA and Fluoride on Arg295Ala and Lys296Ala <i>B. subtilis</i> PPase Mutants | 162 |
| 5.3 | Effect of EDTA and Fluoride on Site-Specific <i>B. subtilis</i> PPase Mutants . . | 163 |

List of Abbreviations

A/C/G/T Adenine/Cytosine/Guanine/Thymine

aa Amino acid

A₆₉₁ Absorption read at 691 nm

AMP Adenosine 5'-monophosphate

ANSA 1-amino-2-naphthol-4-sulphonic acid

ATP Adenosine 5'-triphosphate

BCA Bicinchoninic acid

BCIP/NBT 5-Bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium

BES N,N-bis[2-hydroxyethyl]-2-aminoethane sulfonic acid sodium salt

BSA Bovine serum albumine

bp Base pair

DNA Deoxyribonucleic acid

dNTP Deoxynucleoside triphosphate

dsDNA Double-stranded deoxyribonucleic acid

DTT Dithiothreitol

EDTA Ethylenediaminetetraacetate

FPLC Fast protein liquid chromatography

g Earth's gravitational field, $g = 981 \text{ cm s}^{-2}$

IPTG Isopropyl-beta-D-thiogalactopyranoside

kb Kilo base

kDa Kilo Dalton

L-, D- Prefixes for stereoisomers of chiral compounds (Latin *laevus* and *dexter* mean left and right)

MOPS 3-[N-morpholino]propane-sulfonic acid

OD₆₀₀ Optical density read at 600 nm

PAGE Polyacrylamide gel electrophoresis

P_i Inorganic phosphate

ppGpp guanosine 5'-diphosphate 3'-diphosphate

PP_i Inorganic pyrophosphate

pppGpp guanosine 5'-triphosphate 3'-diphosphate

psi Pounds per square inch

PVDF Polyvinylidene difluoride

rpm Revolutions per minute

sc Super coiled (plasmid DNA)

SDS Sodium dodecyl sulfate

TBE TRIS/borate/EDTA

TEMED *N,N,N',N'*- tetramethylenediamine

TOF MS ES Time-of-flight electrospray ionisation mass spectrometry

TRIS Tris[hydroxymethyl]-aminomethane

Triton X-100 Polyethylene glycol P-1,1,3,3-tetramethylbutylphenyl ether (4-octylphenol polyethoxylate)

TTBS Tris-buffered salt solution with Tween

TWEEN Polyoxyethylene (20) sorbitan monolaurate

UHQ water Ultra high quality water

UV Ultraviolet

v/v Volume per volume

w/v Weight per volume

Chapter 1

Introduction

1.1. Phosphates in Biological Systems

1.1.1. Phosphate Compounds

Phosphorus is one of the fundamental biological elements. It is found in a variety of forms, both organic, as all major subunits of life, e.g. amino acids, nucleotides, sugars, phospholipids, and inorganic. Among inorganic ones, most important from the biochemical point of view are phosphoanhydrides and phosphate mono- and diesters of alcohols (Rusnak 2000). The simplest among inorganic anhydrides of phosphate is inorganic pyrophosphate (PP_i) formed by two orthophosphate (P_i) residues linked by high-energy phosphoanhydride bond (Fig. 1.1, p. 15).

Phosphate molecules may form longer structures known as polyphosphates (Poly $P_{(n)}$) with the number of phosphoryl groups (n) varying from three to several hundred (Wood 1985) (Fig. 1.2, p. 15). ‘Soluble’ long-chain polyP are mainly polymers of 10 to 50, while ‘insoluble’ long-chain polyP of over 50 P_i residues (Schröder et al. 2000). Sometimes

pyrophosphate is classified as the simplest member of inorganic polyphosphates (Kulaev 2000), next to high-molecular-weight PolyP.

1.1.2. Role in Cell Metabolism

Inorganic Pyrophosphate

Inorganic pyrophosphate (PP_i) is generated in all the major biosynthetic pathways, such as activation of amino acids by aminoacyl-tRNA synthetases in polypeptide synthesis, activation of carbohydrates by uridyl transferases in polysaccharide synthesis, phosphodiester bond formation (polymerases) in polynucleotide synthesis, the acylation of CoA (acetyl-CoA synthetase) in fatty acid synthesis, as well as the synthesis of nucleosides and lipids. In fact, PP_i is being generated at all times, from the cell growth to the continuous tissue turnover in an established organism (Wood 1977).

For a long time PP_i was believed to be a waste product of the biosynthetic nucleoside triphosphate (NTP) dependent reactions, hydrolysed by cytoplasmic pyrophosphatases (PPases) immediately after being produced, and thus making biosynthetic reactions irreversible (Kukko-Kalske and Heinonen 1985, Mansurova 1989). Regulation of the biosynthesis intensity was recognized as the most important function of PP_i (Mansurova 1989). Besides, without PP_i hydrolysing activity in cells under conditions where formation of PP_i is unrestricted, it is estimated that the intracellular concentration of this compound would reach about 3 M within 1 hour in *E. coli* (Kukko-Kalske and Heinonen 1985) and quickly inhibit the cell growth. The actual observed concentration ranges from 0.5 mM (logarithmic phase of growth of *E. coli* cells) to 1 mM (Heikinheimo et al. 1996a), but apparently even 2 mM PP_i inhibits growth of some bacteria (Lahti 1983).

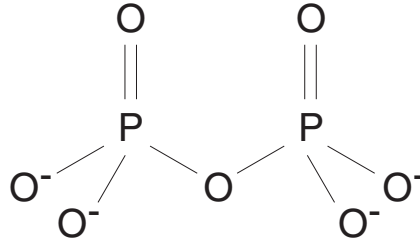


Figure 1.1: Inorganic pyrophosphate (PP_i)

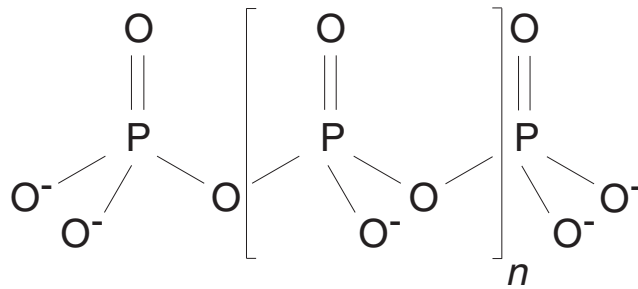


Figure 1.2: Polyphosphates (Poly $P_{(n)}$); n is the number of phosphoryl groups in the polyP molecule

Studies of the last 40 years however, have proved the presence of PP_i in a cell to be of much more significance.

An Energy Donor. Although lower than that of ATP, the free energy of PP_i hydrolysis is sufficient to serve as an alternative to ATP source of energy (Wood 1977). Under conditions similar to those *in vivo* (pH 7.0, 1 mM free Mg^{2+} and $38^\circ C$) the free energy of ATP hydrolysis is calculated as -7.6 kcal/mole and that of PP_i hydrolysis as -5.27 kcal/mole (Wood 1977). The energy conserved in PP_i is utilized in both primitive and higher organisms (Wood 1985, Mansurova 1989). Baltscheffsky (1967) demonstrated that PP_i hydrolysis accompanied both a more rapid and a more extensive reduction than ATP of a cytochrome b in *Rhodospirillum rubrum*. Varma et al. (1983) reported a *Clostridium* species able to utilize inorganic pyrophosphate as a source of energy for growth. In fact, microorganisms were found to contain specific PP_i -dependent enzymes responsible for the phosphorylation of intermediates in glycolysis and gluconeogenesis. In *Propionibacterium shermanii* and *Entamoeba histolytica* (Wood 1977) PP_i replaces the functions of ATP in the reactions catalyzed by carboxytransphosphorylase (replacing P-enolpyruvate carboxykinase) to produce P-enolpyruvate. Also, pyruvate, phosphate dikinase produces pyruvate (normally produced by pyruvate kinase) and PP_i phosphofructokinase (PP_i -PFK) produces fructose-1,6-diP (normally produced by ATP phosphofructokinase, ATP-PFK). Also, PP_i acetylkinase (in place of ATP acetylkinase) generates acetylphosphate in *Entamoeba histolytica*, while PP_i serine kinase produces phosphoserine in *Propionibacterium shermanii*.

PP_i as an energy donor and PP_i -dependent enzymes were identified in plants as well. In contrast to most organisms, there is no significant pyrophosphate hydrolysing activity by PPases in the cytoplasm of higher plants (Mertens 1991). Alkaline pyrophosphatase

is almost exclusively localized in plastids, however the cytosol was estimated to contain 0.2 – 0.3 mM PP_i (Weiner et al. 1987). Interestingly, higher plants (and *Euglena gracilis*) contain comparable cytoplasmic activities of PP_i -PFK and ATP-PFK (and fructose-1,6-biphosphatase, FBPase), while in all other PP_i -PFK-containing organisms there is little if any ATP-PFK (and FBPase) (Mertens 1991). Under stress conditions (anaerobiosis, uncoupling of mitochondrial respiration, wounding), PP_i -PFK could replace ATP-PFK because of an energetic advantage (improves the ATP yield of glucose degradation, especially in fermentative metabolism). In a model cytoplasm (Davies et al. 1993) the standard free energy change (ΔG°) and consequently the actual free energy change (ΔG) values for ATP and PP_i hydrolysis were shown to depend on pH and free Mg^{2+} . Although at normal pH values (7.2 – 7.3) ATP has a greater phosphorylating potential than PP_i (ΔG_{ATP} about -50.9, ΔG_{PP_i} about -27.3 kJ/mole at 25°C and 0.4 mM free Mg^{2+}), low values of pH and free Mg^{2+} favour the potential of PP_i to substitute for ATP. In anaerobiosis (lowered pH, decrease of $[\text{ATP}]/[\text{ADP}]$ ratio but constant PP_i levels), values of $\Delta G'_{\text{PP}_i}$ and ΔG_{PP_i} become closer to those of $\Delta G'_{\text{ATP}}$ and ΔG_{ATP} , respectively. Increased hydrolysis of PP_i under the conditions described, may be an important factor in maintaining critical cell functions (Davies et al. 1993). Furthermore, reactions catalysed by UDP-glucose pyrophosphorylase (UGPase, sucrose breakdown) and pyrophosphate:fructose-6-P phosphotransferase (PFP, Stitt 1990; glycolysis), both utilizing PP_i and identified in plant cells, are found close to the thermodynamic equilibrium (Weiner et al. 1987). In potato tubers (Farré et al. 2000), PP_i is proposed to act as a coordinator to integrate the pathways of cytosolic (PP_i utilization in sucrose breakdown by sucrose synthase and glycolysis by PFP) and plastidial (PP_i generation in starch biosynthesis by ADP-glucose pyrophosphorylase, AGPase) metabolism.

Finally, the pool of PP_i may be utilized also by a PP_i -dependent proton pump (Rea and Sanders 1987). The free energy released during PP_i hydrolysis is found to be similar to that required to move protons across the tonoplast membrane (Weiner et al. 1987).

Inorganic pyrophosphate can be synthesized, as well as ATP, in photosynthesis and oxidative phosphorylation (Wood 1985, Mansurova 1989). PP_i transport across the inner mitochondrial membrane, mediated by the adenine nucleotide translocator in exchange for ADP (1 mole of PP_i per mole of the nucleotide), also occurs (Mansurova 1989). PP_i -dependent generation of a proton-motive force ($\Delta\psi$) across the mitochondrial membrane and PP_i -dependent synthesis of ATP from ADP were also observed (Mansurova 1989).

Recent studies (Urbina et al. 1999) showed that inorganic pyrophosphate is more abundant than ATP in the parasite *Trypanosoma cruzi*. As non-metabolizable analogues of pyrophosphate inhibited the parasite growth both *in vitro* and *in vivo*, PP_i has been proposed to play a crucial role in the survival of this organism.

Regulation of Metabolism. PP_i was found to play an important role in the regulation of macromolecular synthesis and growth as well. It regulates many important enzymes, usually being only an intermediate product of the catalysed pathways. Inorganic pyrophosphate is a potent inhibitor of a second step of tRNA aminoacylation catalysed by aminoacyl-tRNA synthetase, as shown in yeast (Khvorova et al. 1999) and *E. coli* (Wolfson and Uhlenbeck 2002). Most likely, PP_i is covalently bound to the enzyme, and its dissociation is an important step of the reaction, becoming rate-limiting when tRNA's structure is distorted (Khvorova et al. 1999). PP_i was also found to regulate the activity of DNA polymerase (Anderson et al. 1999) and, when present in a too high concentration, leads to pyrophosphorolysis of DNA. In animal tissues pyrophosphate is important in regulation

of calcium-phosphate metabolism. Large deviations in PP_i concentration in tissue liquids may produce a variety of pathologies, such as calcium pyrophosphate dihydrate (CPPD) crystal deposition (Ryan and McCarty 1995).

Polyphosphates

Polyphosphates (polyP) may be found practically in all forms of life, from bacteria (Kuroda and Ohtake 2000) and yeast (Andreeva et al. 2001, Vagabov et al. 2000) to plants (Ruiz et al. 2001) and animals, including human tissues (Schröder et al. 2000)). In prokaryotes, polyP occurs in granules (volutin) along or near chromosomes and near subcellular structures participating in photosynthesis (Keasling et al. 1998). In other lower eukaryotes they may be found on the cell membrane, in vacuoles and the nucleus, but also associated with mitochondria and chloroplasts or in cytoplasm (Keasling et al. 1998). Among human tissues, the richest in polyP are bone-forming osteoblasts, but polyphosphates are also found in blood (plasma, serum, erythrocytes) and nearly all cellular compartments (Schröder et al. 2000). Once designated a ‘molecular fossil’ (Kulaev 2000), polyphosphates are gaining new attention. Many and varied biological functions have been attributed to these molecules.

An Energy Source and Phosphate Donor. Polyphosphates are assumed to serve as storage molecules for energy-rich phosphate (‘phosphagen’). Consequently, they constitute an energy source and direct phosphorylating agent (e.g. for sugar kinases, Wood 1985, Kuroda and Ohtake 2000). They regulate the intracellular level of adenylate nucleotides, serving as a phospho-donor for the kination (by polyphosphate kinase, PPK, providing a nucleoside diphosphate kinase-like activity, Kuroda and Kornberg 1997) of not only ADP but also other NDPs (Shiba et al. 2000). PolyPs maintain the phosphate level in the cell

(Nesmeyanova 2000). In addition, they play the role of chelators of many divalent cations, including Ca^{2+} , K^+ , Mg^{2+} but also heavy metals (Keasling et al. 1998) and act as counterions for basic amino acids (Schröder et al. 2000) or a buffer for alkaline stress (Shiba et al. 1997).

Practical interest in polyP is based on industrial applications, among which is ecological disposal (microbial removal) of pollutant phosphate. Polyphosphate is central to enhanced biological phosphorus removal (EPBR) from wastewater employing polyphosphate-accumulating microorganisms (Keasling et al. 1998). *Microcylindrus phosphovorans* is an example of a microorganism isolated from polyP-accumulating sludge (Kuroda and Ohtake 2000). It accumulates polyP under aerobic conditions and degrades it under anaerobic conditions along with an uptake of glucose.

Recently, the presence of high levels of short-chain condensed polyphosphates has been reported (Moreno et al. 2000) in acidic organelles, acidocalcisomes of trypanosomatid parasites (*Trypanosoma cruzi*, *Trypanosoma brucei* and *Leishmania major*). PP_i utilizing enzymes were also identified in these organisms and implicated as having a critical role in parasite metabolism.

Regulation of Gene Expression. Polyphosphates constitute regulatory factors for gene expression. The polyP level in *E. coli* cells regulates sensitivity to cellular stresses, e.g. exposure to UV and drugs (Tsutsumi et al. 2000) or H_2O_2 and heat shock (Shiba et al. 2000). Studies showed that polyphosphates were necessary for the expression of SOS genes (major DNA damage-inducible genes, elements of the SOS regulon), especially *recA* (Tsutsumi et al. 2000), but also *rpoS* (Shiba et al. 1997), via regulation of mRNA metabolism. The product of *rpoS* is the σ^{38} subunit of RNA polymerase responsible for the expression of

nearly 50 genes involved in adjustments in the stationary phase to nutrient deprivations, high osmolarity and temperature, oxidative stress, etc. (Shiba et al. 2000). It was also found that the adaptation to nutritional stringency such as amino acid starvation (e.g. via activation of intracellular protein degradation) is mediated by the network of stringent responses, e.g. accumulation of stringent factors, ppGpp and pppGpp, and polyP metabolism (Kuroda and Ohtake 2000). Moreover, the studies of Kuroda et al. (2001a) showed that polyP may promote ribosomal protein degradation, the same supplying the amino acids needed to respond to starvation. In fact, (p)ppGpps, the principal signal molecules regulating bacterial adjustments to stress (Shiba et al. 1997), are also polyP related compounds (Nesmeyanova 2000). Interestingly, a key enzyme in the stringent response, guanosine pentaphosphate phosphohydrolase (GPP) was reported to be a long-chain exopolyphosphatase (Keasling et al. 1993). Other important intracellular signalling molecules, dinucleotide polyphosphates ($A_{p_n}A$) representing a new class of second messenger molecules (Martín et al. 1998), are also polyP related. Additionally, polyP was reported to interact physically with RNA polymerase and inhibit transcription *in vitro* (Shiba et al. 2000). As a strong polyanionic polymer it is chemically similar to DNA and RNA, thus likely to bind basic proteins (e.g. histones) or their basic domains (Keasling et al. 1993). Since polyphosphate rapidly accumulates in response to stress and starvation, and then its level gradually decreases, it may play the role of a signal or a precursor of a signal molecule (like [p]ppGpp or cAMP). It may directly affect the transcription machinery (e.g. stability of the transcriptional complexes) by interacting with RNA polymerase or RNA degradosomes (Shiba et al. 2000, Tsutsumi et al. 2000).

Polyphosphates were also established to be directly implicated in developmental changes

in microorganisms, such as switching-over of the genetic program characteristic of the logarithmic growth stage of *E. coli* to the program of slow cell survival under stationary conditions (Kulaev 2000, Kornberg 1999). Interestingly, polyP was found to be essential in pathogenic *Pseudomonas aeruginosa*, not only for various forms of motility (Rashid and Kornberg 2000), but also for the development of biofilms, quorum sensing and virulence (Rashid et al. 2000). It is likely that the role of polyP is linked to its importance in signal transduction (possibly, also *rpoS* gene expression). Studies of Schröder et al. (2000) suggest that long-chain polyP, as well as PP_i , is involved in the mineralization process in bone tissue (inhibits calcium phosphate precipitation). Since it has been already shown to act as a modulator of enzymes involved in signal transduction, polyP might influence the signal transduction process in osteoblasts.

Transport Mediation. High-molecular-weight polyphosphates are involved in the formation of channels (complexes with poly- β -hydroxybutyrate and Ca^{2+}) in cell membranes (Reusch 2000). These channels were found to be implicated in calcium transport across membranes, including those of higher animal cells. They might also play an important role in competence for bacterial transformation (Kornberg 1999).

1.1.3. Role in Molecular Evolution

A theory that inorganic pyrophosphate was a high energy bond donor in the primeval earth (Chi and Kemp 2000) and is one of the key components in the development of the earliest bioenergetic systems (Baltscheffsky and Baltscheffsky 1986), is an attractive one and about 40 years old (Chi and Kemp 2000). The results of many studies lend support to this theory.

PP_i can be obtained by heating of P_i , and possibly was more abundant than ATP on the

primitive earth (Baltscheffsky et al. 1986). Similarly, polyphosphates were present in the prebiotic environment (Reusch 2000) as could be the products of phosphate condensation at high temperatures. Accordingly, they are found in volcanic condensates as the result of volcanic activity through partial hydrolysis of P_4O_{10} (Yamagata et al. 1991).

Inorganic pyrophosphate, found in every cell (see section 1.1.2/Inorganic Pyrophosphate, p. 14), has been shown to constitute an alternative to ATP in both photosynthetic and fermentation-type energy conversion. It replaces ATP in energy and phosphate transfer reactions linked to biological membranes (Wood 1977, Baltscheffsky et al. 1999), but also those of intermediary metabolism, e.g. in propionic acid bacteria (Baltscheffsky et al. 1986, Wood 1977, Wood 1985). Membrane-bound proton-translocating pyrophosphatase (H^+ -PPase) played an important role in the metabolism of phototrophic bacteria (Nyrén and Strid 1991). Having been discovered in a much wider group of organisms, including plants, PP_i -energized H^+ -translocation seems to be phylogenetically more deeply rooted than was previously thought (Drozdowicz et al. 1999). Moreover, recent studies (Baltscheffsky et al. 2001) showed that three out of the strongly conserved aa motifs in the plausible active site of the members of H^+ -pumping inorganic pyrophosphatases family contain unusually high amounts of amino acids most likely available on the primitive earth (Gly, Ala, Asp or Val). The conserved motifs may be the remaining relics of structures involved in very early energy transfer and phosphate metabolism. Interestingly, simple polyanionic (Asp-containing) peptides were shown to possess hydrolytic activity toward PP_i (Henin et al. 1997). The PP_i synthase may have preceded the ATP synthase in biological evolution (Baltscheffsky et al. 1986). Possibly, because it is less susceptible to metabolic control, it is less useful for the demanding requirements of an energy conversion system in higher forms

of life (Baltscheffsky et al. 1997b). Besides, PP_i has preserved to this day some functions that polyP plays in higher eukaryotes (Mansurova 1989). Being a much simpler energy donor than ATP (can only phosphorylate, while ATP – adenylates as well, Baltscheffsky et al. 1986), PP_i could be a part of the first less complex prebiotic system for energy coupling. Interestingly, both sequence and structural similarities between PPase and ATPase proteins have been observed in the catalytic region (Baltscheffsky and Baltscheffsky 1986).

The ubiquity of polyphosphates (see section 1.1.2/Polyphosphates, p. 19) suggests that these simple homopolymers must be fundamental constituents of biological cells and have physiological role(s) fundamental to life (Reusch 2000). Consequently, they could be the components of the earliest cells and polyP might also precede ATP in evolution as a compound coupling energy (Kulaev et al. 2000). In line with this hypothesis, in modern prokaryotes the role of polyP in bioenergetics is significant and polyP metabolism is interrelated with the ATP-ADP-AMP system. Prokaryotic cells also possess a number of enzymes active in polyP/adenyl nucleotide metabolism, enzymes that are not present in the eukaryotes (Kulaev et al. 2000). The discovery of metachromatic granules (volutin, found in bacteria, algae, yeast and protozoa, Ruiz et al. 2001), made of polyP and not nucleic acid (Kornberg 1993) proves the importance of polyP in a cell functioning. Research in recent years supports well the idea of a close interrelationship between nucleic acids and high-molecular-weight polyP (Kulaev et al. 2000, see also section 1.1.2/Regulation of Gene Expression, p. 20). RNA biosynthesis in the nucleus was shown to be coupled to the formation of polyP, possibly implying a role for polyP in the regulation of chromatin activity (Mansurova 1989). The discovery of polyP/poly-R-3-hydroxybutyrate (PHB) channels (Reusch 2000) would also support the hypothesis of the role of polyP in prebiotic evolu-

tion as ‘simpler, more efficient, simpler to produce, easier to assemble, calcium-selective channels (...) are difficult to imagine’ (Reusch 2000) . Interestingly, the role of polyP in the cell has changed during the process of evolution (Kulaev et al. 1997). In prokaryotes its function is predominantly in bioenergetics, in lower eukaryotes (yeast) its main role is to store inorganic phosphate in an osmotically inert form (polyP hydrolysis is critical to support a high rate of growth, Kulaev et al. 2000), while in the higher eukaryotes, the major function of polyP is in gene regulation and involvement in some transport processes.

1.1.4. Enzymes of Phosphate Metabolism

Enzymes involved in the metabolism of phosphate, known also as ‘phosphoryl transfer enzymes’, form one of the largest groups of enzymes in nature. The enzyme-catalysed transfer of phosphoryl groups is ubiquitous in intermediary metabolism and central to the energy balance of all organisms. These reactions are also important for cellular control mechanisms in general (Knowles 1980). Phosphoryl transfer enzymes include transferases (EC 2; International Union of Biochemistry, Nomenclature Committee. *Enzyme Nomenclature*, 1978) and hydrolases (EC 3), however there are also ligases (synthetases, EC 6) and representatives of isomerases (EC 5) in the group.

Phosphatases (alkaline and acid phosphatases, EC 3.1.3.1 and EC 3.1.3.2, ATPases, EC 3.6.1.8), phosphokinases (hexokinase, EC 2.7.1.1) and phosphomutases (phosphoglucosomutases, EC 2.7.5.1, phosphoglyceromutases, EC 2.7.5.3) are the three major groups of enzymes catalysing displacement at phosphorus in phosphoric monoesters or anhydrides. While phosphatases are hydrolases acting on ester bonds (EC 3.1) and transfer phosphate to water, kinases and mutases (transferases transferring phosphorus-containing groups, EC

2.7) use an acceptor molecule other than water and catalyse the phosphoryl group displacement to a different site in the same substrate.

The enzymes that handle phosphoric diesters (transfer of a phosphoric ester moiety rather than unsubstituted phosphoryl group) are also either hydrolytic (nucleases or DNAses, e.g. ribonucleases, EC 3.1.26.1) or nucleotidyl transfer catalysts (RNA, DNA polymerases, EC 2.7.7.6 and EC 2.7.7.7, adenylating enzymes). The ligases, e.g. forming carbon-oxygen bonds (aminoacyl-tRNA and related compounds, EC 6.1.1) and carbon-sulphur bonds (acetyl-CoA synthetase, EC 6.2.1.1), as well as some isomerases representatives also fall into this group.

Phosphotriesterases hydrolyse phosphate triesters. These compounds are unique in carrying no negative charge (Rusnak 2000).

Enzymes contributing to the metabolism of polyP in most microorganisms are polyphosphate kinases (PPKs) and exopolyphosphatases. Other enzymes working on polyP include polyphosphate:AMP phosphotransferase (PAP, some prokaryotic organisms), polyphosphate glucokinase and endopolyphosphatase (eukaryotic organisms). PPK is responsible for polyphosphate synthesis and it transfers the phosphate residue from a high-energy donor (ATP, ADP, glucose-6-phosphate, 1,3-diphosphoglycerate, Keasling et al. 1998) to the polyphosphate chain. However, the reaction may be reversed and generate ATP. The enzyme may also function as a nucleoside diphosphate kinase (NDK), converting GDP, CDP, and UDP to their respective nucleoside tri-phosphates. Exopolyphosphatases (polyphosphate phosphohydrolases, EC 3.6.1.11) remove the terminal phosphate (P_i) from the polyphosphate chain. PAP catalyzes the reversible transfer of a phosphate residue

from polyP to AMP yielding ADP. Polyphosphate glucokinase phosphorylates glucose using polyP as a phosphate donor and endopolyphosphatases cleave polyP internally.

Inorganic pyrophosphatases (EC 3.6.1.1) hydrolyse anhydride bonds in inorganic pyrophosphates. The molecular details of enzyme-catalysed reactions on phosphoric esters, despite extensive studies on their physical-organic chemistry aspects, are still not completely understood (Herschlag and Jencks 1990). Studies on inorganic pyrophosphatases are prompted by the simplicity of PP_i as a substrate and the importance of the enzyme in the control of overall phosphorus metabolism. A detailed understanding of their catalytic mechanisms would elucidate the nature of the mechanism of phosphoryl transfer and the role of divalent cations, usually required for enzyme activity.

Active Site Metal Ions

Most enzymes that catalyse phosphoryl transfer require two or more metal ions for their activity (Rusnak 2000). They are either loosely bound divalent cations (metal-activated enzymes, e.g. Mg^{2+} for all kinases, Ca^{2+} for *Staphylococcal* nuclease) or constitute an integral part of the enzyme active site (metalloenzymes: Zn^{2+} for alkaline phosphatase and *E. coli* DNA polymerase I).

Metal ions are either directly involved in the catalysis (transition state stabilisation, eg. two Zn^{2+} ions in alkaline phosphatase) or playing a structural role in the active site (eg. Mg^{2+} in alkaline phosphatase). They are proposed to interact with the substrate directly (Zn^{2+} in alkaline phosphatase) or via other active site constituents (eg. Mn^{2+} via solvent molecules in protein phosphatase 2C). Interactions with phosphoryl groups effectively neutralise their charge in mono- and diesters, making the phosphorus atom more electrophilic and prone to nucleophilic attack. The role of the nucleophile, also metal

ion coordinated, is played by hydroxide ions or the side chains of active site residues (eg. Ser102 in alkaline phosphatase). In these interactions the ions function as Lewis acids effectively lowering the pK_a of the side chains and facilitating deprotonation to form the more nucleophilic anionic species. Water molecules or other active site residues bound to active site metal ions may also function as general acids that donate protons to the leaving group. In some cases, the neutralisation of the developing negative charge on the leaving group occurs also by a direct metal interaction. Other active site residues may participate in general base catalysis and together with the metal deprotonate a nucleophile. His202 in purple acid phosphatase, homologous His151 in calcineurin and His76 in bacteriophage protein phosphatase has been proposed to participate in the enzyme reaction as active site general acid and/or bases (Rusnak 2000).

1.2. Inorganic Pyrophosphatases

1.2.1. Classification

Inorganic pyrophosphatases (EC 3.6.1.1, Pyrophosphate Phosphohydrolases) fall into two major groups of soluble and membrane-bound enzymes.

Soluble Inorganic Pyrophosphatases

Soluble inorganic pyrophosphatase (PPase) was first discovered in animal tissues in 1928 (Baykov et al. 1999). As key regulators of cellular metabolism (see section 1.2.2), inorganic pyrophosphatases are ubiquitous, constitutive enzymes (Lahti 1983), found in virtually all organisms and cell types in which they have been sought. So far, the enzymes have been

isolated from various prokaryotes (Kuranova et al. 1987, Josse 1966), eukaryotes (Hachimori et al. 1983, Kieber and Signer 1991, du Jardin et al. 1995) as well as archaeobacteria (Richter and Schäfer 1992b, van Alebeek et al. 1994, Meyer et al. 1995). PPases comprise as much as 0.1 – 0.5% of the cell protein (Baykov et al. 1999).

Soluble inorganic pyrophosphatases are divided into two families, Family I and Family II. This classification was introduced by Sivula et al. in 1999, following the identification of *Bacillus subtilis* inorganic pyrophosphatase as a protein completely unrelated to other PPases known so far (see section 1.3.5/The PPase, p. 49). Until then, soluble PPases were divided on the basis of origin, some physical and chemical characteristics into two classes: prokaryotic class A and eukaryotic class B (Young et al. 1998).

Membrane-bound Inorganic Pyrophosphatases

According to some recent sources (Maeshima 2000), the name ‘membrane-associated’ PPases should refer exclusively to membrane-bound PPases that do not show proton pump activity. All the remaining membrane-bound PPases able to transport protons across the membrane would be classified as a separate family under the common name of H⁺-PPases.

H⁺-PPases. In general, the name H⁺- PPases refers to vacuolar H⁺- translocating inorganic pyrophosphatases, known also as V-PPases. They are more recently discovered than soluble enzymes (Baltscheffsky et al. 1997a) and found both in bacteria and eukaryota (Baltscheffsky et al. 1999) but also archaea (Drozdowicz et al. 1999). The most thoroughly investigated are those found on the vacuolar membrane of plant cells (V-PPases, Rea and Poole 1993, Barkla and Pantoja 1996) and in the membranes of the phototrophic proteobacterium *Rhodospirillum rubrum* (Rea and Poole 1993, Baltscheffsky et al. 1998).

Plant H⁺-PPases are found primarily in vacuolar and Golgi membranes (Drozdowicz and Rea 2001). They appear to operate predominantly in a hydrolytic mode and to require millimolar concentrations of K⁺ for their activity (Drozdowicz et al. 1999). On the contrary, phototrophic bacterial H⁺-PPases (found in membranes of chromatophores) are freely reversible and relatively insensitive to monovalent inorganic cations (Drozdowicz et al. 1999).

Vacuolar H⁺-translocating inorganic pyrophosphatases (H⁺-PPases, V-PPases) were shown to form a distinct protein family (Baltscheffsky et al. 1997a, Baltscheffsky et al. 1999). They are ubiquitous in plants and in the higher plants V-PPase sequences show 85% identity (Baltscheffsky et al. 1999). However, the phototrophic bacterium *Rhodospirillum rubrum* sequence was shown to belong to the same family (Baltscheffsky et al. 1998). Moreover, isolation of a sequence-divergent V-PPase (PVP) from the facultatively aerobic hyperthermophilic archaeon *Pyrobaculum aerophilum* (Drozdowicz et al. 1999) indicates that these pumps are much more widely-distributed than was suspected before. Although showing only 38% sequence identity (58% similarity) with both the prototypical V-PPase from *Arabidopsis thaliana* (AVP) and the one from *Rhodospirillum rubrum* (RVP), the enzyme possesses most of the characteristic structural features of the group. Additionally, V-PPases were recently discovered also in the parasitic protozoa, such as *Trypanosoma brucei* (Rodrigues et al. 1999), *Trypanosoma cruzi* (Drozdowicz et al. 2000), *Leishmania donovani* (Drozdowicz et al. 2000), *Plasmodium falciparum* (Drozdowicz et al. 2000), *Plasmodium berghei* (Marchesini et al. 2000) or *Toxoplasma gondii* (Drozdowicz et al. 2000). Above all, it appeared that plants contain two categories of V-PPases (Drozdowicz et al. 2000): K⁺-activated (type I, AVP1-like) and K⁺-insensitive (type II, AVP2-like).

The vacuolar H⁺-PPases are intrinsic membrane proteins with a deduced (from the

cloned DNA) length of 600 to 800 amino acid residues in a subunit (Rea and Poole 1993, Baltscheffsky et al. 1998, Drozdowicz et al. 1999). The predicted subunit molecular masses are between 70 and 115 kDa (McIntosh and Vaidya 2002). However, the ones determined experimentally by SDS-PAGE are always lower: 56 kDa, about 65kDa and 64 kDa for *Rhodospirillum rubrum* (Baltscheffsky et al. 1998), plant V-PPase (Rea and Poole 1993), and the overexpressed *Pyrobaculum aerophilum* V-PPase (Drozdowicz et al. 1999), respectively. This is due to the extreme hydrophobicity of these proteins (McIntosh and Vaidya 2002). Radiation-inactivation analysis suggested that a functional conformation for the PP_i hydrolysis was a dimer of identical subunits (Maeshima 2000). However, all of the catalytic and H⁺- translocating properties of V-PPase are found to be played by a single polypeptide (Drozdowicz et al. 2000) of conserved membrane spanning topology (15 – 16 transmembrane ‘spans’, McIntosh and Vaidya 2002). Yet, the possibility of the functional multimerization is not excluded (McIntosh and Vaidya 2002). Vacuolar H⁺-PPase has been proposed as a fourth category of ion translocase, distinct from F-, P- and V-ATPases (Rea et al. 1992).

Interestingly, despite no overall similarity between the sequences of vacuolar H⁺-PPases and soluble (Family I) PPases (Rea et al. 1992), functionally important similarities between them seem to occur. An example may be the metal- and substrate-binding region (D)X₇KXE¹ found in the putative active site region of all membrane-bound PPases (Baltscheffsky et al. 1999). Thus, although the organisms are evolutionarily remote, their PPases seem to share convergent motifs due to interacting with the same substrates, inhibitors and cofactors (MgPP_i, Mg₂PP_i, Ca²⁺, Mg²⁺, Rea et al. 1992, Rea and Poole 1993).

¹aspartic acid, D is replaced by glutamic acid, E in soluble Family I

Membrane-associated, Non-Proton Pumping. Membrane-associated PPases have been discovered in the plant thylakoid membrane and plant mitochondria. Their molecular structure is more complex than that of the vacuolar enzymes (Lundin et al. 1991). The catalytic, PP_i hydrolysing, subunit is a dimer of much shorter (than vacuolar ones) polypeptides, falling in the range of 20 – 35 kDa (an exception is the 55 kDa spinach chloroplast thylakoid H^+ -PPase). It is a soluble protein that binds noncovalently to a protein proton-pumping complex in the inner mitochondrial membrane. In this respect, the PPase resembles mitochondrial F-ATPase (Lundin et al. 1991). Interestingly, not only the putative catalytic residues (Rea et al. 1992), but the overall sequences of mitochondrial H^+ -PPases were found to be similar to those of soluble PPases, Family I (Lundin et al. 1991, Cooperman et al. 1992). The catalytic subunit of the mitochondrial *Saccharomyces* H^+ -PPase is 49% identical to the soluble PPase from the same source (Lundin et al. 1991). At the same time they show no sequence identities with V-PPases (Rea et al. 1992).

1.2.2. Function in the Cell

All inorganic pyrophosphatases control the level of intracellular inorganic pyrophosphate by hydrolysing it to two inorganic phosphates (P_i). This provides a thermodynamic pull for essential biosynthetic reactions that utilise nucleoside tri-phosphates, including protein, nucleic acid and polysaccharide syntheses (see also section 1.1.2/ Inorganic Pyrophosphate, p. 14). However, this essentially similar function is performed in different ways strictly dependent on the enzyme category and its localisation in a cell.

Soluble Inorganic Pyrophosphatases

Soluble PPases have been found to be essential for cell viability in bacteria, both *E. coli* (Chen et al. 1990) and *Bacillus subtilis* (Ogasawara 2000), as well as in yeast playing a key role in a cellular metabolism (Lundin et al. 1991). The enzymes give a 10^{10} fold increase in the reaction rate as compared with noncatalyzed hydrolysis of PP_i in water, corresponding to turnover numbers within the range of $100 - 500 \text{ s}^{-1}$ for PPases of different origin (Baykov et al. 1999). The energy released in the hydrolysis of PP_i according to our present knowledge is dissipated as heat.

Membrane-bound Inorganic Pyrophosphatases

All membrane-bound H^+ -pumping PPases (H^+ -PPases) are directly involved in bioenergetic reactions, conserving some of the energy liberated during PP_i hydrolysis as transmembrane H^+ gradients.

Prokaryotic enzymes however are known to function physiologically both as ion pumps and as PP_i synthases, forming PP_i at the expense of light energy (Baltscheffsky et al. 1997a). Utilising the photosynthetically derived cellular PP_i reserves they would contribute to the maintenance of H^+ gradient across photosynthetic membranes when, due to insufficient energy supply, the direct H^+ -coupled ATP synthesis cannot be sustained or its rate is not high enough (Drozdowicz et al. 1999, Baltscheffsky et al. 1997b). *Rhodospirillum rubrum* H^+ -PPase is the first and only known alternative energy coupling system in biological membranes to the well-known ATP synthase (ATPase, H^+ -ATPase, Baltscheffsky et al. 1986). Inorganic pyrophosphate constitutes 10 – 12% of the total products of light induced phosphorylation in the chromatophores of *Rhodospirillum rubrum* (Baltscheffsky et

al. 1997b). It was shown to be the alternative energy-rich compound that could drive some energy requiring reactions there (Baltscheffsky 1967), including ATP synthesis in the dark (Baltscheffsky et al. 1998).

Eukaryotic H^+ -PPases, best known as plant vacuolar ones (see section 1.2.1/ H^+ -PPases, p. 29), function as ion pumps solely. They pump H^+ from the cytosol into the vacuole at the expense of PP_i and are considered to contribute to the generation of the transtonoplast H^+ -electrochemical gradient that is utilised in secondary active transport of variety of solutes. Although not essentially critical under optimal metabolic conditions, this energy conservation could be vital under conditions of stress (see also section 1.1.2/An Energy Donor, p. 17) such as anoxia due to flooding, when the activity of V-ATPase is impaired, or under chill, where reduction in ATP supply may result in a temporary inactivation of V-ATPase (Davies 1997). In contrast to cellular ATP levels that change dramatically upon stress conditions such as darkness or rapid changes in the respiration rate, PP_i levels stay constant. Stabilisation of PP_i levels and operation of V-PPase and other PP_i -linked enzymes would provide a back-up system for metabolism under stress, serving not only to diminish ATP consumption, but stabilising cytoplasmic pH as well. Interestingly, H^+ -PPase activity is regulated by cytosolic free Mg^{2+} , which in turn is dependent on ATP levels in a cell (Rea and Poole 1993). As the enzyme is specifically dependent on the presence of K^+ at the cytosolic face of the membrane, it has been proposed that vacuolar H^+ -PPase catalyses the co-ordinate translocation of both H^+ and K^+ into the vacuole lumen. This implies that the enzyme is involved in the regulation of turgor as postulated by Rea and Poole (Rea and Poole 1993).

Finally, mitochondrial PPases from the group of membrane bound non-pumping pro-

tons PPases ('membrane-associated PPases'), in addition to hydrolysis of PP_i and energy coupling, have been suggested to take part in the calcium-mediated response to certain hormones (Lundin et al. 1991).

1.3. Family I Inorganic Pyrophosphatases

1.3.1. Classification

Family I PPases (Sivula et al. 1999) are found in the Domains *Archaea*, *Bacteria* and *Eukaryota* with the greatest number of identified members in the Proteobacteria (Ilias et al. 2001). The Family comprises most of the known pyrophosphatase sequences (previous classes A and B) and is divided into three subfamilies: prokaryotic (Ia, previous class A), plant and animal/fungal pyrophosphatases (Ib and Ic, respectively, previous class B).

Subfamily Ia, Prokaryotic PPases

Prokaryotic pyrophosphatases constitute the largest group and over forty sequences of their representatives are now available in GenBank (Ilias et al. 2001). Twelve of the genes coding for prokaryotic PPases have been cloned so far (Table 1.1).

Prokaryotic PPases are either hexamers or tetramers (Young et al. 1998) with subunit sizes in the range of 162 – 220 amino acids (Sivula et al. 1999) and a molecular mass about 20 kDa (Young et al. 1998). The specific activities of these PPases are typically of 1000 $\mu\text{mol min}^{-1}(\text{mg protein})^{-1}$ (Young et al. 1998). The best-studied representative of this group is the *Escherichia coli* PPase (E-PPase) which is a dimer of trimers (homohexamer).

| Organism | References |
|---|---------------------------|
| <i>E. coli</i> | Lahti et al. 1988 |
| <i>Thermoplasma acidophilum</i> | Richter and Schäfer 1992a |
| <i>Sulfolobus acidocaldarius</i> | Meyer et al. 1995 |
| thermophilic bacterium PS-3 | Maruyama et al. 1996 |
| <i>Bartonella bacilliformis</i> | Mitchell and Minnick 1997 |
| cyanobacterium <i>Synechocystis</i> sp. | Gomez et al. 1997 |
| <i>Sulfolobus</i> sp. strain 7 | Wakagi et al. 1998 |
| <i>Thermus thermophilus</i> | Satoh et al. 1998 |
| <i>Bacillus stearothermophilus</i> | Satoh et al. 1999 |
| <i>Helicobacter pylori</i> | Oliva et al. 2000 |
| <i>Aquifex aeolicus</i> | Hoe et al. 2001 |
| <i>Aquifex pyrophilus</i> | Hoe et al. 2002 |

Table 1.1: Cloned Subfamily Ia, prokaryotic PPases

Subfamily Ib, Plant PPases

Plant PPases bear closer similarity to prokaryotic than to animal and fungal PPases, e.g. have the same type of deletions in their genes as prokaryotic representatives (Sivula et al. 1999). There are five species found to contain proteins of this group so far (*Oryza sativa*, *Arabidopsis thaliana*, potato *Solanum tuberosum*, *Zea mays* and barley *Hordeum vulgare*, Sivula et al. 1999) and three genes have been cloned, namely from *A. thaliana* (Kieber and Signer 1991), *S. tuberosum* (du Jardin et al. 1995) and *H. vulgare* (Visser et al. 1998). Subunits of these proteins are in the range of 211 – 216 amino acids (Sivula et al. 1999) and consequently their molecular mass would be over 20 kDa.

Subfamily Ic, Animal/Fungal PPases

Fungal enzymes are found in *Saccharomyces cerevisiae* cytoplasm and mitochondria, *Kluyveromyces lactis*, *Schizosaccharomyces pombe* and *Pichia pastoris*, while animal PPases in *Drosophila melanogaster*, *Caenorhabditis elegans*, *Bos taurus* and *Homo sapiens* (Sivula et al. 1999). Six of these genes have been cloned, namely *K. lactis* (Stark and Milner 1989), *S. pombe* (Kawasaki et al. 1990, Baykov et al. 1999), *S. cerevisiae* cytoplasm (Kolakowski Jr et al. 1988) and mitochondria (Lundin et al. 1991), *Bos taurus* retina (Yang and Wensel 1992) and human heart (Fairchild and Patejunas 1999). Proteins have been overexpressed from the last four genomic sequences.

Enzymes of this group are mostly larger than their counterparts from other subfamilies. They have subunits of 280 – 292 amino acids (Sivula et al. 1999) and a molecular mass around 32 kDa (Young et al. 1998). They are generally dimers with the exception of tetrameric mitochondrial members (Young et al. 1998). Specific activities are more diverse but essentially lower than those of the prokaryotic subfamily (100 to 800 $\mu\text{mol min}^{-1}\text{mg protein}^{-1}$, Young et al. 1998). This class is more variable including fungal and animal PPases, as well as those of human tissues. *S. cerevisiae* inorganic pyrophosphatase (Y-PPase) is the most- and best-studied member of the group (Baykov et al. 1999).

1.3.2. Conservation of Primary Structure

Despite evident differences in their subunit sizes and the molecular assemblies (see section 1.3.1), there are sequence similarities between Family I PPases (Cooperman et al. 1992, Meyer et al. 1995). Internal identities of plant versus prokaryotic group range from 27% to 49% and of plant versus animal/fungal PPases – from 20% to 29% (Sivula et al. 1999). How-

ever 17 residues are conserved in all 37 sequences currently available in the GenBank (Sivula et al. 1999) and strikingly 13 of these residues are at the active site. These are mostly aspartic acid residues (Asp115/65², Asp117/67, Asp120/70, Asp147/97, Asp152/102), lysines (Lys56/29, Lys154/104, Lys193/142), glutamic acid (Glu48/20, Glu58/31), tyrosines (Tyr93/55, Tyr192/141) and finally arginine (Arg78/43). The other four conserved residues are threonine (Thr99/61), glycine (Gly141/91) and located at the active site another glycine (Gly94/56) and aspartic acid (Asp71/42). Aspartic acid has been demonstrated to be important for substrate binding in E-PPase (Avaeva et al. 1996) but not essential for enzymatic activity in the yeast counterpart (Heikinheimo et al. 1996b). Another five residues (Tyr89, Pro118, Gly132, Phe189, Lys198, *S. cerevisiae* numbering) are conserved in about 34 sequences (Sivula et al. 1999). While mutational and functional analysis have demonstrated Tyr89 and Lys198 are not essential for catalysis, Pro118 seems to be very important for the active site cavity structure and is located in the highly conserved region (Sivula et al. 1999).

Though Family I appears to combine enzymes from widely divergent species (only modest overall sequence similarity), high sequence conservation of the active site, together with kinetic and mutagenesis data, provide evidence that common functional elements (active site structure) are evolutionarily highly conserved from bacteria to mammals. This gives a strong support for the suggestion that the catalytic mechanism is also highly conserved (Cooperman et al. 1992, Heikinheimo et al. 1996b).

The subunit interface of prokaryotic PPases is generally poorly conserved (Baykov et al. 1999), except His136 (*E. coli* numbering) involved in three-center ionic, hydrogen-bonding interaction (conserved in 17 of the 23 sequences), Tyr30 (conserved 20/23 se-

²equivalent residues in both *S. cerevisiae* and *E. coli* enzymes: Y-PPase/E-PPase

quences), Val41 (conserved 13/23 sequences) and Val84 (conserved 14/23 sequences) (Sivula et al. 1999). The latter cases suggest that hydrophobic interactions are important for intratrimeric contacts in these Family Ia enzymes (Sivula et al. 1999). In contrast, subunit interface residues of animal/fungal PPases are very well conserved and Arg51, Trp52 and Trp279 (*S. cerevisiae* numbering) are found in all nine sequences, while His87 is conserved in 5/9 cases (Sivula et al. 1999). In terms of the subunit interface, plant PPases exhibit some similarity to the prokaryotic enzymes in the region corresponding to their ionic triad (H-H-D of the prokaryotic PPases becomes R-D-K in the plant ones, Sivula et al. 1999).

1.3.3. Catalytic Properties

Family I PPases require divalent cations for catalysis, with Mg^{2+} conferring the highest activity and being the physiological activator (Baykov et al. 1999). Zn^{2+} , Mn^{2+} and Co^{2+} are the only other metal ions shown to permit enzyme activity, both for *E. coli* and yeast PPase (Josse and Wong 1971, Cooperman 1982). These enzymes are highly specific for PP_i but are also able to hydrolyse two other naturally occurring substrates, inorganic tri- and tetra-phosphates, in the presence of Mg^{2+} (Baykov et al. 1999). The PP_i -hydrolyzing activity of yeast PPase is maximal at pH 6.5 – 7 (Baykov et al. 1999). Interestingly, in the presence of Zn^{2+} , Mn^{2+} and Co^{2+} (but not Mg^{2+}) Y-PPase becomes able to hydrolyse monoesters of pyro- and triphosphate, such as ADP and ATP but at a rate of about 1/100 of that measured with $MgPP_i$ (Baykov et al. 1999). It has been suggested that substrate specificity is determined by the metal ion(s) bound to protein, rather than to substrate (Baykov et al. 1999).

Physiological PPase inhibitors include phosphate and, possibly, polyphosphates (Baykov

et al. 1999). Ca^{2+} strongly inhibits enzymes *in vitro* (Avaeva et al. 2000, Samygina et al. 2001). In the presence of Ca^{2+} a PP_i molecule occupies a catalytically incompetent location, compared to that used in the presence of activating ions. Ca^{2+} replaces Mg^{2+} at the M2 metal binding site and is unable to generate an attacking hydroxide ion, while an unhydrolyzable CaPP_i complex competes with MgPP_i . PP_i analogues, diphosphonates, inhibit, quite strongly, mammalian cytoplasmic PPases but not mitochondrial and bacterial. Finally, fluoride is a very strong PPase inhibitor of eukaryotic PPases (Baykov et al. 1979). It binds with different rates and stoichiometry to two enzyme intermediates (Baykov et al. 2000, Pohjanjoki et al. 2001). Prokaryotic enzymes were shown to be inhibited by fluoride in a rapidly reversible manner (Baykov et al. 1999).

1.3.4. Enzyme Structure

PPases in general are homooligomers. All bacterial and yeast PPases share the structural core of the monomer that consists of 8 β -strands and 2 α -helices (Fig. 1.3, p. 42). Five of the β -strands (1, 4, 5, 6 and 7) are arranged in a highly twisted β -barrel with four excursions, which makes the overall structure belong to the oligonucleotide-binding (OB) fold (Baykov et al. 1999). The barrel is capped on top with α -helix B and on the bottom with a loop between $\beta 5$ and $\beta 6$. The top surface of the β -barrel and the excursions form the active site (Kankare et al. 1996). Eukaryotic PPases are longer than prokaryotic ones (Cooperman et al. 1992), mainly because of N- (27 residues) and C-terminal (59 residues) extensions (Fig. 1.3, p. 42).

Active site structures of the *E. coli* and yeast enzymes are in general very similar (Baykov et al. 1999). The one of Y-PPase is only slightly more closed and presumably more rigid

than the E-PPase one. These conformational differences are attributed to the presence of two metal binding loops found among the most different regions in the two structures: Glu101-Asn116 and Met143-Trp153 in yeast, while Ser63-Gly66 and Lys93-Ala103 in *E. coli*. The former loop is much shorter in E-PPase, and the latter has a very different conformation despite the same length (Heikinheimo et al. 1996a).

1.3.5. Catalysis

The proposed mechanism of catalysis is based on the two high-resolution (2.2 and 2.0 Å) structures of yeast PPase (Heikinheimo et al. 1996a): one of the active enzyme (in a complex with the two activating metal ions, $\text{Mn}_2\cdot\text{Y-PPase}$) and the other of the product complex (with the product as well, $\text{Mn}_2\cdot\text{Y-PPase}\cdot[\text{MnP}_i]_2$). Mn^{2+} was chosen for crystallisation due to its preferable, compared to the physiological Mg^{2+} , properties such as tighter binding and better diffraction pattern (Rusnak 2000). According to the proposed model, the active site is highly charged and virtually every lone pair (product, substrate, transition state) is coordinated by metal ions, hydrogen bonds to protein side-chains or by hydrogen bonds to ordered water molecules.

Phosphate Binding Sites

PPases bind two P_i at the active site with substantially different affinities (Heikinheimo et al. 1996a). P1, the higher affinity site, is assigned as the leaving group, whereas P2 as an electrophilic group (Fig. 1.4, p. 45). Five amino acid residues interact through their side-chains with phosphate: three, namely Arg78 (bidentate coordination), Tyr192, Lys193 at subsite P1, and two – Lys56 and Tyr93 – at P2. More weakly bound electrophilic P_i

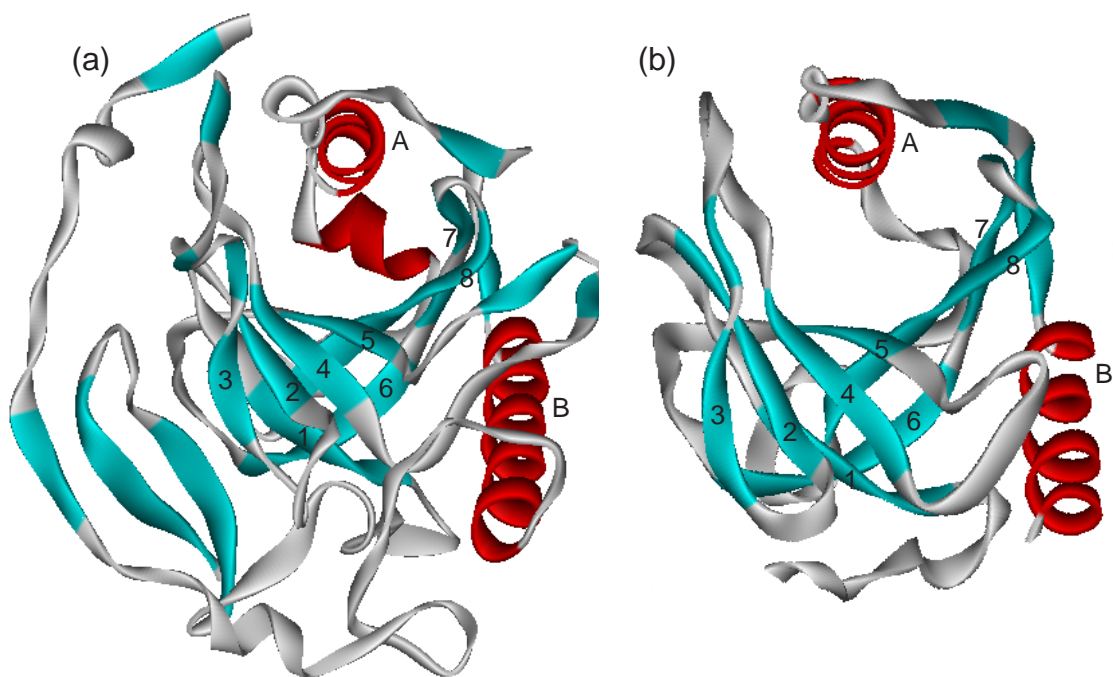


Figure 1.3: Overall fold of yeast (a) and *E. coli* (b) PPases monomers. Structural data of Tuominen et al. (1998) and Kankare et al. (1996) respectively, were downloaded from OCA PDB (2002). Figures were generated using DS ViewerPro 5.0 (Accelrys Inc. 2002). Helices are red, β -sheets are cyan and coils are grey. The marked β -sheets (1 – 8) and helices (A and B) constitute consensus secondary elements found in both structures

dissociates before the leaving group. On the other hand, tighter binding of the leaving group in the higher affinity site should correlate with higher conjugate acidity, and thus a faster hydrolysis rate. The combination of hydrogen bond donors and binding of metal ions (Arg78, Tyr192, Lys193, three water molecules and two metals, M3 and M4) may be enough to make a good leaving group, with no other need for protonation in the transition state (Heikinheimo et al. 1996a). In contrast, P2 is coordinated to metal ions (relatively weak acidic Me^{2+} like Mg^{2+} and Mn^{2+} at M1, M2, M3 and M4; also three water molecules) and only two aa residues, which overall might not be enough to lower the pK_a sufficiently.

However, recent studies of Zyranov et al. (2001) has shown that the relative affinities of P1 and P2 and the order in which P_i is released during PP_i hydrolysis depend on the nature of the metal ion cofactor used. P_i molecule containing the electrophilic phosphorus acquiring oxygen from water is released first from the site P2 only in the presence of Mg^{2+} . The affinities of P1 and P2 for MgP_i differ then only 1 – 9-fold. In the presence of Mn^{2+} and Co^{2+} P2 is a tighter binding (high affinity) subsite and the difference between the subsites affinities is at least 250-fold. Thus, P_i from P2 is released second.

Metal Ions

For the yeast PPase, although active in the presence of four metal ions, three Mg^{2+} confer the highest activity at neutral pH. In contrast, *E. coli* PPase requires four Mg^{2+} per active site to be active (Heikinheimo et al. 1996b). Most likely, in E-PPase binding of the fourth metal ion is necessary to enable deprotonation of Wat1, while in the Y-PPase the interactions (i.e. bridging M1 and M2) are strong enough to lower the pK_a of Wat1. This would be consistent with the Y-PPase active-site cavity being slightly more closed than the one of E-PPase and the M1 and M2 sites being closer together than in the *E. coli* structure.

Interestingly, the *E. coli* can bind even a fifth divalent cation (Baykov et al. 1996) and this was shown necessary for the efficient catalysis in some E-PPase active-site variants (Volk et al. 1996).

In the proposed model two sites M1 and M2, at the base of the active site, bind the activating manganese ions in the absence of substrate (Fig. 1.4, p. 45). Mn^{2+} at M1 is coordinated by three aspartic residues, namely Asp115, Asp120 and Asp152, while the remaining coordination sites are filled by water molecules and the coordination is nearly octahedral. Mn^{2+} at M2 is directly coordinated to only one protein residue, Asp120. The rest of the coordination sites are filled by water molecules (Mn_2 -Y-PPase structure). M1 and M2 are separated by a two-water molecule bridge (Tuominen et al. 1998). Upon binding of the substrate the gap between the two sites narrows from 4.9 Å to 3.7 Å and one of the water molecules is expelled. The one left, Wat1, is the proposed nucleophile.

The third metal ion at the M3 site coordinates another water molecule proposed to play a role of a general acid. The fourth metal binding site, M4, is on the opposite to M3 side of the phosphate molecule. While M3 is coordinated to P1, P2 and Glu58/31, M4 is coordinated to P1, P2, Asp152/102 and Asp 147/97 (Fig. 1.4, p. 45).

Nucleophilic Water

A hydroxide ion (Wat1) plays the role of the nucleophile (essential base). It is stabilised by coordination to two metal ions, M1 and M2, and a hydrogen bond to Asp117 (Fig. 1.5, p. 47). Its abnormally low pK_a depends on the integrity of the active site structure and any perturbation, even conserved mutations, results in the increase of its value (Salminen et al. 1995, Pohjanjoki et al. 1998). Because of its negative charge it can only attack P2 (the electrophilic phosphorus) if the negatively charged oxygens on P2 are charge-shielded

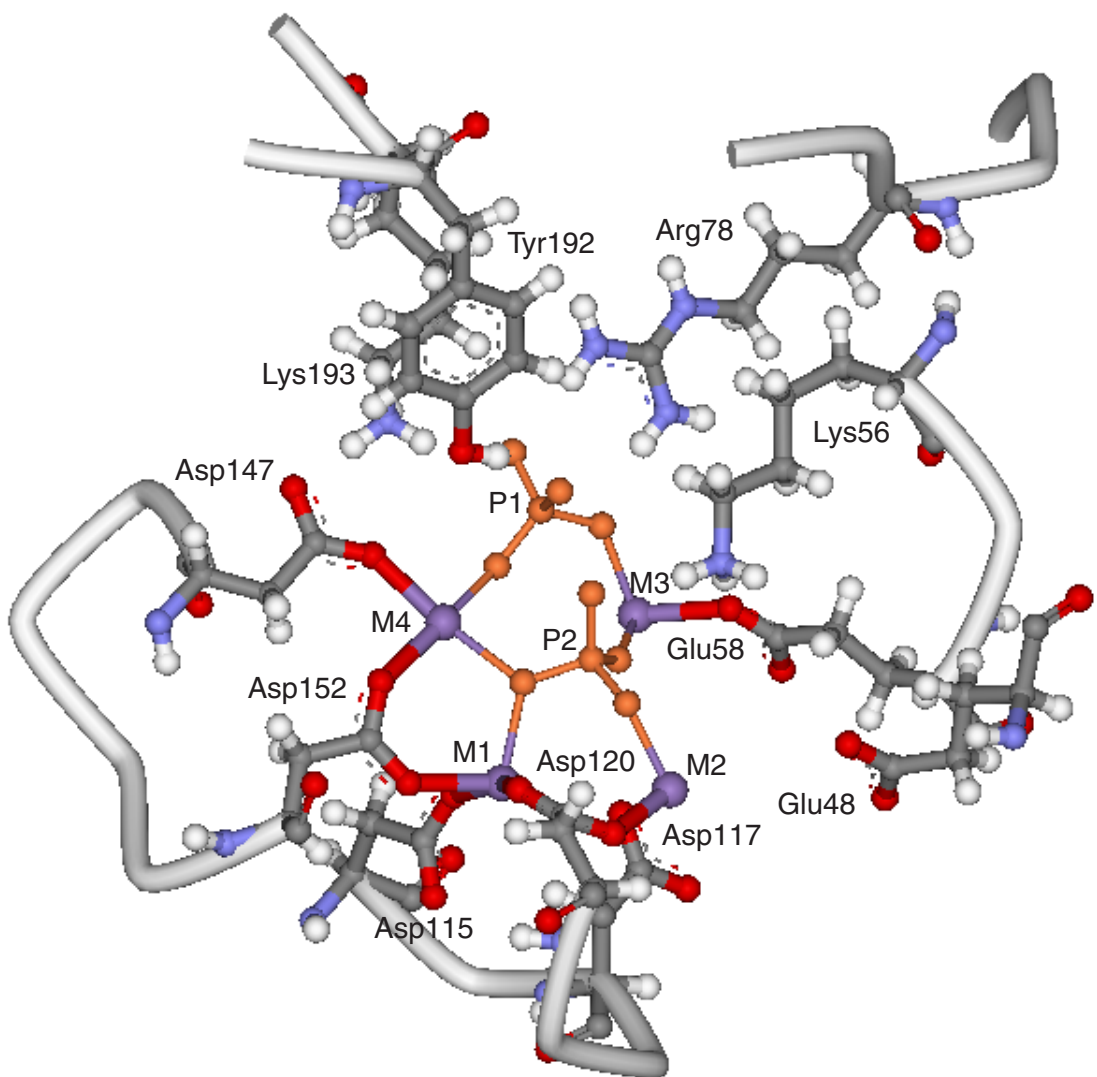


Figure 1.4: Structure of the yeast PPase (PPase:Pi₂) active site. Structural data of Heikinheimo et al. (2001) was downloaded from OCA PDB (2002). Figure was generated using DS ViewerPro 5.0 (Accelrys Inc. 2002). Phosphate binding sites, P1 and P2 and metal binding sites, M1-M4 are labelled, as are conserved active site residues

(Heikinheimo et al. 1996a). Studies of Pohjanjoki et al. (2001) have shown Asp120 to be important for exact placing of M1 and M2 and consequently for the position of Wat1. The other two aa residues found in the first coordination sphere of M1, Asp115 and Asp152, seem to have a smaller effect on this nucleophile position.

General Acid

Another water molecule (Wat6) stabilised by coordination to M3 (increases its acidity) is a strong candidate for the general acid which protonates the leaving group phosphate (P1, Fig. 1.5, p. 47).

The Mechanism

According to the Heikinheimo et al. model, PPase catalyses PP_i hydrolysis by lowering the pK_a of the leaving group in site P1 and by forming the hydroxide anion that functions as a stronger nucleophile than water in a nucleophilic attack (Fig. 1.5, p. 47). Shielding of the charge on the electrophilic phosphorus (P2) permits attack by the hydroxide ion. The hydrolysis proceeds by direct attack of water and does not require a phosphorylated enzyme intermediate. Electrophilic P_i (P2) acquiring oxygen from the nucleophilic water is the first to dissociate from the enzyme. It dissociates together with M3 exposed at the top of the active site in this first leaving MP_i complex (Tuominen et al. 1998).

However, studies of Zyranov et al. (2001) have shown the order in which P_i is released to depend on the nature of the metal ion cofactor used. P_i dissociates first from the site P2 only in the presence of Mg^{2+} . The order of the P_i dissociation is reversed in the presence of Mn^{2+} and Co^{2+} (see also section 1.3.5/Phosphate Binding Sites, p. 41). In general, the evidence presented (Zyranov et al. 2001) provides strong support for the above

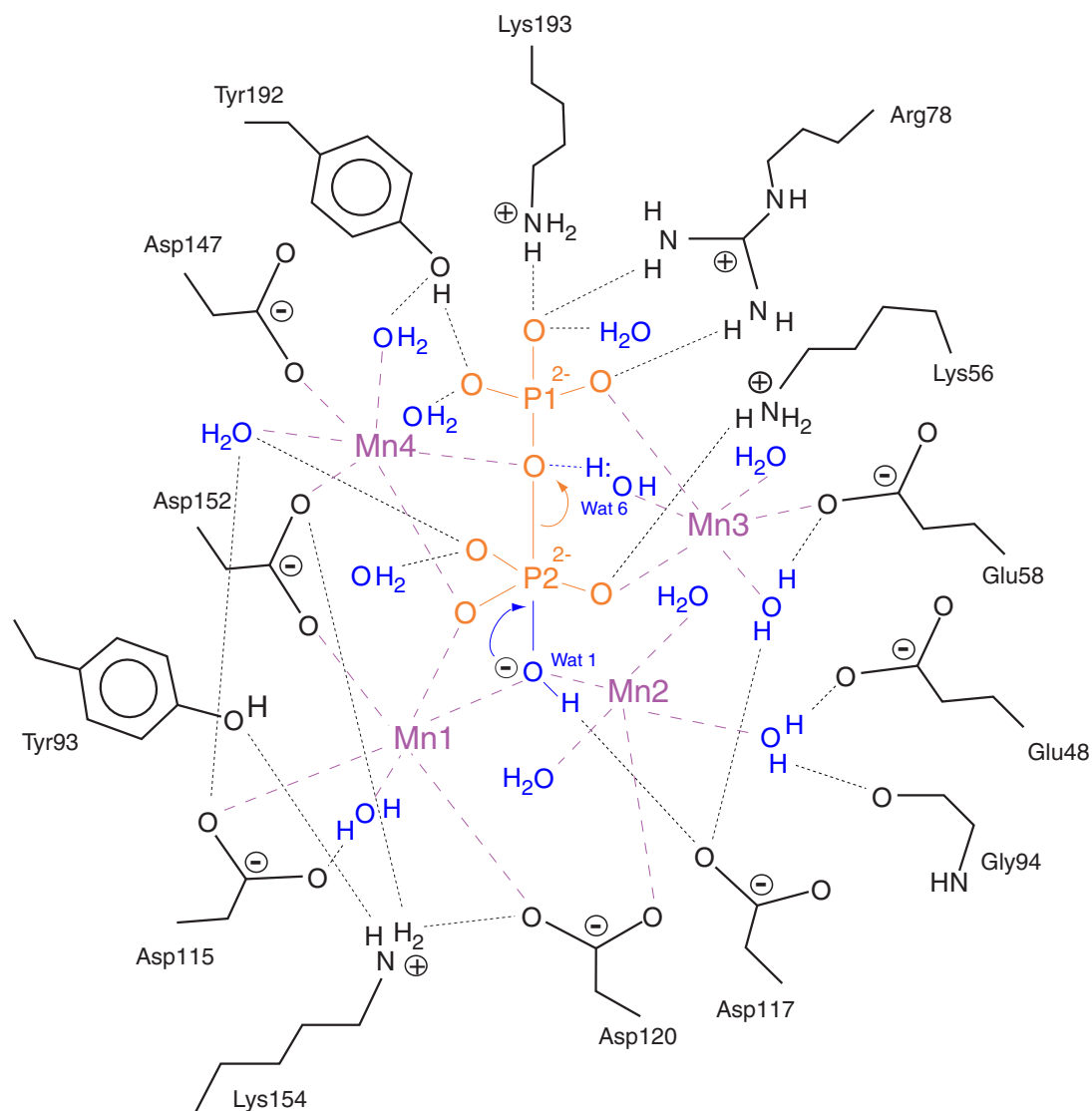


Figure 1.5: The proposed catalytic mechanism for Family I PPases (according to Heikinheimo et al. 1996a). Active site is shown in two dimensions. Hydrogen bonding is shown with *dotted* and metal coordination with *dashed* lines. The flow of electrons is indicated with *arrows*. The reaction involves a hydroxide ion (Wat1), coordinated to Mn1 and Mn2, which hydrolyzes the substrate P_2O_7 -Mn₂. P1 represents the leaving and P2 – the electrophilic phosphoryl groups. The required general acid is a water molecule (Wat6) coordinated to Mn3 and hydrogen bonded to P1

mechanism, involving water/hydroxide addition to the phosphorus atom located in the site P2 of PPase. Following hydrolysis, the relaxation process permits a new Wat1 to diffuse in. This mechanism is chemically analogous to the two-metal ion mechanism proposed for other phosphoryl transfer enzymes, e.g. polymerases and integrases (Tuominen et al. 1998).

Harutyunyan (Heikinheimo et al. 1996a) has proposed an alternative model of catalysis that differs in the identity of the water nucleophile suggested to be associated with M2 and Tyr93. However, Harutyunyan's model was based on a lower-resolution structure (2.4 Å versus 2.0 Å) and recent studies of (Pohjanjoki et al. 2001) strongly support the identification of Wat1, not the one placed in the vicinity of Tyr93, as the nucleophile.

1.4. Family II Inorganic Pyrophosphatases

– a New Line of Research

1.4.1. *Bacillus subtilis* PPase, the First Representative of a New Family

Bacillus subtilis

Bacillus subtilis is the best-characterised member of the Gram-positive bacteria, a group known to be important for the environment, medicine and industry. It is an aerobic, rod-shaped, endospore-forming bacterium commonly found in soil, water sources and in association with plants. *Bacillus subtilis* and some other members of the genus (*B. amyloliquefaciens*, *B. licheniformis*) are an important source of industrial enzymes (e.g. α -amylases, alkaline proteases) being able to secrete them directly into the growth media at concentrations in excess of 10 grams per litre (Jensen et al. 2000). Consequently, they are used in the large-scale manufacture of the industrially important enzymes.

Under conditions of nutritional starvation *Bacillus subtilis* stops growing and shows

increased metabolic diversity, e.g. induction of motility and chemotaxis, production of macromolecular hydrolases and antibiotics. Under these conditions it is also able to differentiate into a state of natural genetic competence, that allows it to bind to and to take up exogenous DNA (Solomon and Grossman 1996). If growth is not re-established, the cells are induced to form extremely resistant spores (Kunst et al. 1997, Setlow 1994). In fact, *Bacillus subtilis* provides an accessible model for investigating the general stress response in Gram-positive bacteria (Vijay et al. 2000), as in organisms studied thus far, this response is mediated by orthologues of the σ^B transcription factor of *B. subtilis*, which via association with RNA polymerase directs transcription of over 100 stress genes of diverse function.

The PPase

The primary interest in *Bacillus subtilis* inorganic pyrophosphatase came from *Bacillus subtilis* being a spore forming organism and the importance of PPase activity for driving biosynthetic pathways (see section 1.2.2), prominent in spores during germination. In 1967 Tono and Kornberg identified the PPase activity both in spores and vegetative cells and reported the enzymes for both morphological states were indistinguishable (Tono and Kornberg 1967). The enzyme was isolated and purified to near homogeneity by Kuhn and Ward (1998), who also presented a more detailed characterisation of the protein. An earlier report on the purification of *B. subtilis* PPase by Shimizu et al. (1997) contained a lot of discrepancies. The enzyme was found to be hexameric with a 24 kDa subunit and most efficiently activated by Mg^{2+} (properties characteristic of a Family Ia, *E. coli*-type enzyme).

When characterised, *Bacillus* PPase appeared unexpectedly different from all the other inorganic pyrophosphatases studied so far.

Physical and Biochemical Characteristics. *Bacillus* PPase differed not only in size (number of subunits, their molecular mass) but gave also appreciably higher specific activity values. Also distinctive was its absolute dependence on Mn^{2+} or Co^{2+} ions for activity (Kuhn and Ward 1998) and its fluoride insensitivity (Table 1.2).

Remarkable effects of Mn^{2+} on the stability and catalytic activity of *Bacillus subtilis* PPase were observed for the first time by Tono and Kornberg in 1967. Kuhn and Ward (1998) reported that separate activation by low concentrations of Mn^{2+} ions is an absolute requirement for the hydrolysis of the substrate, $MgPP_i$. The activation would convert the enzyme from a catalytically inactive to a catalytically active form.

| | Family I | | | Family II |
|------------------------------|--------------|----------|---------------|--------------------|
| | Class A* | Class B* | | Class C* |
| Origin | Prokaryotic | Plant | Animal/Fungal | <i>B. subtilis</i> |
| Subunit Mol. Mass (kDa) | ≈ 20 | > 20 | ≈ 32 | 34 |
| Subunit Size (no. of aa) | 162-220 | 211-216 | 280-292 | 309 |
| No. of Subunits | 4-6 | NA | 2 | 3-4 |
| Specific Activity† | 1000 | NA | 100-800 | 6000 |
| Fluoride Sensitivity | YES | NA | YES | None |
| Mn^{2+}/Co^{2+} Dependence | None | NA | None | ABSOLUTE |

Table 1.2: Comparison of Family I and *B. subtilis* inorganic pyrophosphatases properties (acc. to Young et al. 1998 and Sivula et al. 1999); * classification before 1999; †, $\mu\text{mol } P_i/\text{ml}/\text{min}$ formed; NA, no information available

Primary Sequence. Computer analysis of the protein database confirmed that there was no similarity of *Bacillus* PPase to other PPases (Young et al. 1998), known to show similarity to each other (see section 1.3.2). Lack of homology between the sequence of PPase from *Bacillus* and previously sequenced pyrophosphatases did support the view that inor-

ganic pyrophosphatase from *Bacillus subtilis* was unrelated to other known enzymes (Table 1.2). Interestingly however, the *Bacillus* enzyme shared considerable sequence similarity (over 50%) with predicted proteins of unknown function. These proteins were found in *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Streptococcus gordonii* (Young et al. 1998) and *Streptococcus mutans* (Shintani et al. 1998). The sequence alignment of these first Family II representatives is shown in Fig. 1.6, p. 52. Sequence comparisons revealed a conserved pattern of four His residues along with potentially active conserved groups of Asp, Ser, Glu, Lys and one Arg. Moreover, significant conservations were found at both, N- and C- termini. At the C-terminus, the conserved sequence was eight residues long (294-SRKKQVVP-301, *B. subtilis* numbering) in all five group representatives identified. Shintani et al. (1998) suggested it was a fingerprint for *Bacillus* type, Family II PPases.

Overall, the *Bacillus subtilis* enzyme was proposed to be the first representative of a new class of inorganic pyrophosphatases (Young et al. 1998, Shintani et al. 1998), designated at that time class C (Young et al. 1998), to distinguish from the existing classes A and B (see section 1.2.1/Soluble Inorganic Pyrophosphatases, p. 28, also Table 1.2). In 1999 Sivula et al. proposed re-classification of inorganic pyrophosphatases based on the primary structure and all *Bacillus*-like proteins were categorized as Family II, to distinguish from other known enzymes (class A and B), grouped now under a common name of Family I (section 1.3.1, also Table 1.2).

1.4.2. Research Perspectives

Discovery of a new family (Family II) of inorganic pyrophosphatases, enzymes unique not only in terms of physical/biochemical characteristics, but also their primary structure,

```

      *           20           *
S.mutans  : -MSKILVFGHNPDSDAIGSVMAYVYKRQ----- : 29
S.gordonii : -MSKILVFGHNPDSDAIGSYAFAYAREA----- : 30
B.subtilis : -MEKILVFGHNPDSDTQCSIAIADKKNK----- : 29
M.jannasch : ---MRYVFGHNPDSDSTASIVLAVELDCY----- : 28
A.fulgidus : MEHVYVFGHNPDSDVCSIAIAYVWKNWKEGGNVA : 38
T.maritima : -LERVYVFGHNPDSDVCSIAIGYAHKNNVE----- : 31
      6 GH2NPD1D16 S1 A 7

      40           *           60           *
S.mutans  : ----LGVDAQAVALCPNEEFAVLDYFGIQAPPVVK : 63
S.gordonii : ----YGLDTEAVALCPNEEFAVLDYFGVAAPRVIT : 64
B.subtilis : ----LGFNAEPVRLCPVNGEFAVLDYFGQESPRLV : 63
M.jannasch : -----PARLGINPNEVLRKFGVMEPELIES : 56
A.fulgidus : KMMKIEAAKPVICGVNPEKVLVEKFGFVPEIMTN : 76
T.maritima : ----KGMTFIPARSGELTNEELVLYKFGMNF----- : 60
      G4 E1 9 L F P

      80           *           100          *
S.mutans  : AQAEGAKOVLIDHNEFOOVIADIRVEVVEVDHHRV : 101
S.gordonii : AKAEGAEQVLIDHNEFOOVADIAVEVYGVVDHHRV : 102
B.subtilis : AANE-VNGVLIDHNERQOIKDIEVOVLEVDHHRV : 100
M.jannasch : AKGK---EVLIDHSEKSOVFDDEGKLIADHHRV : 91
A.fulgidus : GEKG---KVALVDHSEKAOVDCIDAEVVAVDHHRV : 111
T.maritima : -----VLIDHNEITQPEGVFAEVLLEVDHHRV : 90
      6 L DH E Q1 6 2 26 66DH56

      120          *           140          *
S.mutans  : ANFETANPEYMRLEFVCSSESIYRQKNGVAIP--- : 136
S.gordonii : ANFETANPEYMRLEFVCSSESIYRQKHSVAIP--- : 137
B.subtilis : ANFETANPEYMRLEFVCSSESIYRQKNNVAIP--- : 135
M.jannasch : G-LTITPEPPLYAKFVCSSTATVLAELYFDADIDIGGK : 128
A.fulgidus : GDVTTPEPPLFVNLVCSSTATVLAELFDATGVEIP--- : 146
T.maritima : GGLSTLNPVFFYNBFVCSSTATVLAELNGKVAIP--- : 125
      T 2P6 EVG 11166 79 2 6 6

      160          *           180          *
S.mutans  : ----KEMAGVLLSGVSDNLLKSPVTHASVPAVME : 169
S.gordonii : ----KEMAGVLLSGVSDNLLKSPVTHPTKAIAP : 170
B.subtilis : ----KEMAGVLLSAGVSDNLLKSPVCTDQVAAKE : 168
M.jannasch : KKELKPEMAGVLLSAGVSDVLLKSPVTTDLKEMAKK : 166
A.fulgidus : ----KEMAGVLLSGLVSDNLLKSPVTTDLKEMVAE : 179
T.maritima : ----REMAGVLLSGLVSDNLLKSPVTTDLKEMMANF : 158
      46AG66LS 66SD1677K T D A

      200          *           220          *
S.mutans  : LAEMAGVLDLEAGVGLALVAGKAGTN-ASKTAAQVDIDA : 205
S.gordonii : LAEMAGVLDLEAGVGLALVAGKAGTN-ASKSAEQLDIDA : 206
B.subtilis : LAEMAGVLDLEAGVGLNLVAGKAGAD-SKKTVEELSLDA : 204
M.jannasch : LAEMAGVLDLEAGVGLALVAGKAGV-GKLLKPEELNMD : 203
A.fulgidus : LAEMAGVLDLLEAGVGLALVAGKAGV-DDLTAQDLEKRDY : 216
T.maritima : LADVAKVLDLEAGVGLALVAGKAGV-PEDVDPALVLEKRDV : 195
      LA 6A 6 29 6 6 466 D

      *           240          *           260          *
S.mutans  : KTEENGVSQVRAQNTVDINELERQNEEAAKASQ : 243
S.gordonii : KTEENGNNVRAQNTVDIAELERQAEAAEKAI : 244
B.subtilis : RETGSKKVEAQNTVDIEDKQRQAEAVSKVV : 242
M.jannasch : RNDNGKRVGGOEVIDVSESKKEDYKLEEK : 241
A.fulgidus : RNDNGKRVGGOELVDLSESRIDEYEAKKMK : 254
T.maritima : RVEEGEESFAQNTVDFSTLKEKERMNTKTLK : 233
      R 9 7 6 06 D 6 7 6

      *           280          *           300          *
S.mutans  : AANGYSDVLEMTDINSNIEI-ALGNNTDKVPAANF : 281
S.gordonii : ADNGYSDVLEMTDINSNIEI-AIGSNMDKVEAANF : 282
B.subtilis : AEKNLDLLEMTDINSNIEI-AIGNEAAKVEKAFNV : 280
M.jannasch : KNEGYDLEMTDINSNIEI-AIGNEAAKVEKAFNV : 278
A.fulgidus : EEGYAGVLEMTDINSNIEI-AIGNEAAKVEKAFNV : 292
T.maritima : GEGVKKVLEMTDINSNIEI-AIGNEAAKVEKAFNV : 270
      7 6714 62 1 6 E AF

      *           320          *
S.mutans  : TLNNHAPLACAVSRKKQVVEQTESFNG-- : 310
S.gordonii : VLNNHAPLACAVSRKKQVVEQTESFNA-- : 311
B.subtilis : TLNNTALKGVSRKKQVVEQTDAMAE-- : 309
M.jannasch : KVEGNSVLEGVSRKKQVVEQERAYNG-- : 307
A.fulgidus : KLEGKSVLEGVSRKKQVVEQEKAFAL-- : 322
T.maritima : EKDGLFLKGVSRKKQVVEQGEVLRER : 301
      2 L G 6SRKK4 VF 6

```

Figure 1.6: Multiple sequence alignment of the first Family II members, recognized at the time I was starting the project (except *T. maritima*, identified shortly after). Sequences shown include: *Streptococcus mutans*, *Streptococcus gordonii*, *Bacillus subtilis*, *Methanococcus jannaschii*, *Archaeoglobus fulgidus* and *Thermotoga maritima*. *Thermotoga maritima* sequence is unusual being 548 aa long, thus about 250 aa from the central part of the sequence has been omitted in order to give the alignment. Sequences were aligned using ClustalW (Higgins et al. 1996). Residues fully conserved in all sequences are in *black boxes*. *Grey boxes* denote residues of groups conserved including conservative replacements

prompted research into this entirely new and fascinating area (Family I has been already characterised in great detail, see section 1.3).

It seemed most likely, that bearing no sequence similarity to other pyrophosphatases, while being highly similar among their own group, Family II enzymes shared a distinct active site together with a distinct mechanism of action. Both were of special interest especially in terms of the remarkable (at least $6 \times$ higher than of Family I counterparts, see Table 1.2, p. 50) activity of Family II proteins.

Family II PPases were found present in Gram-positive pathogens, e.g. *Streptococcus*, whilst absent from human cells (animal/fungal subfamily of Family I). *Streptococcus mutans*, a bacterium most commonly isolated from dental plaques, is responsible for dental caries and tooth decay (Leo et al. 1990). Because in *Bacillus subtilis* the PPase gene was shown essential for bacterial survival (Ogasawara 2000) and therefore it is likely to be essential in other organisms too, the *B. subtilis* enzyme seemed to be a potential drug design target for development of new antibiotics.

The existence of two distinct families of PPases suggested the possibility of two distinct evolutionary paths for the development of this universal, key enzyme function in the cell (see section 1.2.2), an example of convergent evolution. It seemed attractive to elucidate why two different pyrophosphatase activities were necessary (activities may differ even within a genus, e.g. *Bacillus stearothermophilus* versus *Bacillus subtilis* enzyme) and if there are any environmental factors modulating the Family I/Family II occurrence.

Last but not least, inorganic pyrophosphatase belongs to ‘phosphoryl transfer enzymes’ important for the control of overall phosphorus metabolism, and at the same time utilizing probably one of the simplest substrates possible. A detailed understanding of its catalytic

mechanism would elucidate the nature of the mechanism of phosphoryl transfer and the role of divalent cations in it.

Overall, the *B. subtilis* protein made an excellent subject for research directed at addressing the issues presented above.

Aims of the Project

The main objective of the project was the identification of parts of the molecule of *Bacillus subtilis* PPase crucial for enzyme activity. This was to aid the search for the active site and the mechanism of action of Family II enzymes, apparently different from those of Family I counterparts. Studies focused on the investigation into the role of the C-terminal fingerprint (Shintani et al. 1998) or signature, as one of the most significantly conserved regions in the protein sequence.

Chapter 2

Materials and Methods

2.1. Bacterial Strains and Plasmids

Two bacterial strains, derivatives of *E. coli* were used in this work, namely DH5 α (Hanahan, 1983; Bethesda Research Laboratories), for maintenance and storage of bacterial clones, and BL21(DE3) (Grodberg and Dunn, 1988; Novagen) for the overexpression purposes. DH5 α is *recA*⁻ (lacking the recombinase A gene) for insert stability of recombinants (and *endA*⁻, lacking the endonuclease A gene, for improving the quality of plasmid DNA). BL21(DE3) carries the T7 RNA polymerase gene (λ DE3 lysogen) under the control of the *lacUV5* promoter. It is deficient in both *lon* and *ompT* proteases, which may interfere with isolation of intact recombinant proteins. Before use, the strains were grown from a single colony on freshly prepared agar plates and incubated at 37°C overnight. Plates were stored at 4°C for up to 1 month. For long-term storage, bacterial strains were frozen at -70°C as 8% glycerol stocks (section 2.5).

Plasmid pET11c (Studier and Moffatt, 1986; Novagen) was used as a vector for cloning

purposes. It contains a 25 bp *lac* operator sequence immediately downstream from the 17 bp promoter region. Binding of the *lac* repressor at this site reduces transcription by T7 RNA polymerase. This provides a second control mechanism, in addition to repression at *lacUV5*, to suppress basal expression in λ DE3 lysogens. The vector map is shown in Fig. 2.1, p. 57.

2.2. General Protocols and Procedures

Water Distillation

Water was distilled routinely using a standard laboratory glass still and was stored in glass containers at room temperature.

Water Deionization (Ultra High Quality Water)

Deionised ultra high quality water (UHQ water) was used for cloning purposes (e.g. PCR, samples for agarose gels) as well as to prepare samples for mass spectrometry (MS) or buffers for FPLC. For FPLC purposes water was prepared routinely using an Elgastat UHQ (Elga). Water for PCR or MS was from an Elgastat Option 4 system (Elga) and sterilized in aliquots by autoclaving (section 2.2/Autoclaving, p. 58). It was then stored at room temperature.

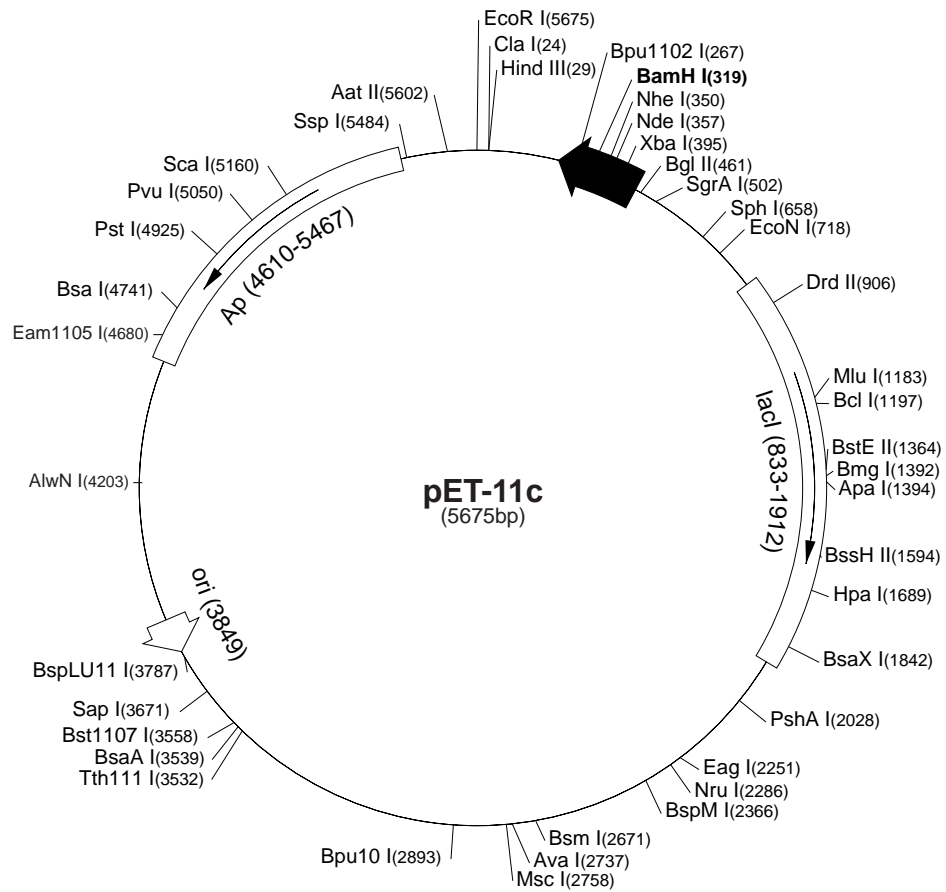


Figure 2.1: The map of pET11c vector (adopted from Novagen, 1998)

Sterilization

Autoclaving

Media for culturing bacteria, water and other solutions, were sterilized routinely by autoclaving using a Heinicke steam sterilizer (704-8000-DSE, Hotpack). Sterilization was at 121°C and 15 psi for 15 min.

Filter-sterilization

All solutions containing temperature-sensitive compounds (e.g. ampicillin) were sterilized by passing through a pre-sterilised 0.2 μm filter (Acrodisc, Pall Gelman Laboratory or Sartorius, Minisart). The same procedure was used to prepare samples or buffers for FPLC (0.45 μm filter, Pall Gelman Laboratory).

2.3. Laboratory Equipment of General Use

Centrifuges

All sample preparations of up to 1.5 ml volume were centrifuged routinely using an MSE Micro Centaur centrifuge (up to 13,000 rpm). When refrigeration was necessary, a Beckman GS-15R bench centrifuge was used. The Beckman GS-15R was used for spinning samples up to 200 ml volume (4 \times 50 ml rotor, up to 4,800 rpm).

Spectrophotometer

For all absorbance (A/OD) measurements a UV-VIS Cecil CE 1020 spectrophotometer was used.

2.4. Materials, Solutions and Buffers

All the solutions and buffers were made with distilled water, unless stated otherwise.

2.4.1. Suppliers

The chemicals and reagents used were purchased from BDH supplies unless stated otherwise. TRIS was from Boehringer Mannheim (Roche), EDTA from Fisher Scientific and sodium chloride (NaCl) from BDH or Fisher Scientific.

2.4.2. Reagents for General Use

All these chemicals were obtained from Fisher Scientific.

Ethanol

100% ethanol as well as its 70% solution were used routinely in plasmid DNA isolation. 100% ethanol was used also as a disinfectant, e.g. to clean the surfaces for aseptic work, or as a component of some solutions. It was stored at room temperature.

Methanol

100% methanol was used in the procedure of electrophoretic transfer of proteins onto immobilising membrane and immunodetection of blotted antigens, but also as a component of some solutions. It was stored at room temperature.

Isopropanol (Propan-2-ol)

100% isopropanol was used to isolate plasmid DNA using Qiagen kits. It was stored at room temperature.

Glycerol

100% glycerol was added as a component of some solutions. Its 80% solution was used for making stocks of bacterial strains for storage at -70°C . The solution was sterilized by autoclaving (section 2.2/Autoclaving, p. 58) and stored at room temperature.

2.4.3. Media

LB-broth (Luria Broth or Luria-Bertani Medium) was used routinely for growing bacterial cells. Liquid and solid media (for agar plates) contained 1% casein hydrolysate (Life Technologies), 0.5% yeast extract (Life Technologies) and 1% sodium chloride. They were made up to a volume by the addition of distilled water and sterilized by autoclaving (section 2.2/Autoclaving, p. 58). Solid media contained also 1.2% agar (Merck) added together with other solid ingredients. Agar plates (approximately 25 ml of media per petri dish) were poured under aseptic conditions and allowed to set. Solidified media were stored at 4°C and liquid media at room temperature, for up to 2 – 3 weeks.

Antibiotics

Ampicillin (Sigma) was used for positive selection. When required, it was added aseptically to autoclaved media cooled to about 50°C, to give a final concentration of 0.1 mg/ml. A 10 mg/ml filter-sterilised (section 2.2/Filter-sterilization, p. 58) stock solution was stored frozen at -20°C.

2.4.4. Isolation and Purification of Plasmid DNA

Resuspension Buffer

The solution contained 50 mM glucose (Fisons Chemicals), 10 mM EDTA and 25 mM TRIS/HCl pH 8.0. Optionally, lysosyme (muramidase from egg white, Sigma) was added to a final concentration of 4 mg/ml (4 x 41,400 units) prior to use. The solution was stored at room temperature.

Lysis Solution

Lysis solution, 0.2 M NaOH and 1% SDS, was prepared freshly from 10 M NaOH and 10% SDS stock solutions (stored at room temperature) by dilution with distilled water.

Neutralization Solution

Potassium acetate was made by combining 60 ml of 5 M potassium acetate, 11.5 ml of glacial acetic acid and 28.5 ml sterile distilled water. It was stored at room temperature.

Phenol/chloroform

A ready-to-use stock phenol:chloroform:isoamyl alcohol (25:24:1) pH 8.0 solution was used.

It was stored at 4°C.

TE Buffer

The solution was 10 mM TRIS/HCl pH 8.0, 1 mM EDTA. It was stored at room temperature.

TNE Buffer

The working solution contained 10 mM TRIS/HCl pH 8.0, 100 mM NaCl and 1 mM EDTA. It was prepared as 10× stock and stored at room temperature.

Qiagen Kits Solutions

Plasmid Mini (cat. no. 12123), QIAfilter Plasmid Maxi (cat. no. 12262) and HiSpeed Plasmid Maxi (cat. no. 12662) Qiagen Kits were used for plasmid DNA purification. QIAquick PCR Purification (cat. no. 28104) and Gel Extraction (cat. no. 28704) Kits were used to purify plasmid DNA. All the solutions supplied with the kits were used and stored according to the manufacturer's instructions.

2.4.5. Restriction and Modifying Enzymes

NdeI was purchased from Life Technologies, while *BamHI*, *SacI* (cloning purposes) and *HindIII* were from MBI Fermentas. *EcoRV*, *SacI* (qualitative analysis of truncated variants) and *Asp700* were obtained from Boehringer Mannheim (Roche). Calf intestine al-

kaline phosphatase (CIAP) and T4 DNA ligase (generation of site-directed mutants) were purchased from MBI Fermentas. T4 DNA ligase (truncated mutants construction) was obtained from BioLabs. All enzymes were supplied with corresponding 10 × concentrated reaction buffers. Enzymes and buffers were stored at -20°C.

2.4.6. Bacterial Transformation

Calcium Chloride

Calcium chloride (dihydrate, Fisons Chemicals) was used as a 100 mM solution. It was sterilized by autoclaving (section 2.2/Autoclaving, p. 58) and stored at room temperature.

2.4.7. Agarose Gel Electrophoresis

DNA Ladder (Marker)

A GeneRuler 1kb DNA Ladder from MBI Fermentas was used routinely as a DNA marker. The fragment sizes (bp) are as follows: 10,000, 8,000, 6,000, 5,000, 4,000, 3,500, 3,000, 2,500, 2,000, 1,500, 1,000, 750, 500 and 250. The GeneRuler 1kb DNA Ladder was supplied together with a 6× Loading Dye Solution (0.2% bromophenol blue, 0.2% xylene cyanol FF, 60% glycerol, 60 mM EDTA). Both solutions were stored at -20°C. The recommended amounts for use were 1:1:4 of GeneRuler, Loading Dye Solution and deionised water, respectively.

DNA Sample Buffer

The solution contained 10 mM TRIS/HCl, 50% (w/v) glycerol and 0.2% (w/v) bromophenol blue. It was stored frozen at -20°C. Alternatively, the loading dye solution supplied by MBI Fermentas (see above) was used.

TBE Buffer

TBE was 89 mM TRIS, 89 mM boric acid (Fisons Chemicals) and 0.2 mM EDTA. It was prepared fresh from a 5 × stock (stored at room temperature) by dilution with distilled water.

Ethidium Bromide

Ethidium bromide 10 mg/ml aqueous stock solution was purchased from Sigma. It was stored at room temperature. The working solution was prepared fresh prior to use by adding 2.5 μ l of ethidium bromide stock to each 50 ml of 1 × TBE buffer.

Agarose

Agarose was purchased from Sigma and, unless otherwise indicated, was used at a concentration of 0.8% (w/v). The solution was prepared by adding 0.28 g of solid agarose to 35 ml of 1 × TBE buffer and heating briefly (interrupted by occasional mixing) in a microwave oven to allow all the agarose to dissolve.

2.4.8. Protein Overexpression

IPTG

IPTG (Melford Laboratories Ltd) was stored frozen at -20°C as a filter-sterilized (section 2.2/Filter-sterilization, p. 58) 0.1 M stock solution.

2.4.9. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Protein Standard (Marker)

MW-SDS-70L by Sigma (molecular weight range 14,000 – 70,000 Da) was used as molecular weight marker. The proteins included in the kit are as follows: bovine albumin (66,000), egg albumin (45,000), rabbit muscle glyceraldehyde-3-phosphate dehydrogenase (subunit 36,000), bovine erythrocyte carbonic anhydrase (29,000), bovine pancreas PMSF treated trypsinogen (24,000), soybean trypsin inhibitor (20,000) and bovine milk α -lactalbumin (14,200). The marker working solution was prepared according to the manufacturer recommendations by reconstituting the lyophilized contents of separate vials with 1 \times Laemmli sample buffer (see below) and mixing them together in appropriate amounts. The solution was incubated at about 100°C for at least 60 s. The solution was stored at -20°C as 100 μ l 4 \times concentrated aliquots. It was diluted prior to use with 1 \times Laemmli sample buffer and stored at 4°C. 2 – 4 μ l of the marker were used per lane.

Alternatively, BioRad Prestained SDS-PAGE Standard with the protein molecular weight range from 111,000 to 20,500 Da was used. This includes phosphorylase B (111,000), bovine serum albumin (73,000), ovalbumin (47,500), carbonic anhydrase (33,900), soybean trypsin inhibitor (28,800) and lysosyme (20,500). However, when the standard was used for SDS-

PAGE the band pattern obtained did not correspond accurately to the one claimed by the manufacturer (carbonic anhydrase migrated quicker, suggesting molecular weight of 31,000 Da).

2× Laemmli Sample Buffer

The working solution of 2 × Laemmli sample buffer (Laemmli 1970) contained 0.125 M TRIS/HCl pH 6.8, 20% (w/v) glycerol, 12% (w/v) SDS, 0.01% (w/v) bromophenol blue and 0.1 M DTT (Sigma). It was stored at 4°C for up to several weeks. For long-term storage at -20°C, DTT was prepared separately as a 1 M solution in distilled water. Prior to use it was added (1 vol: 9 vol) to the mixture of remaining ingredients.

Resolving Buffer

0.75 M TRIS/HCl pH 8.8 was used as a resolving buffer. It was stored at 4°C.

Stacking Buffer

1.25 M TRIS/HCl pH 6.8 was used. It was stored at 4°C.

Reservoir Buffer

The working solution was made of 25 mM TRIS, 192 mM glycine (Fisher Scientific) and 0.1% (w/v) SDS. The buffer was stored as a 10 × concentrate at room temperature.

Acrylamide Stock Solution

The ready-to-use stock acrylamide gel solution, 30% (w/v) acrylamide and 0.8% (w/v) bis-acrylamide was ProtoGel from National Diagnostics. It was stored at 4°C.

Ammonium Persulphate

10% (w/v) ammonium persulphate (Fisher Scientific) solution was prepared fresh. It was stored at 4°C and used within the next 3 – 4 days.

SDS

10% (w/v) SDS solution was used. It was stored at room temperature.

TEMED

The ready-to-use TEMED stock solution was purchased from Fisher Scientific. It was stored at 4°C.

Water-Saturated Butanol

The solution was made by vigorous mixing of approximately 5 volumes of 100% n-butanol (Fisons Chemicals) with 1 volume of distilled water. The mixture was left to stand to allow the two phases to separate. It was stored at room temperature.

Coomassie Stain

Staining solution contained 0.025% Coomassie Brilliant Blue R-250, 40% methanol and 7% acetic acid. The dye was dissolved in methanol, mixed with acetic acid and the solution was filtered (optional). Distilled water was added and the mixture was kept at room temperature.

Destain Solution

A solution of 10% (v/v) methanol and 5% (v/v) acetic acid was used. It was stored at room temperature.

2.4.10. Electrophoretic Transfer of Proteins (Western Blotting)

Transfer Buffer

The buffer contained 48 mM TRIS, 38 mM glycine (Fisher Scientific), 0.037% (w/v) SDS and 20% (v/v) methanol. It was stored at room temperature.

PVDF Membrane

Sequi-Blot PVDF Membrane, 0.2 μm , purchased from BioRad Laboratories was used.

2.4.11. Immunodetection of Blotted Proteins

With the exception of the 'stop' solution (see below) all solutions were prepared just prior to use.

Blocking Buffer

Blocking buffer was 2% (w/v) solution of dried skimmed milk (Marvel Original, Premier Beverages) in 0.05 M TRIS/HCl pH 7.5 and 0.15 M NaCl.

TTBS

Solution was made of 20 mM TRIS/HCl pH 8.0, 0.5 M NaCl and 0.05% (w/v) Tween 20 (Fisons Chemicals).

Antibodies

Serum prepared from a guinea pig immunised with purified, overexpressed *B. subtilis* PPase was used as the primary antibody. It was stored at 4°C and used at 1:10 000 dilution (v/v, 3 μ l in 30 ml) in TTBS.

Sigma Anti-Guinea Pig IgG, Alkaline Phosphatase Conjugate, raised in a goat was used as a secondary antibody. It was stored at 4°C and used at 1:10 000 dilution in TTBS.

Alkaline Phosphatase Development Buffer

The solution contained 0.1 M TRIS/HCl pH 9.5, 0.05 M MgCl₂ (hexa-hydrate) and 0.1 M NaCl.

Substrate Solution

A ready-to-use stock solution of 5-Bromo-4-chloro-3-Indolyl Phosphate, p-Toulidine Salt (Sigma) together with Nitroblue Tetrazolium in 25 mg tablets (Sigma), were used as a BCIP/NBT substrate. One tablet of NBT was dissolved in 1 ml of deionized water, 330 μ l mixed with 10 ml of alkaline phosphatase development buffer and, subsequently, with 33 μ l of BCIP stock solution.

'Stop' Solution

0.01 M EDTA solution was used. It was diluted prior to use from 0.1 M EDTA, pH 7.0 stock stored at room temperature.

2.4.12. Blot Preparation for Protein Sequence Analysis

Blot Stain

A solution of 0.2% Coomassie Brilliant Blue R-250 in 50% methanol and 1% acetic acid was used. It was stored at room temperature.

Blot Destain

50% methanol was used and the solution was stored at room temperature.

2.4.13. FPLC

Standard Buffer

50 mM TRIS/HCl pH 8.0 in ultra high quality (UHQ) water was used routinely as a standard buffer. It contained EDTA at either 0.1 mM or 1 mM concentration. NaCl was added to the buffer as required. The solution was filtered (section 2.2/Filter-sterilization, p. 58) and degassed (water pump) prior to use. It was stored at 4°C at both working strength and as 5 × stock.

Solutions for Column Maintenance

2 M NaCl solution, 2 M NaOH and 75% acetic acid were used to clean the MonoQ ion exchange column. All were stored at room temperature.

2.4.14. Inorganic Pyrophosphatase Activity Assay

Basic Medium

A solution of 25 mM BES (Sigma) pH 7.0 and 0.4 mg/ml BSA (Sigma) was used. It was stored frozen at -20°C.

Substrate Solution

Substrate solution was prepared fresh prior to use by mixing equal volumes of 50 mM DL-serine (Sigma), 2 mM sodium pyrophosphate ($\text{Na}_4\text{P}_2\text{O}_7$, Sigma), 2 mM EDTA, pH 9.0 (solution A) and 10 mM MgCl_2 (solution B). Solution A was stored frozen at -20°C, solution B at room temperature.

When activity was assayed against tripolyphosphate, the substrate solution was prepared as described above except tripolyphosphate (Fluka) was used in place of sodium pyrophosphate. The solution was prepared fresh prior to use.

EDTA Solution

0.1 M EDTA was used as a 'stop' reagent as well as an EDTA stock for activity studies. pH was adjusted to 7.0 with NaOH and the solution was stored at room temperature. For assays of enzyme activity, EDTA stock was diluted 25 × in basic medium to give a working solution 4 mM in EDTA.

Manganese and Fluoride Solutions

Manganese and fluoride for activity studies were stored, at room temperature, as 0.1 M stocks of manganese chloride (MnCl_2 , Sigma) and sodium fluoride (NaF , Sigma). A working solution, 4 mM manganese, was prepared fresh by diluting the stock in basic medium. The working solution of fluoride was made prior to use by mixing with substrate solution at the ratio 1:89.

Fiske-Subbarow Reagent

Reagent was prepared fresh prior to use by mixing 9 ml of distilled water, 2.5 ml of ammonium molybdate reagent and 1 ml of ANSA reagent. The chemicals were purchased from Sigma, unless stated otherwise. Ammonium molybdate reagent was 2.5% (w/v) ammonium molybdate (tetrahydrate, Sigma and Fisher Scientific) dissolved in 2.5 M H_2SO_4 (Fisons Scientific) and stored at room temperature. ANSA reagent was made of 0.25% (w/v) ANSA (1-amino-2-naphthol-4-sulphonic acid), 13.3% (w/v) sodium metabisulphite ($\text{Na}_2\text{S}_2\text{O}_5$) and 0.5% (w/v) sodium sulphite (Na_2SO_3). The mixture was heated to dissolve and filtered through Whatman filter paper (Qualitative 1). It was stored in a dark bottle protected from light at room temperature.

Phosphate Standard Solution

A 0.4 mM solution of sodium phosphate (NaH_2PO_4 , Sigma) in distilled water was used. It was prepared from a 4 mM stock. Both solutions were stored at room temperature.

2.4.15. Protein Assay

Coomassie Reagent and Albumin Standard

Ready-to-use Coomassie Protein Assay Reagent (Pierce) was used. It was stored at 4°C. The reagent was supplied together with an Albumin Standard (aqueous solution of 2 mg/ml BSA in 0.9% w/v saline, 0.05% w/v sodium azide). The albumin standard was stored at room temperature.

BCA Reagents and Albumin Standard

The BCA Protein Assay Kit was purchased from Pierce. It contained BCA Reagent A (sodium carbonate, sodium bicarbonate, bicinchoninic acid and sodium tartrate in 0.1 M sodium hydroxide), Reagent B (4% aqueous solution of cupric sulphate) and Albumin Standard (2 mg/ml BSA in 0.9% saline and 0.05% sodium azide). It was stored at room temperature.

2.5. Glycerol Stocks of Bacterial Strains

For long-term storage 8% glycerol stocks of some BL21(DE3) clones (the full-length PPase, 5AATR, 17AATR, Arg295Ala, Lys296Ala, Arg295Ala/Lys296Ala, Val299Gly/Val300Trp) were made as described in the pET System Manual (Novagen). A single colony of an appropriate clone was inoculated into 50 ml of LB broth with ampicillin and incubated with vigorous shaking at 37°C until the OD₆₀₀ reached 0.6 – 0.9. 0.9 ml was removed from the culture and mixed by vortexing with 0.1 ml of sterile 80% glycerol. The stocks were stored frozen at -70°C.

2.6. Isolation of Plasmid DNA

Isolation of plasmid DNA is a standard requirement for any DNA analysis and/or manipulation purposes. In this work the alkaline lysis procedure was used, the most common method of DNA isolation. This technique exploits the difference in topology and consequently, in denaturation and renaturation characteristics of small, covalently closed circular plasmid DNA molecules and much larger chromosomal DNA. Under alkaline conditions (pH 11) both DNAs are denatured. High pH changes the condition of ionizable groups (ionizing certain groups and deionizing others), while SDS (ionic detergent) disrupts cell membranes and destabilizes all hydrophobic interactions holding various macromolecules in their native conformation. Rapid neutralization of the alkaline medium with a high salt potassium acetate buffer causes the chromosomal DNA – being too large to renature correctly – to form an insoluble aggregate. Small plasmid DNA molecules remain in solution due to their covalently closed nature promoting interstrand rehybridization. Rapid neutralization also precipitates denatured proteins (and cell debris) along with insoluble potassium dodecyl sulphate. The latter assists in the entrapment and precipitation of the high molecular weight chromosomal DNA. Soluble and insoluble material are separated by centrifugation. The scale of the procedure may vary, depending on the amount of DNA that is required.

2.6.1. Small Scale Isolation ('Miniprep')

Standard Alkaline Lysis Procedure

Small scale isolation produces a maximum of a few micrograms of DNA. The standard alkaline lysis procedure used was in principle according to Sambrook et al. (1989). It was

used mostly when the cloning work for the truncated mutants was done. Cells from 1.5 ml overnight bacterial culture were recovered by a brief spin at 13,000 rpm (up to 2 min). Often, they were frozen at this stage and the protocol was continued the following day or when required. Cells (after thawing on ice) were resuspended by vortexing in 100 μ l of resuspension buffer. EDTA in the buffer chelates the Mg^{2+} needed for enzymes that degrade DNA (DNAses). The cell suspension was mixed gently, but thoroughly by inverting the tube, with 200 μ l of lysis solution and left up to 5 min on ice or at room temperature. Longer incubation periods lead to the release of genomic DNA and, consequently, the contamination of the plasmid DNA preparation. Finally, 150 μ l of neutralization solution was added, the preparation was mixed gently, but thoroughly (to avoid localized potassium dodecyl sulphate precipitation) by inverting the tube, and left on ice for up to 30 min or put at $-20^{\circ}C$ for 15 min. Samples were then centrifuged at 13,000 rpm for up to 15 min at room temperature and the supernatant was carefully collected. The centrifugation step was repeated, if necessary, to remove all traces of precipitated material. This step helps to remove all the cellular debris, including cell wall debris with covalently bound chromosomal DNA and precipitated proteins.

The supernatant was then mixed with an equal volume (about 450 μ l) of phenol/ chloroform, vortexed and centrifuged for 2 min at 13,000 rpm and $4^{\circ}C$. Phenol/chloroform mixture denatures proteins but not nucleic acids. Centrifugation of the emulsion formed produced a lower, organic phase, separated from the upper, aqueous phase by an interface of denatured protein. The aqueous solution was recovered with special care taken not to disturb the protein interface. The phenol/chloroform step was omitted whenever possi-

ble, especially when removal of contaminating proteins was not as crucial for further DNA applications (e.g. digestion with restriction enzymes or alkaline phosphatase).

Finally, the DNA preparation was mixed with two volumes of 100% ethanol and the DNA allowed to precipitate out of solution on ice (alternatively in the freezer) for at least 15 min. After centrifugation (at least 15 min, 13,000 rpm, 4°C), the DNA pellet was washed with 1 ml of 70% ethanol, vortexed and centrifuged again. The resulting DNA pellet was air-dried and redissolved in a small amount (up to 30 μ l) of sterile distilled water or, alternatively, TE buffer. The DNA solution was stored frozen at -20°C. This procedure yielded up to 1 μ g of DNA from 1.5 ml of overnight culture.

Qiagen Plasmid Mini Kit Protocol

The Plasmid Mini Kit by Qiagen was used according to the manufacturer's instructions. Qiagen plasmid DNA isolation protocols are based on the modified alkaline lysis procedure. However, the deproteinisation step of phenol/chloroform extraction is replaced by the column chromatography and the supernatant material (after centrifugation following the neutralisation step) is applied to the Qiagen Anion-Exchange Resin. The resin binds the plasmid DNA under appropriate low-salt and pH conditions, while all the contaminants (RNA, proteins, dyes, low-molecular-weight impurities) are removed by a medium-salt wash. DNA is eluted in a high-salt buffer and then desalted and concentrated by isopropanol precipitation.

As the yield of the plasmid DNA was not satisfactory, the Special Applications Protocol for low-copy-number plasmids was used. DNA, isolated from 10 ml overnight bacterial culture and isopropanol precipitated from the lysate, is redissolved in TE buffer and loaded

onto the Qiagen resin. Then, it is purified according to the general protocol. This procedure yielded up to 7 μg of DNA from 10 ml of overnight culture.

2.6.2. Large Scale Isolation ('Maxiprep')

Two different Qiagen kits were used for the isolation of DNA in bigger quantities. These were the QIAfilter Plasmid Maxi and HiSpeed Plasmid Maxi Kit. Both were used according to the manufacturer's instructions. Both protocols rely on the modified alkaline lysis as a cell disruption method.

Qiagen QIAfilter Plasmid Maxi Kit Protocol

QIAfilter Plasmid Maxi Kit was used e.g. for DNA isolation from DH5 α clones of the original PPase, 5AATR and 17AATR for the sequencing purposes. The purification protocol follows essentially the Mini Kit one but uses larger culture and reagent volumes. Additionally, QIAfilter Maxi Cartridges are used to clear (filter) bacterial lysates without centrifugation. The protocol yielded up to about 50 μg from about 250 ml of overnight culture.

Qiagen HiSpeed Plasmid Maxi Kit Protocol

HiSpeed Plasmid Maxi Kit was used while cloning most of the site-specific mutants. The purification procedure follows essentially the Maxi Kit one. Similarly, it employs Maxi Cartridges to clear bacterial lysate instead of centrifugation. However, use of the HiSpeed Tip packed with Qiagen Resin HS, in place of the conventional Qiagen-tip, reduces the time of DNA binding, washing and eluting. Eluted DNA, mixed with isopropanol, is applied to the QIAprecipitator Module where it is bound and dried and then released with an

appropriate buffer. The HiSpeed Maxi Kit yielded up to about 70 μg of DNA from about 200 ml of overnight culture.

2.7. Standard DNA Manipulation Techniques

2.7.1. Digestion with Restriction Endonucleases

Cleavage by restriction enzymes (or restriction endonucleases) plays a key role in all the gene manipulation (cloning) procedures. These enzymes, part of the bacterial restriction/modification (or restriction/methylation) systems, allow bacteria to monitor the origin of incoming DNA and to destroy the molecules recognized as foreign. In particular, the enzymes termed type II restriction endonucleases are important in all aspects of molecular biology. They recognize specific sequences (usually 4 – 6 bp in length) and of specific base composition, and cleave the DNA into fragments in a defined manner. The sequences recognized are usually palindromic, that is they read the same in both directions on each strand, and when cleaved, they leave a cohesive-ended fragment. The ends produced from different molecules but by the same enzyme are complementary (also ‘cohesive’ or ‘sticky’) and consequently, anneal to each other.

The standard reaction mixture contained DNA to be cut, the appropriate enzyme, enzyme buffer (from $10 \times$ stock) and sterile distilled water. All digests were done at 37°C . When 5AATR truncated mutant was made, 15 μl (1.5 μg) of vector pET11c DNA was digested overnight with about 1.5 μl (7.5 units) of *NdeI* in a total volume of 25 μl . For *BamHI* digest that followed, 41 μl of DNA were used with 1 μl (10 units) of the enzyme in the total volume of 50 μl . The digestion was done for 4 hours 20 min. For 17AATR

truncated variant, 10 μl of vector DNA was digested overnight with 0.5 μl (5 units) of *Bam*HI in a total volume of 15 μl . The *Nde*I digest, also overnight, was with 1.5 μl (7.5 units) of the enzyme in a total volume of 20 μl .

The mutated inserts (up to 15 μl) were treated twice overnight (to ensure digestion) with up to 7.5 units of *Nde*I in the total volume of 20 μl . *Bam*HI treatment was for about 4 hours and about 13 μl of DNA with 10 units of enzyme were used in the total volume of about 15 μl .

For re-cloning of site-directed mutants, a vector (original PPase clone in pET11c) was digested with *Sac*I and *Bam*HI, both simultaneously and sequentially. For the double digest 40 μl (6.5 μg) of DNA was treated overnight with 1 μl of each enzyme (10 units) in the total volume of 80 μl . For the sequential digestion, up to 46 μl (8 μg) of DNA was used with 15 – 20 units of *Sac*I in the total of about 60 μl . The reaction was incubated for overnight or for 1 hour 15 min. Similarly, the *Bam*HI digestion that followed was done both overnight and for 1.5 h and about 30 μl (4 – 7 μg) of DNA was used with 10 – 20 units of enzyme in a total volume up to 60 μl . Mutagenic inserts were treated sequentially and up to 50 μl (9 μg) of DNA was digested first with 10 – 15 units of *Sac*I in the total volume of 60 μl for 1 hour and 15 min or 100 μl overnight. Then, up to 30 μl of DNA was digested with 10 – 20 units of *Bam*HI in the total volume of up to 40 μl for 1.5 h or in 60 μl overnight.

2.7.2. Dephosphorylation

To reduce the background in cloning experiments, phosphatases which remove the 5'-terminal phosphate groups from a nucleic acid strand are used. By dephosphorylating the

restricted cloning vector they prevent its re-ligation, eliminating any possibility of dimer formation or re-circularization of the vector.

The restricted, linearized vector DNA was treated with calf intestine alkaline phosphatase before ligation. When the truncated variants were constructed, up to 200 ng of DNA and 1 μ l (1 unit) of enzyme were used and the reaction was at 37°C for 25 min, in a total volume of 30 μ l. For site-directed mutagenesis, 1 – 3 μ g of DNA were treated with 0.25 – 0.5 unit of enzyme for 30 min at 37°C in the final volume of 50 μ l.

2.7.3. Ligation

The cohesive ends produced by certain restriction enzymes anneal and can be joined covalently by DNA ligase to create artificially recombinant molecules. The ligase forms a covalent bond between the 5'-phosphate at the end of one strand and the 3'-hydroxyl of the adjacent one. Ligation of the specific DNA fragment with a specialized DNA carrier molecule (vector) is necessary in order to be able to propagate the DNA of interest.

When making the truncated mutants, 1 μ l (400 units) of T4 DNA ligase was used and the vector/insert ratios 1:9 for 5AATR and 1:3 for 17AATR (total DNA about 100 ng) in a final volume of 10 μ l. When the site-directed variants were made, ligation was done with 2.5 units of T4 DNA ligase and at the vector/insert ratios 1:10 (total DNA 100 – 200 ng), except for Lys296Ala where 1:5 was also used. The total reaction volume was 20 μ l. A control ligation reaction with vector but no insert DNA present was set up at the same time to assess the likelihood of vector re-ligation (reaction background). The reaction was performed overnight at room temperature.

2.7.4. Bacterial Transformation

The recombined (cut and ligated) DNA can be replicated indefinitely (cloned) within microbial cells. Bacterial transformation, the process by which bacterial cells take up naked DNA molecules, is an essential step in many cloning experiments. Bacterial cells are induced or made competent (i.e. competent for transformation) by treatment with calcium chloride at low temperature (0 – 5°C) in the early logarithmic phase of growth. The bacterial cell membrane is permeable to chloride ions, but not to calcium ions. As the chloride ions enter the cell, water molecules accompany the charged particle. This influx of water causes the cells to swell and is necessary for the uptake of DNA, however the exact mechanism of this process is unknown. Calcium chloride treatment has to be followed by a brief increase in temperature, termed heat shock (optimum at 42°C). *E. coli* cells survive at this temperature due to the expression of the heat shock genes.

Preparation of Competent *E. coli* Cells (Calcium Chloride Procedure)

The procedure based on the one from Sambrook et al. (1989) was used routinely. A 240 μ l aliquot of the overnight *E. coli* culture was inoculated into 60 ml of fresh LB broth and grown to OD₆₀₀ of 0.3. 50 ml of this culture was transferred to a sterile 50 ml tube, left on ice for 10 min and the cells were collected by centrifugation at 4,800 rpm for 10 min at 4°C. The pellet was drained of the residual medium, resuspended in about 12 ml of a sterile ice-cold 100 mM CaCl₂ and left on ice for 20 min. Cells were harvested as before and resuspended in about 2 ml of sterile ice-cold 100 mM CaCl₂. If not used immediately, the cells were stored at 4°C and used within the following 12 – 24 hours as the transformation efficiency increases then 4 to 6 \times .

Transformation of Competent Cells with Plasmid DNA

For transformation of DNA from 5AATR and 17AATR into BL21(DE3) the modified method from Sambrook et al. (1989) was used. To 100 μl of competent *E. coli* cells, 1 – 2 μl plasmid DNA was added, the cells were mixed gently and incubated on ice for 1 hour. This was followed by a heat shock of 2 min in a 42°C water bath. Cells were then resuspended at room temperature in 0.5 ml of LB broth and incubated for 90 min at 37°C (water bath) with an occasional shaking, to allow expression of the antibiotic resistance genes carried by the plasmid. Finally, the cells were plated onto LB agar plates containing ampicillin.

For all the other transformations the protocol used was according to the pET System Manual by Novagen. According to the 9th edition, 20 μl aliquots of competent cells were pipeted into pre-chilled tubes, mixed gently with about 1 μl of plasmid DNA preparation or up to 4 μl (40 ng of DNA) of ligation mixture, and left on ice for 5 min. The tubes were then heated for 30 s in a 42°C water bath and placed on ice for 2 min. Following the addition of 80 μl of LB medium at room temperature, the transformation mixtures were incubated for 60 min at 37°C. A rotating shaker or water bath were used. In the latter case, cell suspensions were mixed gently from time to time. Cells were plated onto LB agar plates containing ampicillin.

2.8. Purification of DNA (Clean-up)

Usually, DNA at subsequent stages of a cloning procedure requires purification to remove the contaminants from the previous step, e.g. enzymes, buffer components, etc.

2.8.1. Purification of DNA Fragments from Solution

Purification of DNA in solution was done by ethanol precipitation (preceded by phenol/chloroform extraction if necessary) or, more routinely, using the QIAquick PCR Purification Kit by Qiagen. The Qiagen technique is based on the silica-gel membrane QIAquick spin-columns. DNA, in the presence of a high concentration of chaotropic salts and at an appropriate pH, adsorbs to the column membrane while contaminants pass through and impurities are washed away.

Ethanol precipitation was used mostly during construction of the truncated mutants. The DNA preparation (e.g. restriction enzyme digest) was mixed with two volumes of ice-cold 100% ethanol and frozen (put briefly into liquid nitrogen, or held at -70°C for up to 30 min or at -20°C for at least 30 min). The solution was then centrifuged (13,000 rpm, 5 min) and the resulting pellet was washed with 1 ml of 70% ethanol, centrifuged again and air-dried at room temperature. Optionally, before the ethanol precipitation (and phenol/chloroform extraction) the DNA preparation was diluted with $1 \times$ TNE (mixed with $10 \times$ concentrate and made up to $100 \mu\text{l}$ with distilled water). An equal volume of the phenol/chloroform was added to the DNA solution, the mixture was mixed well by vortexing and centrifuged at 13,000 rpm for 5 min. The upper, aqueous phase was carefully collected.

During the cloning work on site-specific protein variants, DNA was purified mainly using the QIAquick PCR Purification Kit by Qiagen according to the manufacturer's instructions. Briefly, 1 volume of the DNA sample mixed with 5 volumes of Buffer PB (allowing efficient recovery of DNA fragments starting from 100 bp) was applied to the QIAquick spin column and centrifuged (microcentrifuge, 13,000 rpm) for 1 min. Flow-through was discarded and

the column washed with 0.75 ml of Buffer PE (centrifuged as before, flow through discarded and centrifuged again). DNA was eluted with Buffer EB (up to 50 μ l) or sterile distilled water (for optimum yield, water pH was adjusted to 7.0 – 8.5).

2.8.2. Recovery of DNA Fragments from Gels

The purification of DNA fragments from agarose gels was done using Qiagen QIAquick Gel Extraction Kit according to the manufacturer's instructions. The procedure followed enzymatic reactions and was done to separate digested DNA from its uncut fraction, but mostly to isolate an appropriate DNA fragments prior to the ligation reaction.

2.9. Agarose Gel Electrophoresis

DNA electrophoresis in horizontal agarose gels is used to separate DNA molecules larger than about 100 bp according to their size. Migration rates of the DNA molecules in the gel are inversely proportional to the logarithms of their molecular weights. Electrophoresis is followed by staining with ethidium bromide which complexes with DNA by intercalating between stacked base-pairs. When illuminated with ultraviolet light, ethidium bromide exhibits a strong orange/red fluorescence which makes DNA visible. Comparing the distance migrated by the DNA band of interest with marker DNA fragments of known size allows a good visual estimation of the DNA size. It is a very useful method to check not only the purity of DNA, but also the intactness of a DNA preparation or the extent of an enzymic reaction (separates different molecular configurations of a plasmid DNA molecule).

An 0.8% agarose solution (section 2.4.7/Agarose, p. 64) was cooled to about 40 – 50°C and poured into an assembled gel tank. A comb was placed in the right position to allow

formation of the wells. After the gel has set (usually 7.5×10 cm, about 7 mm thick), the comb was removed and a sufficient volume of $1 \times$ TBE buffer containing ethidium bromide (section 2.4.7/Ethidium Bromide, p. 64) added to cover the surface of the gel (usually about 65 ml). Samples were loaded (up to $20 \mu\text{l}$) and the apparatus connected to the Power Pac 300 by BioRad. The electrophoresis was run routinely at 90 – 94 V for about 40 min, until the bromophenol blue dye (present in the DNA sample buffer) was about 0.5 cm from the bottom of the gel. To ensure good separation and then isolation of 339 bp site-specific mutant inserts from 6234 bp vector fragments (following the restriction digest while re-cloning site-specific variants, section 2.12.2) a 1.5% agarose gels were prepared. Also, a larger volume of agarose solution was used to allow deeper well formation.

2.10. Polymerase Chain Reaction (PCR)

The polymerase chain reaction is an extremely useful technique to quickly and reliably amplify specific DNA sequences up to a million-fold *in vitro*. Based on a primer extension reaction, PCR is performed in a thermal cycler with each cycle (up to 40) consisting of three standard steps.

The initial, denaturation step, denatures the template DNA with the target DNA to be specifically amplified by heating it to above 90°C for up to 2 min. The two intertwined strands separate producing the necessary single-stranded DNA template for the DNA polymerase.

In the next, annealing step (30 – 60 s), the temperature is reduced to approximately $40 - 60^\circ\text{C}$. It allows the two oligonucleotide primers, present in excess, to bind to their

complementary sites which flank the target DNA. The oligonucleotides serve as primers (provide a free 3'-hydroxyl group) for DNA synthesis by a thermostable DNA polymerase.

Finally, the temperature is raised to the optimum for the polymerase (approximately 74°C) and in this last, extension step (up to 1 min) the DNA synthesis proceeds from both of the primers until the new strands have been extended along and beyond the target DNA. Due to successive cycles of denaturation, annealing and extension, the increase in the amount of DNA produced is exponential. The use of a thermostable polymerase, most commonly *Taq* DNA polymerase isolated from the thermophilic bacterium *Thermus aquaticus*, allows the dissociation of newly formed complementary DNA and subsequent annealing of primers to the target sequence with minimal loss of enzymatic activity.

2.10.1. Generation of Truncated Coding Sequences

Truncated coding sequences of *B. subtilis* PPase were synthesized by PCR in a Hybaid PCR Express Thermal Cycler using Ready-To-Go Beads by Pharmacia Biotech. The reaction mixture included 1 ng (40 μ M) of *B. subtilis* PPase clone DNA (gene *yybQ* in pET11c), 20 pmol (0.8 μ M) of each of the two, 3'- and 5'-oligonucleotide primers (see 2.11.1), 200 μ M each deoxynucleoside triphosphate (dATP, dCTP, dGTP, dTTP), 10 mM TRIS/HCl pH 9.0, 1.5 mM MgCl₂, 50 mM KCl, stabilizers and about 1.5 unit of *Taq* DNA Polymerase, all in the final volume of 25 μ l of distilled sterilized water. The reaction was cycled 25 times, with the programme based on the one used for cloning of *B. subtilis* PPase (Young et al. 1998):

1. Initial denaturation (1 cycle), 95°C, 2 min,
2. Actual programme (25 cycles):

- Denaturation, 95°C, 1 min,
 - Primer annealing, 60°C, 1 min,
 - Extension, 72°C, 1 min,
3. Final extension (1 cycle), 72°C, 5 min.

2.10.2. Site-specific Modification of Coding Sequence

Site-specific modification of *B. subtilis* PPase coding sequence by PCR was done essentially according to the instructions of the Quick Change Site-Directed Mutagenesis Kit (Stratagene). The reaction mixture included about 50 ng of *B. subtilis* PPase clone DNA (gene *yycQ* in pET11c), 125 ng of each of the two, 3'- and 5'-oligonucleotide primers (see 2.11.1), 0.5 mM each deoxynucleoside triphosphate (dATP, dCTP, dGTP, dTTP), 20 mM TRIS/HCl pH 8.8, 2 mM MgSO₄, 10 mM KCl, 10 mM (NH₄)₂SO₄, 0.1% Triton X-100, 0.1 mg/ml nuclease-free BSA and 2.5 units of *Pfu Turbo* DNA polymerase, all in the final volume of 50 μ l of sterilized UHQ water. The amplification reaction was done using a Hybaid PCR Express Thermal Cycler. The reaction was cycled 18 times with the exception of 16 times for Ser294Ala mutant. The following programme was used:

1. Initial denaturation (1 cycle), 95°C, 30 s,
2. Actual programme (18/16 cycles):
 - Denaturation, 95°C, 30 s,
 - Primer annealing, 55°C, 1 min,
 - Extension, 68°C, 13.13 min¹,

¹for Arg295Ala, Lys296Ala, Arg295Ala/Lys296Ala and Val299Gly/Val300Trp 13.12 min were used

3. Final extension (1 cycle), 72°C, 5 min.

2.11. Oligodeoxynucleotides (Primers)

All the oligonucleotides used in this study were synthesised by Alta Bioscience on 3 PE Biosystems Expedite machines. Primers were used for truncation and site-specific modification of *B. subtilis* PPase coding sequence as well as for sequencing of the new DNA sequences produced. The clockwise (5') primer used to generate truncated mutants was Sephadex G25 purified, while all the primers for site-directed mutagenesis were purified on a reverse phase cartridge.

2.11.1. Mutagenic Primers

Truncated Mutants

A list of oligonucleotides used for construction of the truncated mutants of *B. subtilis* PPase is given in Table 2.1. The clockwise (5') primer was identical for both 5AATR and 17AATR. The oligonucleotides were constructed to introduce unique *EcoRI* and *NdeI* restriction endonuclease sites at the 5'-end and *HindIII* and *BamHI* sites at the 3' end of the amplified fragments. The primers were used also to screen the cloned DNA of enzyme variants for the presence of mutant inserts (Hybaid PCR Express Thermal Cycler).

Site-Specific Mutants

Primers used for site-directed mutagenesis of *B. subtilis* PPase coding sequence are shown in Table 2.2. Only sequences of clockwise primers are shown as those of the anticlockwise

ones were complementary. The only exception were the primers made for Lys296Ala, thus both sequences are shown.

| Mutant | Primer Direction | Primer Sequence |
|--------------|--------------------|--|
| 5AATR 17AATR | Clockwise (5') | 5'-GGA ATT CCA TAT GGA AAA GAT ACT TAT TTT CGG AC-3' |
| 5AATR | Anticlockwise (3') | 5'-GGA AGC TTG GAT CCT TAT GTT AAG ACA GGA ACG ACT T-3' |
| 17AATR | Anticlockwise (3') | 5'-GGA AGC TTG GAT CCT TAA ACG CCT TTT AAG AGG GCT G-3' |

Table 2.1: Sequences of primers used in construction of *Bacillus subtilis* PPase truncated mutants

2.11.2. Sequencing Primers

A list of primers used for sequencing is given in Table 2.3, p. 91. They were used for sequencing of the original *Bacillus subtilis* PPase clone as well as of both truncated *B. subtilis* proteins. Both T7 terminator and T7 promoter primers were used also to screen the clone DNA of truncated enzyme variants for the presence of mutant inserts (Hybaid PCR Express Thermal Cycler) while T7 terminator primer and SEQ2 were used for sequencing of the site-directed mutants as well.

2.12. Construction of Mutant Proteins

2.12.1. Truncated Mutants

Truncated *Bacillus subtilis* proteins were made using PCR and standard cloning techniques, including digestion with restriction enzymes, ligation and transformation of *E. coli* cells. DNA for PCR was WB600 *yjbQ* gene cloned in pET11c (*B. subtilis* PPase clone, Fig. 2.2-1,

| Mutant | Primer Sequence |
|----------------------|---|
| Arg295Ala | 5'-C CTC TTA AAA GGC GTT GTT TCC <u>GCT</u> AAA AAA CAA GTC GTT CCT G-3' |
| Arg295Ala/ Lys296Ala | 5'-GCC CTC TTA AAA GGC GTT GTT TCC <u>GCT</u> <u>GCA</u> AAA CAA GTC GTT CCT GTC TTA AC-3' |
| Lys296Ala | 5'-AAA GGC GTT GTT TCC CGT <u>GCA</u> AAA CAA GTC GTT CCT GTC-3' 5'-TAA GAC AGG AAC GAC TTG TTT <u>TGC</u> ACG GGA AAC AAC GCC-3'* |
| Ser294Ala | 5'-CTC TTA AAA GGC GTT GTT <u>GCC</u> CGT AAA AAA CAA GTC GTT C-3' |
| Ser294Cys | 5'-CC CTC TTA AAA GGC GTT GTT <u>TGC</u> CGT AAA AAA CAA GTC G-3' |
| Ser294Thr | 5'- CC CTC TTA AAA GGC GTT GTT <u>ACC</u> CGT AAA AAA CAA GTC G-3' |
| Ser294Val | 5'-GCC CTC TTA AAA GGC GTT GTT <u>GTC</u> CGT AAA AAA CAA GTC GTT C-3' |
| Val299Gly/Val300Trp | 5'-GGC GTT GTT TCC CGT AAA AAA CAA <u>GGC TGG</u> CCT GTC TTA ACA GAC GC-3' |
| Val299Ile | 5'-GTT TCC CGT AAA AAA CAA <u>ATC</u> GTT CCT GTC TTA ACA GAC GC-3' |
| Val300Ile | 5'-GTT TCC CGT AAA AAA CAA GTC <u>ATT</u> CCT GTC TTA ACA GAC GC-3' |
| Val300Trp | 5'-GTT TCC CGT AAA AAA CAA GTC <u>TGG</u> CCT GTC TTA ACA GAC GCA ATG G-3' |

Table 2.2: Sequences of primers used for site-directed mutagenesis of *Bacillus subtilis* PPase. Mismatches with the original sequence are indicated in bold. Codons for the mutated amino acids are underlined; *, anticlockwise (3') primer sequence

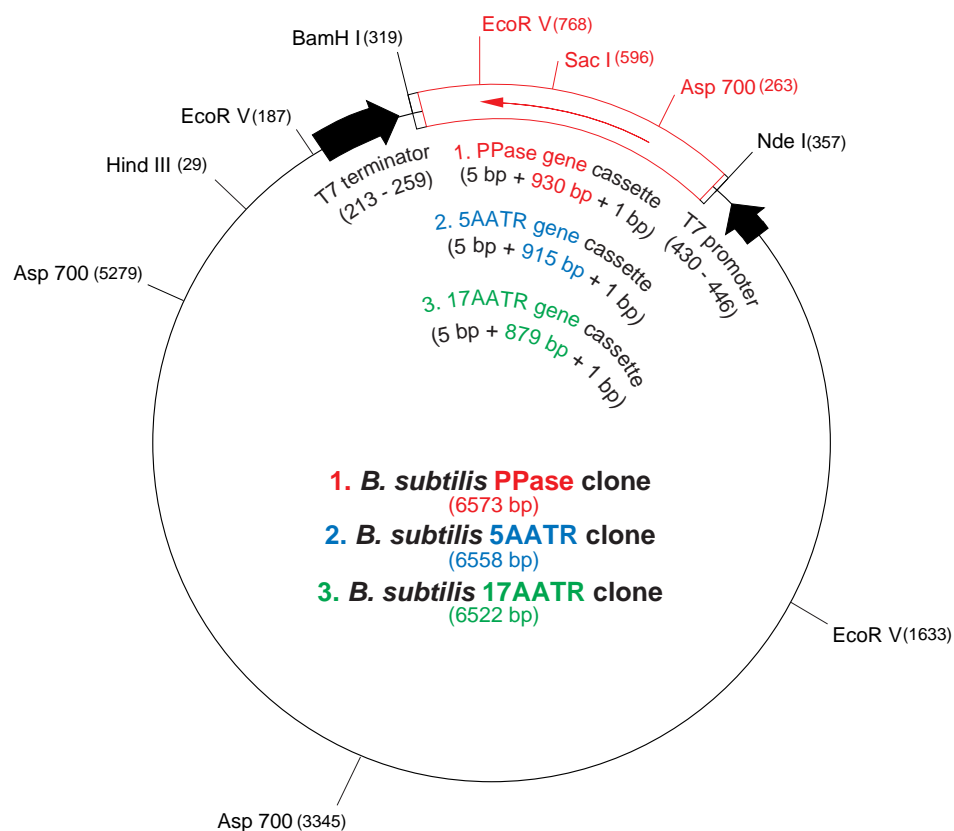
| Primer Name | Primer Direction | Primer Sequence |
|--------------------------------|--------------------|-----------------------------------|
| T7 Promoter Primer (Novagen) | Clockwise (5') | 5'-TAA TAC GAC TCA CTA TAG GG-3' |
| T7 Terminator Primer (Novagen) | Anticlockwise (3') | 5'-GCT AGT TAT TGC TCA GCG G-3' |
| SEQ1 | Clockwise (5') | 5'-AGA GGT TAT CGA CCA TCA CC-3' |
| SEQ2 | Clockwise (5') | 5'-AAC TTG CAC GGA CCA AGA CGT-3' |
| SEQ3 | Anticlockwise (3') | 5'-CAA GCC GTA TTC TTC AGC-3' |
| SEQ4 | Anticlockwise (3') | 5'-CTG GCG TTC GTT ATG GTC AA-3' |
| SEQ5 | Anticlockwise (3') | 5'-GCT TCT AAT TCA GCT TGG CG-3' |
| SEQ6 | Clockwise (5') | 5'-GCG TCT TGT TGA AAC AGC TGC-3' |

Table 2.3: Primers used for sequencing

p. 93) as described in Young et al. (1998). The amplification reaction details are given in section 2.10.1 and the primers used are described in section 2.11.1/Truncated Mutants, p. 88 (see Table 2.1, p. 89). The oligonucleotides were designed to amplify the coding region of the *B. subtilis* PPase gene inserting stop codons to eliminate either 15 or 51 3'-terminal nucleotide residues.

Both the amplified products, engineered to contain *Bam*HI and *Nde*I sites, and vector pET11c (Fig. 2.1, p. 57) were digested sequentially with *Bam*HI and *Nde*I (section 2.7.1). Following the first digest, the linearized vector DNA was separated from its uncut fraction in agarose gel and then extracted. The restricted vector DNA was additionally treated with calf intestine alkaline phosphatase (section 2.7.2). The mutant PCR products were ligated into the plasmid pET11c vector (Fig. 2.2, p. 93). Ligation was overnight at room temperature (section 2.7.3).

The ligated DNA was transformed into *E. coli* DH5 α for maintenance and storage. LB agar medium containing ampicillin was used for selection of transformant colonies. Plasmid DNA was screened for the presence of mutant inserts by PCR using T7 terminator and T7 promoter primers (Fig. 2.2, p. 93) as well as primers used for cloning (section 2.11.1/Truncated Mutants, p. 88). Transformants were also analysed by restriction enzyme digestion using *Eco*RV, *Asp*700 and *Sac*I with *Hind*III. The predicted restriction patterns are shown in Fig. 2.2, p. 93. Finally, the mutant DNA was sequenced in both directions (section 2.13).



| Restriction Enzymes | Restriction Pattern (bp) | | |
|-------------------------|--------------------------|-------------|-------------|
| | Clone Variant | | |
| | PPase | 5AATR | 17AATR |
| EcoR V | 4229 | 4229 | 4229 |
| | 2045 | 2045 | 2045 |
| | 299 | 284 | 248 |
| Asp 700 | 3252 | 3252 | 3252 |
| | 1934 | 1934 | 1934 |
| | 1387 | 1372 | 1336 |
| Sac I / Hind III | 5944 | 5944 | 5944 |
| | 629 | 614 | 578 |

Figure 2.2: Scheme of the *B. subtilis* full-length PPase (1), 5AATR (2) and 17AATR (3) clones in pET11c with their restriction map. The PPase, 5AATR and 17AATR gene cassettes include genes of the respective protein variant: PPase *yybQ* (red), 5AATR (blue) or 17AATR (green), as well as extra base pairs introduced by the primers (5 bp + 1 bp). The table presents the restriction patterns obtained while cutting the constructs with different combinations of restriction enzymes, used to screen the truncated variants (section 2.12.1, p. 92)

2.12.2. Site-Specific Mutants

All site-directed mutants were made using the QuickChange Site-Directed Mutagenesis Kit by Stratagene according to the manufacturer's instructions. DNA for PCR was the *Bacillus subtilis* WB600 *yjbQ* gene in pET11c (*B. subtilis* PPase clone, Fig. 2.2-1, p. 93). The amplification reaction details are in section 2.10.2 and of the mutagenic primers used in section 2.11.1/Site-Specific Mutants, p. 88 (see Table 2.2, p. 90). Following PCR, DNA was treated with *DpnI* to digest the methylated, non-mutated parental DNA template. The circular, nicked dsDNA was subsequently transformed into Epicurian Coli XL1-Blue supercompetent cells, where the nicks in the mutated plasmid were repaired. Selection of transformant colonies was done using ampicillin in LB agar medium. Following the transformation into Epicurian Coli cells the transformant DNA was sequenced (section 2.13) in one direction.

After the desired mutations were confirmed, the mutant inserts were re-cloned into vector DNA of a confirmed sequence. The sequenced *B. subtilis* PPase clone and the transformant DNAs (variants of the specifically mutated PPase clone) isolated from Epicurian cells were treated, mostly separately, with *SacI* and *BamHI* (see also Fig. 2.2, p. 93). The band sizes obtained were 6234 and 339 bp, for each construct (the original PPase clone as well as the mutant ones). Following the digestion, DNA fragments were separated and purified from agarose gel (section 2.8.2). The vector DNA was additionally digested with alkaline phosphatase (section 2.7.2). Mutant (339 bp) and vector (6234 bp) fragments were ligated overnight at room temperature (section 2.7.3). The ligated DNA was transformed into *E. coli* DH5 α and sequenced across the restriction sites in both directions.

2.13. DNA Sequencing

DNA sequences of the full-length PPase and truncated proteins were determined by Alta Bioscience and those of the site-directed mutants by Lark Technologies, Inc. Sequencing was done using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit and ABI PRISM 377 automatic DNA Sequencer (Applied Biosystems) for the analysis of sequencing reactions. PE Biosystems 877 robotic system was used for cycle sequencing reactions.

2.14. Protein Overexpression and Isolation

Plasmid DNA extracted from DH5 α clones was used to transform *E. coli* BL21(DE3). A single colony of an overnight *E. coli* culture was suspended in 5 ml of LB media with ampicillin and a new culture grown with shaking at 37°C until the OD₆₀₀ reached about 0.5. The bacterial suspension was then centrifuged at 4800 rpm for 10 min, the supernatant was removed and pellet, resuspended in fresh LB broth (5 ml), stored in the cold room overnight. The following day, 5 ml of bacterial suspension was inoculated into 200 ml LB with ampicillin (added just prior to use) and incubated with shaking at 37°C until the OD₆₀₀ reached 0.45 – 0.7. The contents of the flask was centrifuged at 4800 rpm for 20 min and the pellet resuspended in 200 ml of the fresh media with ampicillin and IPTG added (1 mM).

The protein was over-expressed for 2 h at 37°C (truncated mutants) or 2.5 h at 30°C (the full-length PPase and site-directed mutants). Cultures were grown with shaking at about 200 rpm. For SDS-PAGE analysis, 100 μ l samples withdrawn from the cultures

before and after induction were centrifuged briefly, mixed with an equal volume of 2 × Laemmli sample buffer and heated at 95°C for at least 5 min. 200 ml cultures of induced cells were harvested by centrifugation (3800 g, 20 min), washed in the standard buffer containing 0.1 (Arg295Ala, Lys296Ala, Arg295Ala/Lys296Ala, Val299Gly/Val300Trp) or 1 mM EDTA (the full-length PPase, 5AATR and all the remaining site-specific mutants), re-centrifuged and frozen at -20°C. Usually the following day the cells were thawed on ice and resuspended in 25 ml of the standard buffer containing an appropriate concentration of EDTA. If the washing step did not precede freezing and was done only after the cells has been thawed, there were considerable losses of protein. The bacterial suspension was sonicated on ice in four/five 90 s pulses with about 30 s rests. The sonicated extract was centrifuged at 30,000 g (MSE Hi-Spin 21) for 30 min at 4°C. The supernatant (whole-cell extract) of 17AATR, not sonicated in the buffer with EDTA, was made 0.1 mM in EDTA. All extracts were stored in aliquots at -20°C. For the *E. coli* control, the extract of *E. coli* BL21(DE3) transformed with pET11c alone was prepared in the same way.

2.15. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-polyacrylamide gel electrophoresis was used to analyze samples of overexpressed material qualitatively. It was also used to estimate the relative molecular masses of the overexpressed proteins and to monitor enzyme purification. Proteins are denatured prior to electrophoresis by heating in the presence of an anionic detergent, SDS and a reducing agent, β -mercaptoethanol. Both cause the proteins to unfold completely. SDS binds to the polypeptides giving them a high net negative charge, proportional to the length of the chain. Separation in the gel depends principally on protein molecular mass. Proteins first

pass through the stacking gel, the role of which is to concentrate the sample into a sharp band (utilizing differences in ionic strength and pH between stacking gel and electrophoresis buffer). Subsequently molecules enter the main separating gel, where they are resolved.

2.15.1. Sample Preparation

Protein samples were mixed with an equal volume of $2 \times$ Laemmli sample buffer and heated in a GRANT BT3 Dry Block Heater at about 95°C for at least 5 min. Before loading onto a gel they were centrifuged, 10,000 rpm for about 5 min. If not used immediately, they were frozen at -20°C and only heated prior to application to the gel.

2.15.2. SDS-PAGE

10% polyacrylamide gels, about 8×7 cm (including wells) and 0.7 mm thick, were prepared routinely using the Mighty Small Multiple Gel Caster SE 200 purchased from Hoefer Scientific Instruments (Amersham Pharmacia Biotech). The resolving gel was prepared by mixing 5 ml of resolving buffer, 3.3 ml of acrylamide stock solution, $50 \mu\text{l}$ of ammonium persulphate and 1.55 ml of water. The solution was degassed using a water pump and mixed gently with $100 \mu\text{l}$ of 10% SDS and $5 \mu\text{l}$ of TEMED. It was then poured promptly using the glass Pasteur pipette between the assembled plates (glass and alumina), leaving sufficient space at the top for a stacking gel (about 3 cm). The surface of the resolving gel mixture was overlaid with water-saturated butanol and the gel was allowed to set (about 1 h). After it had polymerized, the butanol overlay was removed and the gel surface rinsed with distilled water to remove traces of butanol and unpolymerized acrylamide.

The remaining space was then filled with the stacking gel solution made of 0.5 ml of stacking buffer, 0.66 ml of acrylamide stock solution, $50 \mu\text{l}$ of ammonium persulphate and

3.77 ml of water, all degassed as before and mixed gently with 50 μ l of 10% SDS and 5 μ l of TEMED. The comb was inserted immediately. After polymerisation (left for up to 1 h), the comb was removed and the wells rinsed with distilled water to remove unpolymerized acrylamide. At least 1 cm of stacking gel is necessary between the bottom of the loading wells and the resolving gel.

The gel was assembled in the Hoefer Mighty Small II vertical gel electrophoresis unit and electrophoresis buffer was added to both chambers. The wells were washed with the buffer using a glass Pasteur pipette and samples (up to 25 μ l) were applied. The gel was run at 16 mA, until the dye (bromophenol blue) present in the sample buffer, totally unretarded, reached the bottom of the gel (about 1 h). The current was switched off and the gel, rinsed with distilled water, removed from between the glass plates.

2.15.3. Staining and Destaining SDS-Polyacrylamide Gels

Proteins on SDS-polyacrylamide gels were detected with the sulphated trimethylamine dye Coomassie Brilliant Blue R-250 (see also section 2.22.1). It is dissolved in acid-methanol solution that fixes proteins in gel via denaturing and consequently precipitating them. The destain solution removes the unbound background dye from the gel, leaving stained proteins visible as clear blue bands.

The gel was shaken in Coomassie stain solution for about 2 hours and then washed in destain solution overnight (optionally, a tissue was placed in the gel box for quicker adsorption of the dye). When the gel was stained overnight, the destaining process was done the following day with the destain solution being changed every few hours.

2.15.4. Drying SDS-Polyacrylamide Gels

Drying of SDS-polyacrylamide gels was done routinely using a Camlab gel drying frame. The destained gel was soaked in 3% glycerol in destain solution for about 2 h with shaking. It was then placed between two sheets of clear Bio Design Gel Wrap membrane (moistened with 10% glycerol). After any air bubbles trapped between the gel and film sheets were removed, it was clamped between the frames and allowed to dry overnight.

2.16. Electrophoretic Transfer of Proteins (Western Blotting)

The technique of Western Blotting allows the transfer (blot) of the pattern of separated by SDS-PAGE proteins onto an immobilising hydrophobic membrane. The bound proteins are then available for analysis by a variety of specific protein-ligand interactions.

Following SDS-PAGE of crude extracts, proteins were transferred onto a PVDF, highly hydrophobic membrane using a semi-dry blotting apparatus by Pharmacia. The anode of the apparatus was moistened with UHQ water. Three pieces of Whatman 3MM chromatography paper (Whatman International Ltd), cut to size of the gel being blotted were pre-soaked in transfer buffer and were placed on the anode. A PVDF membrane, also cut to size of the gel, was pre-wetted by floating and submerging briefly (60 s) in 100% methanol, was soaked in transfer buffer for about 2 min and placed on the top of the filter papers. The polyacrylamide gel was removed from between the glass plates of the PAGE apparatus and soaked for 5 min in transfer buffer. It was then placed carefully on the membrane and covered with three extra pieces of Whatman 3MM paper pre-soaked in transfer buffer. To ensure that all the air bubbles trapped were removed, the glass rod was rolled gently on top

of the stack. Finally, it was covered with the cathode plate, moistened with UHQ water. The transfer was performed at a constant current of 100 mA. After 1 h the current was switched off and the recovered PVDF membrane allowed to air-dry at room temperature for approximately 30 min, to improve the efficiency of antigen binding in the following immunostaining experiment (section 2.17).

2.17. Immunodetection of Blotted Antigens

Immunodetection allows further analysis of blotted proteins by probing with antibodies to detect specific antigens. The primary antibody detects and binds to the protein of interest on the blot, while enzyme-linked secondary antibody (e.g. anti-guinea pig IgG conjugate with alkaline phosphatase) is directed against the primary antibody. Addition of enzyme substrate results in insoluble coloured product precipitated at the site of protein of interest on the blot.

The immunodetection was done essentially as described by Glover et al. (1997). The PVDF membrane was re-wetted by submerging briefly in 100% methanol followed by brief rinsing with distilled water. It was then incubated with 30 ml of blocking buffer for at least 1 h to saturate non-specific protein binding sites. This step, as well as all the following, were done with gentle agitation at room temperature. The membrane was then washed in 30 ml of TTBS three times for 15 min and incubated with 30 ml of TTBS containing the appropriate dilution of primary antibody (see section 2.4.11/Antibodies, p. 69) for 1 h. The membrane was then washed in TTBS as before, to remove unbound antibodies, and incubated for an extra hour with 30 ml of TTBS containing the appropriate dilution of a secondary antibody (section 2.4.11/Antibodies, p. 69). The membrane was then washed

again in TTBS and once with 30 ml of development buffer for up to 15 min to remove any traces of the detergent (Tween 20 can affect depositing of the precipitated substrate and lead to smearing of bands). The colour reaction was started by incubating the membrane in BCIP/NBT substrate solution without shaking, until the bands of interest had reached the desired intensity. To stop the reaction the membrane was washed in 10 mM EDTA and distilled water. It was then air-dried and stored protected from light.

2.18. Protein Sequence Analysis

Protein sequencing was done by Alta Bioscience using a PE Biosystems 473A automatic protein sequencer. The instrument utilises the Edman degradation method to cleave the N-terminal amino acid residue from the polypeptide as a phenylthiohydantoin derivative, which is then identified chromatographically by HPLC. The degradation/analysis cycles are repeated for subsequent aa residues.

Sample Preparation

Protein sample was run on SDS-polyacrylamide gel as described in section 2.15.2 and subsequently blotted onto a PVDF membrane, essentially according to the protocol in section 2.15.4. The only exceptions were soaking of the gel in transfer buffer for 5 min and transfer performed at a constant current of 90 mA. Following the blotting, the PVDF membrane was submerged in a blot stain for 60 s, followed by incubation in a blot destain for 30 min. It was then air-dried between sheets of Whatman 3MM paper and sent to Alta Bioscience.

2.19. Mass Spectrometry

Relative subunit molecular mass of the produced protein variants was analysed by time-of-flight electrospray ionisation mass spectrometry in the School of Chemical Sciences, University of Birmingham. A Micromass LCT time-of-flight mass spectrometer was used.

Sample Preparation

Proteins in crude extracts following overexpression were diluted appropriately with sterile UHQ water to a final protein concentration of about 5 pmol/ml in at least 10 μ l of solution. Low (less than 20 mM) salt concentration was required.

2.20. Fast Protein Liquid Chromatography (FPLC)

FPLC is a patented chromatographic system (Pharmacia) for conducting column chromatography at elevated flow rates using pressures of up to 3 MPa. Various columns (stationary phase) with a small diameter packing material may be used, depending on the type of separation that is preferred. As the solvent is forced into the stationary phase by the flow rate, the sample mixture separates based on the ability to non-covalently interact and/or diffuse into the column material.

2.20.1. Protein Purification. Ion Exchange Chromatography (IEC)

Ion exchange chromatography relies on the attractive forces between molecules carrying charged groups of opposite signs. The chromatography uses an ion exchanger with a charged particle resin that can bind sample molecules reversibly. The one used for the

protein purification was prepacked Mono Q HR 5/5 (Pharmacia), a strong anion exchanger (completely ionized at normal pH values) with a resin containing quaternary ammonium ($-\text{CH}_2\text{-N}^+[\text{CH}_3]_3$) as functional groups. Increasing salt concentration of the mobile phase was used to release and elute the bound proteins from the column matrix. Mono Q has a unique hydrophilic polyether resin formed into highly homogenous beads (10 μm average particle size) and it separates proteins up to $M_r = 10^7$.

The full-length *B. subtilis* PPase was purified using two-step ion exchange chromatography. About 10 ml of filtered (section 2.2/Filter-sterilization, p. 58) whole-cell extract in standard buffer containing 1 mM EDTA was applied to a Mono Q column, equilibrated with the same buffer. Elution was at 0.5 ml/min. For the first 35 – 40 min the column was washed at a constant, 0.14 M NaCl concentration, in order to remove residual proteins and nucleic acids. Subsequently, a linear gradient of NaCl (0.14 – 0.5 M) was applied to elute the adsorbed proteins. This process was repeated with another 10 ml of whole-cell extract. 1 ml fractions were collected and those with the highest concentration of PPase protein (visually assessed from SDS-polyacrylamide gels) were pooled and re-loaded onto the Mono Q column in a single run. The conditions used were as before, except fractions of 0.5 ml were collected. The fraction with the highest specific activity was chosen for further studies. Chromatography was performed at room temperature. 5AATR was purified using the same procedure, except the two runs of the first chromatography step were done in the gradient of 0.14 – 0.6 M NaCl. Also, the elution rate was 1 ml/min and 2 ml fractions were collected in the first two runs and 0.25 ml in the last one.

17AATR was purified in one step. 1 ml of the whole-cell extract was applied to the column equilibrated with the standard buffer containing 0.1 mM EDTA. Elution was at

1 ml/min. The column was washed at a constant, 0.15 M NaCl concentration, then a shallow gradient was applied (0.15 – 0.45 M). 0.3 ml fractions were collected. The native *E. coli* enzyme was purified from the whole-cell *E. coli* extract following the same procedure as used for 17AATR except the fraction size was 1 ml.

2.20.2. Native Molecular Mass Analysis. Size Exclusion

Size exclusion (gel filtration) chromatography utilizes the molecular sieve properties of different porous materials and the separation of molecules is based on their molecular size and shape. The resin consists of porous beads whose accessibility continuously decreases for molecules of increasing size. Consequently, the largest molecules are the quickest to pass through the column and elute first, while the smallest get retarded by the bead pores and elute last. The column used for native molecular mass analysis was Superose 12 FPLC (Pharmacia) with 10 μm average particle size and 24 ml bed volume. It separates proteins with molecular masses in the range between 1×10^3 and 3×10^5 .

Filtered (section 2.2/Filter-sterilization, p. 58) crude extracts and Mono Q purified material were loaded onto a prepacked Superose 12 column. Prior to application, the Mono Q purified material was concentrated $2 \times$ (5AATR) and $4.5 \times$ (full-length PPase) using Centricon Centrifugal Filter Devices YM-30 (Millipore) to get 0.5 ml total sample volumes. The column was equilibrated and then washed with the standard buffer with added EDTA (0.1 mM, experiment with purified material) or without it (the experiment with proteins in crude extracts). The column was used at 0.5 ml/min flow rate. Chromatography was at room temperature.

Calibration Curve

For a given column there is a fixed relationship between solute elution volume and molecular size. Consequently, a calibration curve can be obtained by plotting the logarithms of the known molecular weights of protein standards versus their respective V_e/V_o values (V_e is the elution volume and V_o is the void volume, required for the elution of a large, excluded, molecule, e.g. blue dextran). By measuring the elution volume of the protein of interest and calculating its respective V_e/V_o value, the protein's molecular weight may be estimated.

Two calibration curves were used. Both were made by John Osborne with the Gel Filtration Molecular Weight Markers, MW-GF-200 Kit (Sigma). The calibration curve obtained for the Superose column equilibrated in standard buffer had an equation of the linear regression line $y = -2.0556x + 7.8006$. The square of the correlation coefficient was $r^2 = 0.9925$. The curve was used to estimate the molecular mass of the full-length PPase and its truncated mutants in whole-cell extracts.

The calibration curve obtained for the column equilibrated with standard buffer containing 0.1 mM EDTA had an equation of the linear regression line $y = -2.1554x + 7.8177$ and the square of the correlation coefficient $r^2 = 0.9363$. This curve was used to estimate the molecular mass of the full-length PPase and the truncated mutants in the purified material.

2.20.3. Cleaning of MonoQ Column

The MonoQ ion exchange column was cleaned routinely between purification runs of different protein variants. In brief, the column was attached upside down and washed with 0.5 ml of 2 M NaCl, followed by 2 M NaOH and 75% acetic acid, all at 0.5 ml/min flow rate.

In each case the column was washed at least twice. Finally, it was washed with water for at least 10 min.

2.21. Inorganic Pyrophosphatase Activity Assay

PPase activity was assayed as described by Kuhn and Ward (1998) using sodium pyrophosphate at 0.9 mM as substrate (section 2.4.14/Substrate Solution, p. 71). 10 – 50 μl of protein sample (made up to 50 μl with basic medium, if necessary) was mixed (briefly by vortexing) with 450 μl of substrate solution and incubated for 5 min at room temperature. The reaction was stopped by addition of 45 μl of stop solution (section 2.4.14/EDTA Solution, p. 71) with brief mixing by vortexing. P_i present was determined by mixing (briefly by vortexing) the resulting solution with 0.5 ml of Fiske-Subbarow reagent and reading at 691 nm after 15 min incubation. The intensity of colour developed was the measure of PPase activity. The results were corrected for the amount of P_i in the reaction mixture. Enzyme blank (reaction stopped at zero time; usually $A_{691} = 0.03 - 0.05 \mu\text{mol Pi formed/ml protein sample}$) or reagent blank (no enzyme, basic medium instead; usually $A_{691} = 0.01 - 0.02 \mu\text{mol Pi formed/ml protein sample}$) was always included. All enzyme dilutions were made with basic medium.

Prior to assay, samples were routinely inactivated with 2 mM EDTA (30 min, 37°C) and then activated with 2 mM MnCl_2 (60 min, room temperature). With the exception of Val299Gly/Val300Trp, the inactivation step was omitted for all other site-directed mutants, since they were stored in a buffer containing 1 mM EDTA.

When activity was assayed against tripolyphosphate, it was used, as sodium pyrophosphate, at 0.9 mM.

Phosphate Standard Curve

Sodium phosphate in the range of 0 – 0.2 μmol was used as standard. The appropriate volumes of phosphate standard solution (125 – 500 μl of 0.4 mM standard) were mixed with distilled water added to adjust the final volume to 545 μl (the volume of the reaction mixture). The solution was then mixed with Fiske-Subbarow reagent and after 15 min incubation at room temperature the extent of the colour reaction was determined by reading at 691 nm. The readings were corrected for the A_{691} of the control sample, which contained 545 μl of distilled water instead of phosphate standard solution. A_{691} values obtained were plotted against the amount of P_i (μmol) and the resulting standard curve is shown in Fig. 2.3, p. 108.

2.22. Protein Assay

2.22.1. Bradford Method

Protein content of enzyme samples was determined routinely by the method based on the one of Bradford (Bradford 1976) using bovine serum albumin as standard. At low pH free Coomassie dye has its absorbance maximum at 465 nm but when it binds protein, an immediate absorbance shift occurs to 595 nm. Simultaneously, the colour change of the reagent takes place from green/brown (red/brown) to blue.

The assay was done using the Coomassie Protein Assay Reagent Kit according to the manufacturer's instructions. 20 μl of protein sample was mixed by vortexing with 1 ml of the Coomassie reagent. After about 5 min the extent of the colour reaction was determined

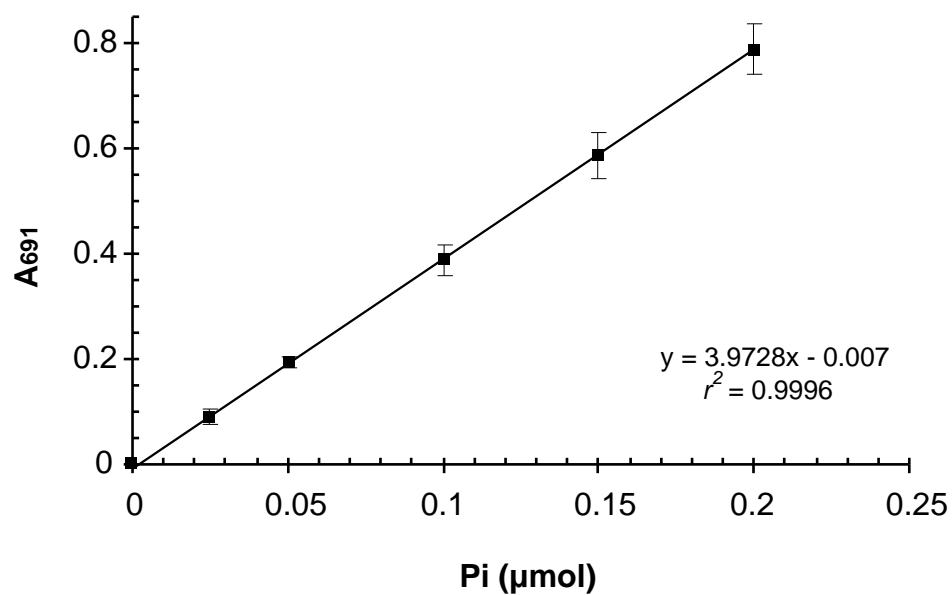


Figure 2.3: P_i standard curve. The appropriate volumes of phosphate standard solution (125 – 500 μl of 0.4 mM standard) were mixed with distilled water added to adjust the final volume to 545 μl (the volume of the reaction mixture). The solution was mixed with Fiske-Subbarow reagent and after 15 min incubation at room temperature the extent of the colour reaction was determined by reading at 691 nm. The readings were corrected for the A₆₉₁ of the control sample, which contained 545 μl of distilled water instead of phosphate standard solution. An equation of the linear regression line (y) as well as the square of the correlation coefficient (r^2) are given.

Each result represents the mean ± SD from three independent assays

by reading the absorbance at 595 nm. The readings were corrected for the A_{595} of the control sample, which contained 20 μl of the standard buffer in place of the protein.

Protein Standard Curve

A set of protein standards was prepared by diluting the 2 mg/ml albumin standard in standard buffer (section 2.4.13/Standard Buffer, p. 70) to get a range from 100 to 500 μg of protein. To each tube containing 20 μl of an appropriate standard 1 ml of the Coomassie reagent was added, the contents were mixed and, after about 5 min, the absorbance was measured at 595 nm. The values were corrected for the A_{595} of the control sample, which contained 20 μl of the standard buffer in place of the protein. The corrected values were plotted against the amount of protein ($\mu\text{g}/\text{ml}$) present and the resulting standard curve is shown in Fig. 2.4, p. 110.

2.22.2. BCA Method

The method employing bicinchoninic acid (BCA) was used as an alternative for the colourimetric detection and quantification of total protein in 17AATR and pET11c control crude extracts. The method depends on the reduction of Cu^{2+} to Cu^{1+} by protein under alkaline conditions (the biuret reaction) and the highly sensitive colorimetric detection of the cuprous cation (Cu^{1+}) in the reaction with BCA. Chelation of two molecules of BCA with one cuprous ion gives an intense purple coloured product with a strong absorbance at 562 nm. The method seems to be more reliable than the Bradford one, as the amount of dye binding does not depend purely on the content of Arg and Lys (Bradford method) but the macromolecular structure of the protein, the number of peptide bonds and the presence of

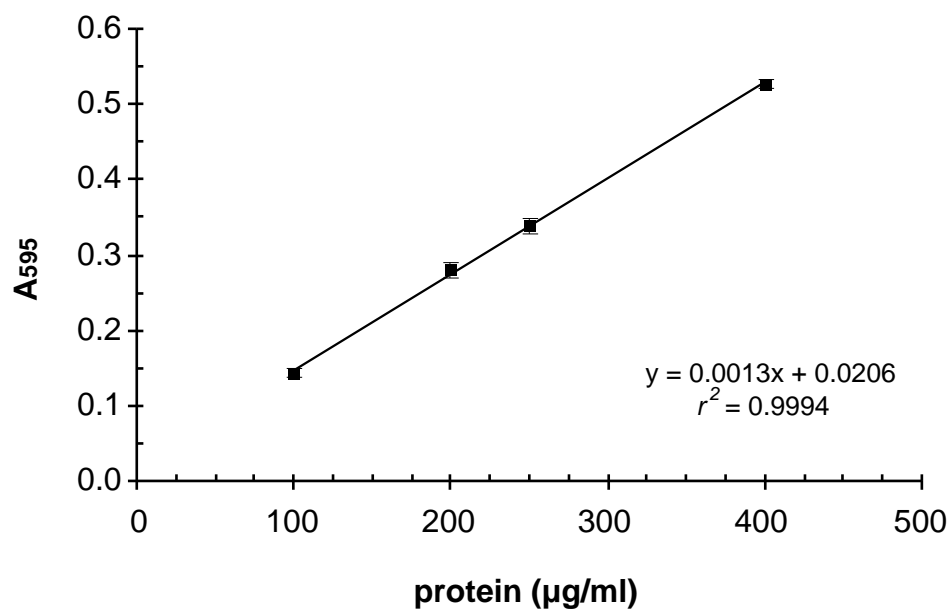


Figure 2.4: Bradford protein standard curve. To each tube containing 20 μl of an appropriate standard (prepared by diluting 2 mg/ml albumin standard in standard buffer, section 2.4.13/Standard Buffer, p. 70) 1 ml of the Coomassie reagent was added. The contents were mixed and the absorbance measured at 595 nm. The readings were corrected for the A_{595} of the control sample, which contained 20 μl of the standard buffer in place of the protein. An equation of the linear regression line (y) as well as the square of the correlation coefficient (r^2) are given. Each result represents the mean \pm SD from one assay performed in triplicate

four aa (Cys, cystine, Trp and Tyr). However, the detection limit for both methods is said to be 20 μg of protein per ml.

The Pierce BCA Protein Assay was done according to the manufacturer's instructions. 50 μl of sample was mixed by vortexing with 1 ml of BSA working reagent (50 parts of reagent A and 1 part of reagent B) and incubated 30 min at 37°C. The readings were corrected for the A_{562} of the control sample, which contained 50 μl of the standard buffer in place of the protein.

Protein Standard Curve

A set of protein standards was prepared by diluting the 2 mg/ml BSA stock in standard buffer (section 2.4.13/Standard Buffer, p. 70) to get a range from 25 to 1000 μg of protein. 1 ml of the BSA working reagent (see above) was added to each tube containing 50 μl of an appropriate standard, the contents were mixed and the absorbance measured at 562 nm. The readings were corrected for the A_{562} of the control sample containing 50 μl of the standard buffer instead of the protein. The corrected values were plotted against the amount of protein ($\mu\text{g}/\text{ml}$) present and the resulting standard curve is shown in Fig. 2.5, p. 112.

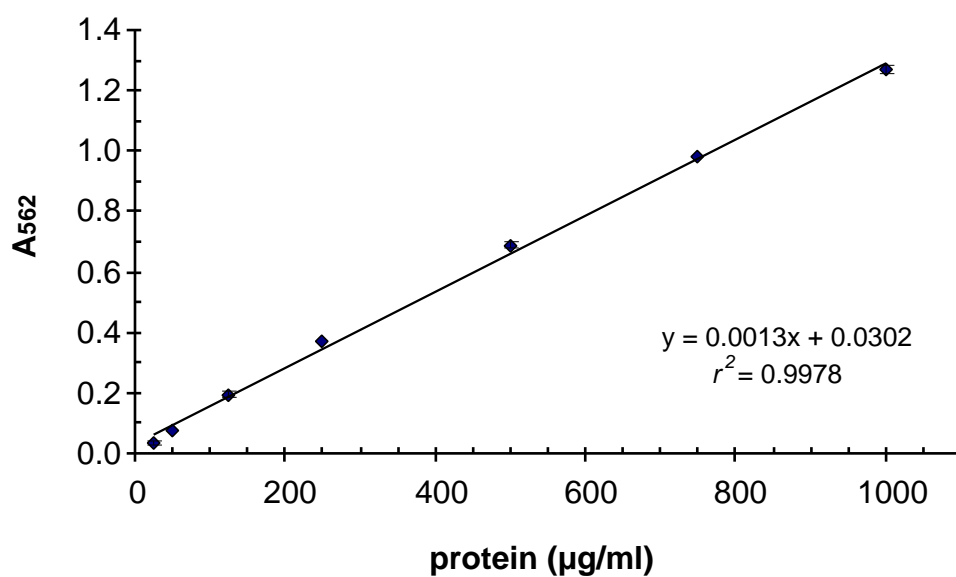


Figure 2.5: BCA assay protein standard curve. 1 ml of the BSA working reagent was added to each tube containing 50 μ l of an appropriate standard, prepared by diluting 2 mg/ml BSA stock in standard buffer (section 2.4.13/Standard Buffer, p. 70). The contents were mixed and the absorbance measured at 562 nm. The readings were corrected for the A₅₆₂ of the control sample containing 50 μ l of the standard buffer instead of the protein. An equation of the linear regression line (y) as well as the square of the correlation coefficient (r^2) are given. Each result represents the mean \pm SD from at least two assays performed in duplicate

Chapter 3

The C-terminus is Essential for Enzyme Activity but Not Subunit Composition

3.1. Introduction

Bacillus subtilis inorganic pyrophosphatase is the first member of the newly identified Family II of PPases (Young et al. 1998, Shintani et al. 1998). The degree of sequence similarity found among proteins of this group is notable (over 50%). Furthermore, these sequences are totally unrelated to the sequences of the well-studied proteins of Family I PPases.

At the time I was starting the project, DNA sequence analysis was the only source of information about the *Bacillus subtilis* enzyme and the other new family members. Sequence alignment of the other four representatives identified at the time, namely *Streptococcus gordonii*, *Methanococcus jannaschii*, *Archaeoglobus fulgidus* (Young et al. 1998) and *Streptococcus mutans* (Shintani et al. 1998), revealed a number of extended regions of similarity (Fig. 1.6, p. 52). However, one of the most remarkable conservations in all members was found near the C-terminus. It was eight residues long (294 – SRKKQVVP – 301; *B. sub-*

tilis numbering) in all the members and proposed as a fingerprint (Shintani et al. 1998) or signature for *Bacillus* type (Family II) PPases (Fig. 5.1, p. 157).

To examine the role of this conserved region, the coding sequence of *B. subtilis* PPase was used to synthesise two different products with truncations at the 3' end. The products were cloned and C-terminally truncated proteins expressed in *E. coli*. Truncations removed from the C-terminus 5 non-conserved residues (5AATR) and 17 amino acids (17AATR) including the conserved assembly, the signature sequence of Family II enzymes. A number of methods were employed for the initial characterisation of mutants. Truncated proteins were used to assess the roles of the C-terminal region in both enzyme activity and protein tertiary structure of the *B. subtilis* PPase.

3.2. Generation of Truncated Mutants

3.2.1. PCR and Cloning

Two 3' truncated variants of the *Bacillus subtilis* PPase gene were synthesised from the *B. subtilis* PPase coding sequence (gene *yybQ*) using PCR. The PPase clone (*yybQ* in pET11c) was used as template. The oligonucleotide primers used are indicated in Materials and Methods (see section 2.11.1/Truncated Mutants, p. 88). PCR products of expected length, 915 for 5AATR and 879 bp for 17AATR (Fig. 3.1, p. 117), were cloned into the plasmid vector pET11c.

Both products and pET11c were digested with *Bam*HI and *Nde*I. The linearized vector DNA was additionally treated with alkaline phosphatase (section 2.7.2). Ligation was overnight at room temperature at the vector/insert ratios 1:9 for 5AATR and 1:3

for 17AATR. The ligated DNA was transformed into *E. coli* DH5 α , for maintenance and storage, with a yield of 10 and 12 colonies for 5AATR and 17AATR, respectively.

Following the transformation into DH5 α the transformant DNA was screened for the presence of mutant inserts by PCR using the commercially available T7 Terminator and T7 Promoter primers (Fig. 2.2, p. 93) as well as the primers used for cloning. Products were also analysed by restriction digest using *EcoRV*, *Asp700* and *SacI* with *HindIII*. The obtained restriction pattern agreed with the predicted one shown in Fig. 2.2, p. 93. Finally, the mutant DNA was sequenced in both directions (see an example in Fig. 5.4, p. 169). The 910 bp fragment was found to be identical to the parent *yybQ* gene except for being 15 nucleotides shorter as designed, and terminating at position 915 (*yybQ* sequence numbering) or Thr304 (Fig. 3.2, p. 117). The 880 bp fragment was, as expected, 51 nucleotides shorter than the parent sequence, terminating at position 879 or Val292 (Fig. 3.2, p. 117). However, this sequence was not identical to the parent and contained one mutation, C for T at position 844. This mutation generated the CTA codon in place of TTA in the parent sequence. As both triplets code for lysine, the translated amino acid sequence was therefore identical to that produced from *yybQ*. For overexpression of the truncated proteins the mutant DNA isolated from DH5 α was transformed into *E. coli* BL21(DE3).

3.2.2. Overexpression, Protein Isolation and Immunostaining

Overexpression of truncated coding sequences cloned in pET11c was done at 37°C and with shaking. Cells were induced for the production of the proteins by IPTG. Overexpression gave two C-terminally truncated *B. subtilis* PPase mutants: 5AATR lacking the 5 non-conserved amino acids and 17AATR that additionally lacked 12 more residues covering

the highly conserved signature region. The C-terminal sequences of the truncated mutants produced in this study are shown in Fig. 3.2, p. 117. After 2-hour induction, single major species were identified by SDS-PAGE in the crude extracts of *E. coli* BL21(DE3) (Fig. 3.3 a, p. 118). About 0.7 g of wet cell paste was obtained from 200 ml of cell culture. Extending the induction period to 3 hours resulted in a significant degradation of both mutant proteins (Fig. 3.3 b, p. 118).

During 2 h overexpression the cells were screened for the formation of inclusion bodies: every 20 min samples were withdrawn from the cultures and checked under the microscope. The formation of inclusion bodies started between 40 and 60 min of induction. The overexpressed protein was extracted from the *E. coli* cells using sonication. The sonicated material was centrifuged (to remove cell debris and any inclusion bodies) and the overexpressed soluble PPase recovered in the supernatant. This preparation is referred to throughout as the whole-cell (crude) extract. As judged by SDS-PAGE, pellets contained negligible amounts of the enzyme (Fig. 3.4, p. 119). The full-length PPase was overexpressed following a similar procedure as used for the truncated proteins.

The truncated proteins reacted strongly with antibodies raised to purified overexpressed *B. subtilis* PPase (Fig. 3.5, p. 119).

3.2.3. Relative Subunit Molecular Mass

The relative subunit molecular masses of 5AATR and 17AATR were determined in crude extracts using TOF MS ES. MS analysis gave single major peaks in the spectrum (example shown in Fig. 3.6, p. 121) and the values were in excellent agreement with those calculated

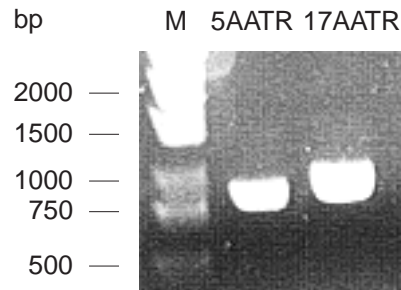


Figure 3.1: Truncated PCR products (0.8% agarose). M – Fermentas 1kb DNA ladder

| | | |
|-------------------|---------------------------------------|-----|
| Full-length PPase | LENNTALLKGVV SRKKQVVP VLTDAMAE | 309 |
| 5AATR | LENNTALLKGVV SRKKQVVP VLT----- | 304 |
| 17AATR | LENNTALLKGV----- | 292 |

Figure 3.2: The C-terminus of truncated *B. subtilis* PPase mutants. The 'SRRK' motif is in *bold*, the postulated fingerprint (Shintani et al. 1998) or signature sequence is *boxed*. The total numbers of amino acids in each protein are indicated on the right

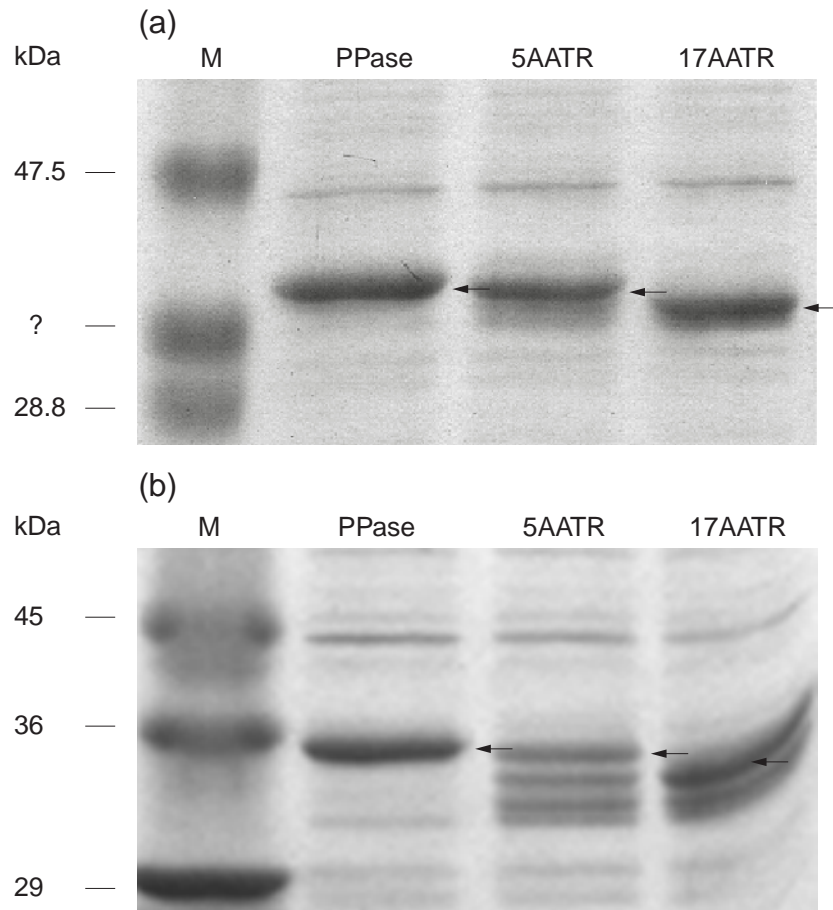


Figure 3.3: Overexpression of truncated *B. subtilis* PPase mutants: (a) SDS-PAGE of proteins following 2-hour induction with IPTG at 37°C (10 μ l of 4 \times diluted whole-cell extract loaded), (b) SDS-PAGE after 3-hour induction (10 μ l of 5 \times diluted whole-cell extract loaded). M – protein marker: BioRad (see section 2.4.9/Protein Standard, p. 66) and Sigma, respectively. Arrows indicate overexpressed proteins (or the main protein bands). Proteins were stained with Coomassie Brilliant Blue

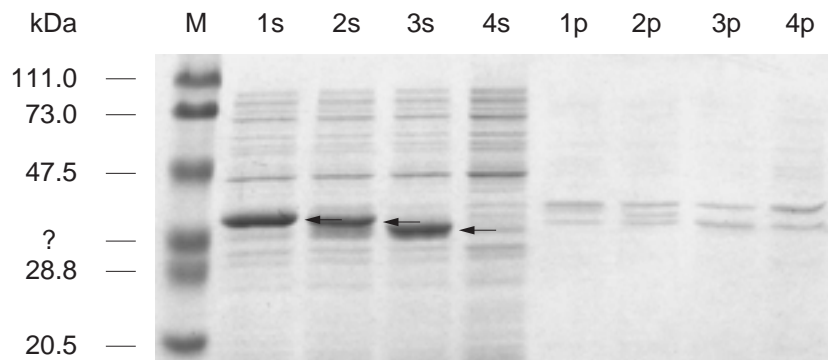


Figure 3.4: SDS-PAGE of supernatant (whole-cell extract, s) and pellet (p) material after sonication following 2-hour overexpression. M – BioRad protein marker (see section 2.4.9/Protein Standard, p. 66), 1 – full-length *B. subtilis* PPase, 2 – 5AATR, 3 – 17AATR, 4 – *E. coli* control. 10 μ l of 4 \times diluted supernatant and 2 \times diluted pellet materials were loaded. Arrows indicate overexpressed proteins. Proteins were stained with Coomassie Brilliant Blue

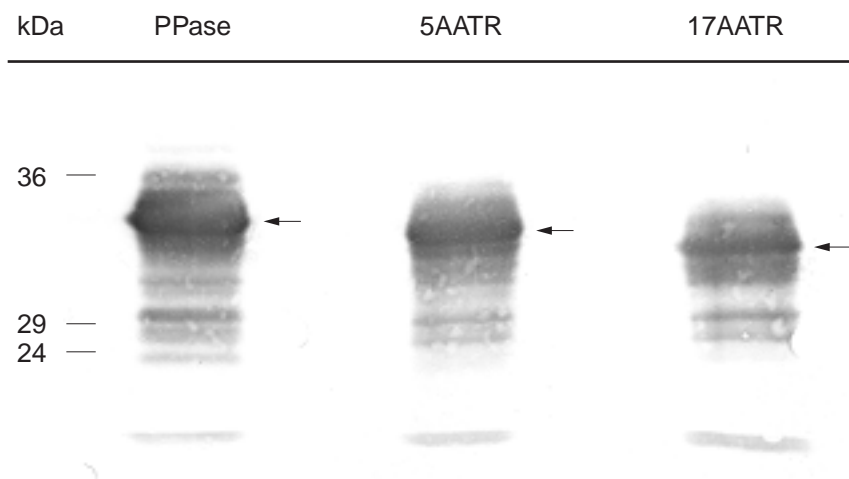


Figure 3.5: Immunostaining of *B. subtilis* PPase and both truncated mutants. Arrows indicate the position of overexpressed proteins as seen on Coomassie Blue stained SDS-polyacrylamide gels. Immunostaining was done using polyclonal antibody against purified overexpressed *B. subtilis* PPase as indicated under Materials and Methods (section 2.17)

from sequence data (Table 3.1). The values estimated from SDS-PAGE by comparison with standards (Fig. 3.3 a, p. 118), corresponded to the ones determined by MS.

| Enzyme variant | Relative Subunit Molecular Mass (M_r) | |
|----------------|---|--------------|
| | Calculated ¹ | Experimental |
| PPase | 33,989 | 33,988 |
| 15AATR | 33,471 | 33,471 |
| 17AATR | 32,135 | 32,136 |

Table 3.1: Molecular mass of *B. subtilis* PPase and its truncated mutants measured by TOF MS ES

3.2.4. Purification

Full-length PPase and 5AATR

The full-length *B. subtilis* PPase was purified using two-step ion exchange chromatography. Half of the whole-cell extract was applied to a Mono Q column equilibrated with 50 mM TRIS containing 1 mM EDTA, pH 8.0. For the first 35 – 40 min the column was washed with buffer at a constant, 0.14 M NaCl concentration. Subsequently, a linear gradient of NaCl (0.14 – 0.5 M) was applied to elute the adsorbed proteins. This process was repeated with the remaining whole-cell extract. The purity of collected fractions in these two runs was assessed using SDS-PAGE and those with the highest concentration of PPase protein (visually assessed) were pooled and re-loaded onto the Mono Q column in a single run, using the conditions as before. A peak of PPase activity eluted from the column at an approximate NaCl concentration of 0.375 M. The fraction with the highest specific activity was chosen for further studies.

¹Subunit molecular mass calculated from the amino acid sequence

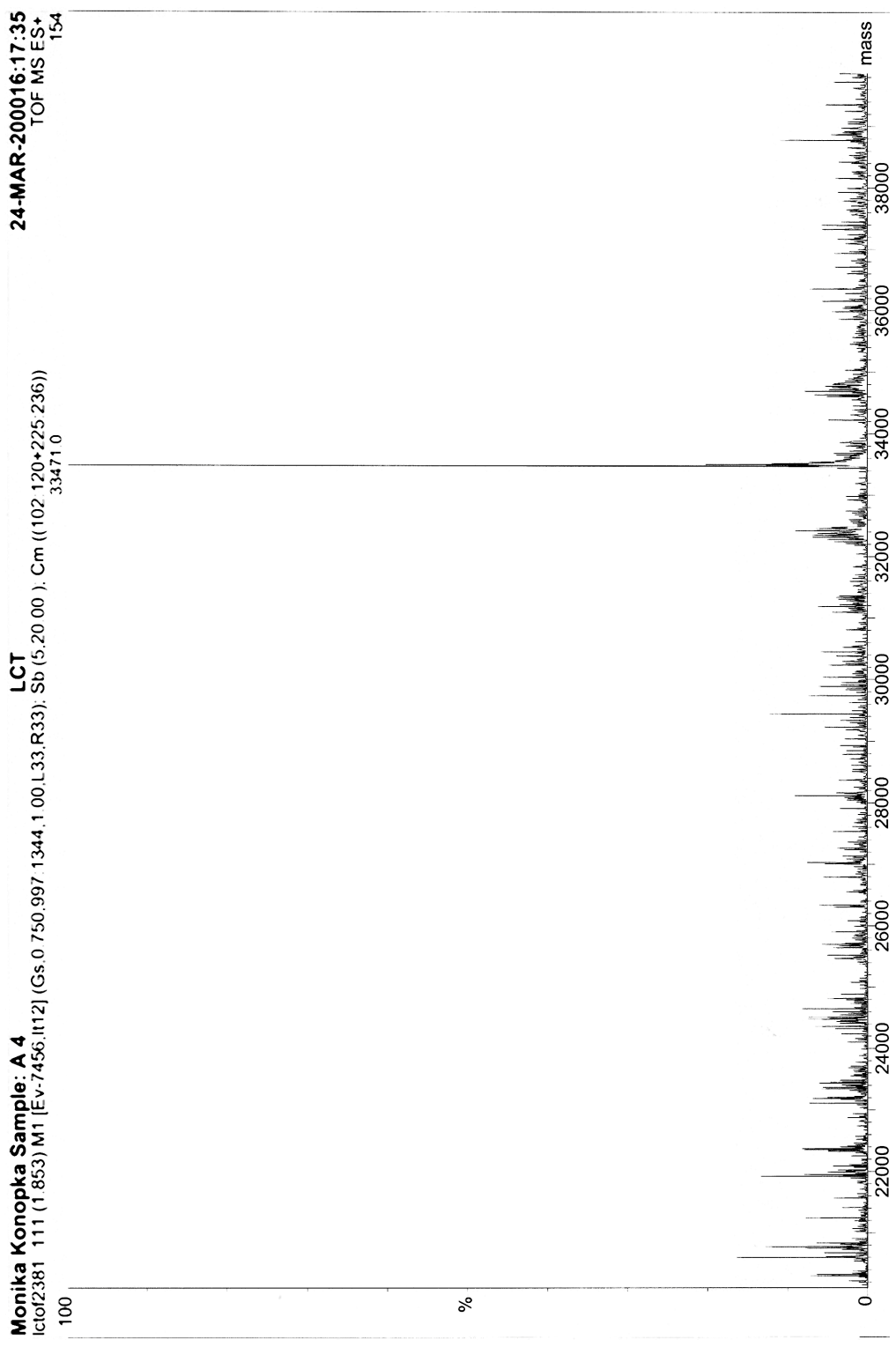


Figure 3.6: TOF ES mass spectrum of 5AATR. The analysis was done in the School of Chemical Sciences, University of Birmingham. A Micromass LCT time-of-flight mass spectrometer was used

5AATR was purified using the same procedure, except the two runs of the first chromatography step were done in the gradient of 0.14 – 0.6 M NaCl. A peak of PPase activity eluted from the column at an approximate NaCl concentration of 0.315 M.

The results of the purification procedures are presented in Table 3.2 and Fig. 3.7, p. 124, showing the proteins at consecutive stages of purification. The 5AATR was more difficult to purify. The final purification factor was about 5.8 for PPase and 5.5 for 5AATR, while the recovery of the enzymes – 6.8% and 6.7%, respectively, all calculated on the basis of the activity in the crude extracts. SDS-PAGE gels stained with Coomassie brilliant blue showed single major bands for the purified fractions and they were at least 95% pure for untruncated PPase (visually assessed) and 90% for 5AATR.

Gel filtration was tried (not shown) as a final, third purification step, but was not satisfactory. No improvement in specific activity was obtained. The specific activity, if not lower, remained of the same value, while the purification yield decreased. The elution profile did indicate further purification due to separation of the low molecular weight, UV absorbing material, eluting as a clear separate peak at the end of the run (not shown). Similarly, an attempt to use an alternative method, hydrophobic interaction chromatography (phenyl-Superose column) was unsuccessful (not shown). The protein did not bind to the resin (buffer as used for ion-exchange chromatography) and most of it would pass through the column eluting in the flow-through fractions.

17AATR and *E. coli* Control

17AATR was also purified using Mono Q. The whole-cell extract was applied to the column equilibrated with 50 mM TRIS containing 0.1 mM EDTA, pH 8.0 only once, in a one step purification. First, the column was washed with the buffer at a constant, 0.15 M NaCl

| Purification step | Total protein (mg) | Total activity (μ mole Pi/ml/min) | Specific activity (μ mole Pi/mg) | Yield (%) | Purification factor |
|-------------------|-----------------------|---|--|--------------|---------------------|
| PPase | | | | | |
| Crude extract | 36.6 | 105 561 | 2 873 | 100 | 1 |
| First MonoQ | 2.7 | 25 317 | 9 367 | 24 | 3.3 |
| Second MonoQ | 0.4 | 7 221 | 16 561 | 6.8 | 5.8 |
| 5AATR | | | | | |
| Crude extract | 27.4 | 14 959 | 545.3 | 100 | 1 |
| (fr. 31+31) | | | | | |
| First MonoQ | 3.8 | 6 609 | 1 752 | 44.2 | 3.2 |
| Second MonoQ | 0.3 | 998 | 3 005 | 6.7 | 5.5 |

Table 3.2: Purification summary of *B. subtilis* PPase and 5AATR mutant

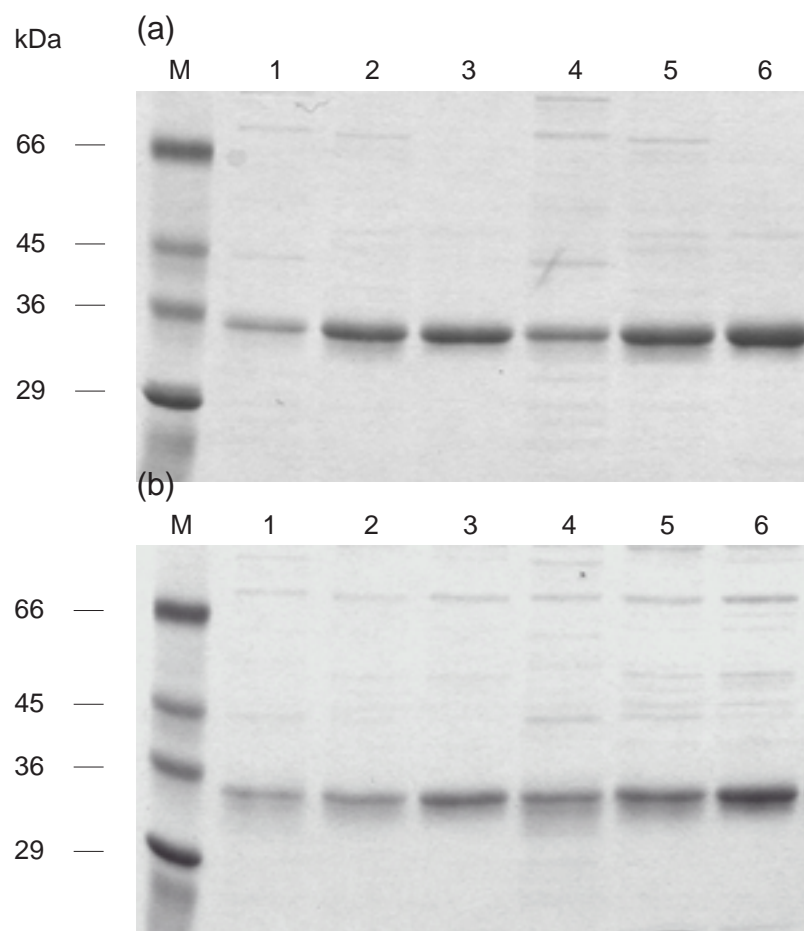


Figure 3.7: Purification of *B. subtilis* untruncated PPase (a) and 5AATR mutant (b): M – Sigma protein marker, 1 and 4 – crude extract, 2 and 5 – first MonoQ, 3 and 6 – second MonoQ. 1.0 μg (lanes 1 – 3) and 1.5 μg (lanes 4 – 6) of protein were loaded. The SDS-polyacrylamide gel was stained with Coomassie Blue

concentration, then a shallow gradient of NaCl was applied (0.15 – 0.45 M). A peak of PPase protein eluted from the column at an NaCl concentration of 0.29 M and retention volume of approximately 31 ml and was clearly separated from a peak of activity that eluted 1.8 ml earlier, at approximately 0.285 M NaCl (Fig. 3.8, p. 126). The elution profiles were the same for the protein with (not shown) and without (Fig. 3.8, p. 126) Mn^{2+} activator.

The estimation of the purity of 17AATR was difficult due to the lack of the mutant activity (see section 3.3.2). The purified fractions showed single major bands on SDS-PAGE gels stained with Coomassie brilliant blue and were at least 90% pure (Fig. 3.9 b, p. 126). 17AATR was even more difficult to purify than 5AATR. SDS-PAGE showed the difference between partially purified PPase and 17AATR material, with 17AATR clearly accompanied by other co-purifying proteins (Fig. 3.9 a, lanes 2 and 3, p. 126).

The *E. coli* PPase was purified from the whole-cell *E. coli* extract following the same procedure as used for 17AATR. A peak of activity eluted from the column at an approximate 0.27 M NaCl concentration (27 ml retention volume, Fig. 3.10, p. 134). The elution profiles, as for 17AATR, were based on the activity assays for the protein not activated with Mn^{2+} however when repeated for the one pretreated with Mn^{2+} , the activity values remained the same. The protein was only partially purified and not recognizable yet on gels by SDS-PAGE. However, this one step chromatography gave the purification factor of about 16 and 50% yield (not shown).

Elution profiles in terms of enzyme activity differed for 17AATR and the *E. coli* enzyme. Interestingly, an additional, experimental run for the mutant and *E. coli* enzyme extracts mixed together (1:1 by volume) showed only one 'peak' that corresponded to the pyrophosphatase activity (Fig. 3.10, p. 134). Overall, the elution profiles of *E. coli* and *E. coli* mixed

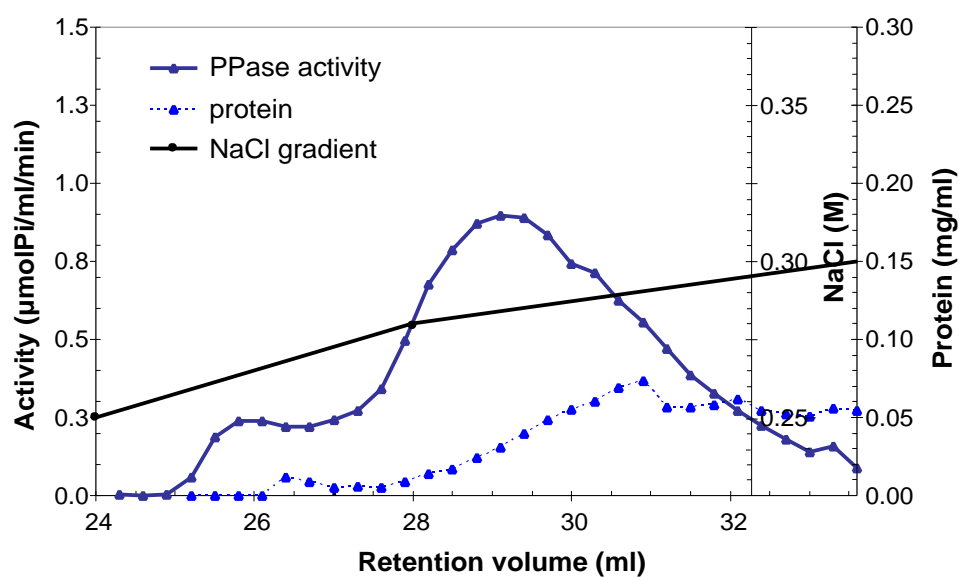


Figure 3.8: Elution profile of protein and pyrophosphatase activity detected in the 17AATR material following MonoQ ion-exchange chromatography

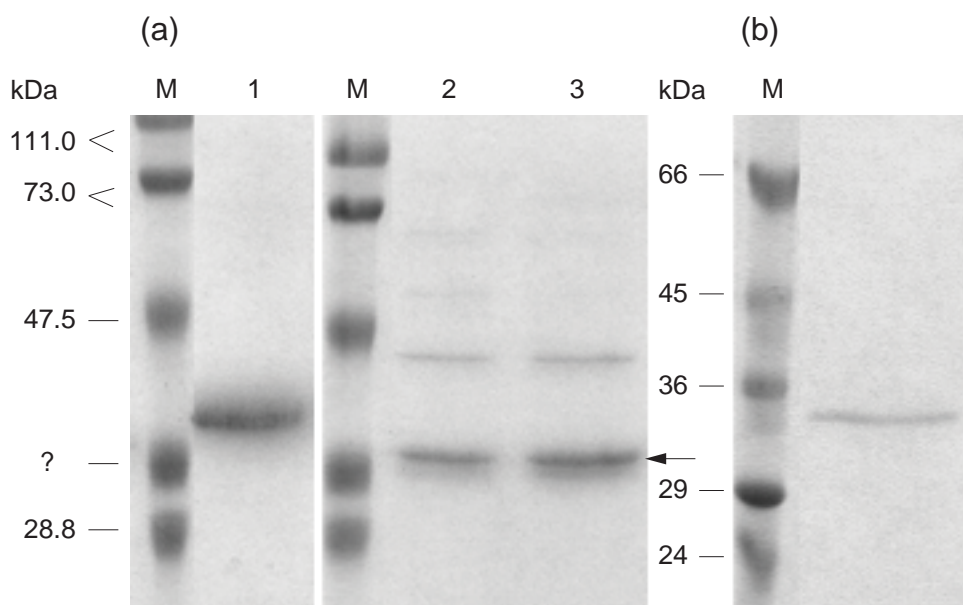


Figure 3.9: Purification of 17AATR: (a) comparison of one-step purification of full-length *B. subtilis* PPase (1) and 17AATR (indicated with an arrow: 2 – the highest activity fraction, 3 – fraction with the highest concentration of protein), M – BioRad protein markers (see section 2.4.9/Protein Standard, p. 66), (b) purified 17AATR, M – Sigma protein marker

with 17AATR in terms of enzyme activity looked very similar (retention volume about 27 ml) and differed from the one of 17AATR alone (retention volume 29 ml, Fig. 3.10, p. 134).

3.3. Enzyme Activity

Bacillus subtilis PPase catalyses hydrolysis of MgPP_i and its activity is absolutely dependent on the presence of Mn^{2+} . Kuhn and Ward (1998) reported that at pH 7, unless a low concentration of Mn^{2+} is present, the accumulation of P_i is time-limited. However, at pH 9 P_i accumulation remains linear over several minutes in the presence of EDTA (absence of free Mn^{2+}). Therefore, prior activation by low concentrations of Mn^{2+} as well as pH 9 were adopted as standard conditions for a routine activity assay.

The overexpression system used for the production of *Bacillus subtilis* PPase and its variants was not ideal in regard to inactive or low-activity mutants. *E. coli* BL21 strain had an intact gene for its native PPase. Consequently, protein isolation after expression yielded overproduced *B. subtilis* PPase together with a native *E. coli* enzyme. In a crude extract of overexpressed *B. subtilis* enzyme *E. coli* activity is insignificant, constituting only 0.1 % of that of the *Bacillus* protein. However, to differentiate between *E. coli* background activity and that of *Bacillus* mutants, the different effects of EDTA and NaF on the two enzymes were exploited. EDTA inhibits the *Bacillus* enzyme but has no effect on the *E. coli* one (Young et al. 1998). NaF has the opposite effect and fluoride insensitivity is unique to family II (Young et al. 1998, Kuhn et al. 2000). Finally, Mn^{2+} has no activation effect on the activity of *E. coli* enzyme (in fact, seems inhibitory, see Table 3.3, p. 130).

3.3.1. The Activity of 5AATR is Significantly Impaired

The highest specific activity observed for the purified untruncated *Bacillus* PPase was 16 560 $\mu\text{mol Pi formed/mg protein}$ (protein concentration used for activation with Mn^{2+} was 1.45 $\mu\text{g/ml}$, see section 4.2.1/Protein Concentration-Dependent Mode of Activation, p. 141). The pyrophosphatase activity recorded in the extract containing 5AATR was significant but specific activity constituted only about 20% of the value recorded for the full-length protein. However, the highest specific activity measured for the purified 5AATR was 2690 $\mu\text{mol P}_i/\text{mg}$ (protein concentration used for activation with Mn^{2+} was 2.66 $\mu\text{g/ml}$), only 16% of that measured for the original PPase.

3.3.2. 17AATR is Not Active

Activity detected in a crude extract of 17AATR was 2.3 $\mu\text{mole P}_i \text{ formed/min/mg protein}$. This was of the same order as that of extracts of *E. coli* BL21(DE3) transformed with pET11c alone (*E. coli* control, 4.7 $\mu\text{mole P}_i / \text{min/mg}$). Additionally, when this activity was assayed for the effect of EDTA and NaF, its properties were indistinguishable from those of the control *E. coli* enzyme in crude extract. It was insensitive to EDTA, inhibited by fluoride and not activated by Mn^{2+} (Table 3.3, p. 130).

At this stage however, in the absence of information, it seemed important to determine if the mutant protein, though carrying quite a serious change to its sequence and predictably – structure, was fully inactivated. We wanted to exclude the possibility that the activity detected in the crude extract could be due to overexpressed 17AATR. The activity assay of the purified 17AATR material showed barely detectable activity, that is 0.03 $\mu\text{mole P}_i \text{ formed/min/mg}$ being 0.02% of the untruncated PPase activity. The presence of 17AATR

in these purified fractions observed by SDS-PAGE was confirmed by N-terminal sequencing. When the activity was assayed for the effect of Mn^{2+} and EDTA, again, its properties were indistinguishable from those of the control partially purified *E. coli* extract. The activity was insensitive to EDTA, inhibited by fluoride and not activated by Mn^{2+} (Table 3.3).

3.4. Native Molecular Mass

Size exclusion chromatography, activity assay and SDS-PAGE of eluted fractions were used to estimate the relative molecular masses of the proteins. The chromatography was conducted both on the purified material in the presence of 0.1 mM EDTA and on the crude extracts – in its absence.

The results are shown in Table 3.4, p. 131. The experimentally determined maximum coefficient of variation for size exclusion chromatography was 11%. In the crude extract (and absence of EDTA) 5AATR was indistinguishable from a tetramer, while untruncated PPase showed a 13% smaller and 17AATR – an 11% greater molecular mass than calculated for proteins with tetrameric structures. In the purified material (and presence of EDTA) PPase and 17AATR were indistinguishable from tetramers and 5AATR was about 11% smaller than calculated for tetrameric protein. Overall, EDTA treatment did not materially affect the result as values of M_r from whole cell extracts and purified material were not significantly different. M_r from gel-filtration was consistent with each of the three proteins being multimers, most likely homotetramers. The more variable estimates made in crude extracts may result from protein-protein interactions.

| Enzyme | Relative activity* (%) | | | | | | | |
|------------------------|------------------------|------------|-----------------------|-------------|-------------|--------------|-----------|------------|
| | None | | 2 mM Mn ²⁺ | | 2 mM EDTA | | 1 mM NaF | |
| | extract | MonoQ | extract | MonoQ | extract | MonoQ† | extract | MonoQ |
| 17AATR | 100 ± 4.6 | 100 ± 45.8 | 69.2 ± 1.8 | 31.1 ± 3.8 | 99.2 ± 1.8 | 161.1 ± 21.4 | 3.9 ± 0.3 | 13.3‡ |
| <i>E. coli</i> control | 100 ± 3.9 | 100 ± 1.8 | 82.7 ± 1.9 | 116.4 ± 2.3 | 104.9 ± 4.2 | 110.6 ± 0.6 | 3.2 ± 0.1 | 10.5 ± 1.7 |

Table 3.3: Effect of manganese, EDTA and fluoride on *E. coli* PPase and 17AATR mutant. Pyrophosphatase activity was assayed in the crude extract and purified material (MonoQ) as described in Materials and Methods (section 2.21) without any pretreatment (none), after an exposure to manganese ions (routine assay conditions: manganese treatment preceded by an exposure to EDTA), following an extra EDTA treatment (an extra exposure after the activation) and in the presence of NaF. Each result represents the mean ± SD from four separate measurements (extracts, except two measurements in the assay with NaF) or three separate measurements (MonoQ); *, specific activity detected in the material of untreated 17AATR (3.25 ± 0.1, extract and 7.5 ± 3.4, MonoQ) or *E. coli* control (5.4 ± 0.2, extract and 131.6 ± 2.3, MonoQ) was taken as 100%, †, enzyme activity was assayed after an exposure to EDTA as a part of the routine activation procedure with manganese, ‡, the result represents one measurement only

| Enzyme variant | $10^{-3} \times M_r(\text{Da})$ | | |
|----------------|---------------------------------|--------------------|-------------------|
| | Calculated from sequence | By gel filtration* | |
| | | Whole-cell extract | Purified material |
| PPase | 33.9 | 120 | 129 |
| 5AATR | 33.5 | 136 | 120 |
| 17AATR | 32.1 | 143 | 124† |

Table 3.4: Calculated and experimental relative subunit molecular masses of the full-length PPase and truncated mutants; *, measurements were done for whole-cell extracts (buffer without 0.1 mM EDTA) and one-step Mono Q purified material (buffer with 0.1 mM EDTA) as described in Materials and Methods (see section 2.20.2), †, material before loading was pre-incubated with 5 mM EDTA, pH 8.0 for 60 min

3.5. Conclusions

Generation of two truncated *Bacillus subtilis* PPase mutants was an important step towards elucidation of the role of the C-terminal region in *B. subtilis* PPase activity.

From all the experiments done the conclusion is that the *Bacillus subtilis* PPase mutant lacking 17 C-terminal residues (V293-E309, Fig. 3.2, p. 117) did not exhibit any pyrophosphatase activity and that the low levels detected in its preparations were of *E. coli* origin only. Besides the fact that properties of the *E. coli* enzyme and that in 17AATR preparation are indistinguishable (inhibition by fluoride, insensitivity to EDTA and lack of activation by manganese, see section 3.3.2, Table 3.3), this conclusion is supported by the following observations: (a), during purification the peak of enzyme activity was never coincident with the peak of 17AATR protein (Fig. 3.8, p. 126) and (b), the maximum amount of 17AATR by SDS-PAGE did not coincide with maximum enzyme activity (Fig. 3.9 a, p. 126). The lack of activity observed for 17AATR is consistent with what the sequence alignments of family II members implied, that the C-terminus with its proposed signature sequence must

be absolutely essential for enzyme function. 17 C-terminal amino acids must be crucial to retain the native, catalytically active form of the inorganic pyrophosphatase. Interestingly however, removal of the five non-conserved amino acids from the C-terminus significantly impairs the enzyme activity and there is at least a five-fold decrease in specific activity.

Gel-filtration experiments gave consistent results implying that the native and truncated proteins are homotetramers. The sizes of truncated proteins and the full-length PPase are indistinguishable by this technique. Therefore it appears that the C-terminus has no direct role in the tertiary structure of the proteins, while both its conserved and non-conserved regions are essential for full enzyme activity.

The highest specific activity recorded for the untruncated *Bacillus* PPase exceeded by more than a factor of 2.5 the specific activity reported by Young et al. (1998). The discrepancy could be explained by the observation of Kuhn and Ward (1998) that the activity depends on the enzyme concentration used for the activation (see section 4.2.1/Protein Concentration-Dependent Mode of Activation, p. 141). The difference in specific activity between the purified PPase and 5AATR could be possibly even more pronounced (greater than the recorded 16%) if the enzymes were assayed following the activation with Mn^{2+} at exactly the same protein concentrations.

As 17AATR is inactive, the activity detected in its crude extracts is of *E. coli* origin. Consequently, the retention volume in terms of the activity detected in the whole-cell extracts of 17AATR should be the same as that for *E. coli* control extracts and mixture of both. This was not the case, and a shift in the elution profile was observed in the extract of 17AATR when compared with that of *E. coli* control (29 ml for 17AATR, 27 ml for *E. coli* and *E. coli*/17AATR mix). This may be explained if protein-protein interactions are

induced by presence of the high levels of mutant protein. Clearly, when present in excessive amounts (extracts with the overexpressed 17AATR) the mutant protein's interaction with the *E. coli* enzyme affects the elution pattern of the latter. In line with the above observations, SDS-PAGE showed that truncation seems to result in more background proteins co-purifying with mutant proteins (Fig. 3.7 a, lane 3 and 6 compared with b, lane 3 and 6, p. 124 and Fig. 3.9 a, lane 2 and 3, p. 126). This could be due to some truncation-induced changes to the surface of the protein molecule, eg. exposure of some hydrophobic patches, which promote protein-protein interactions.

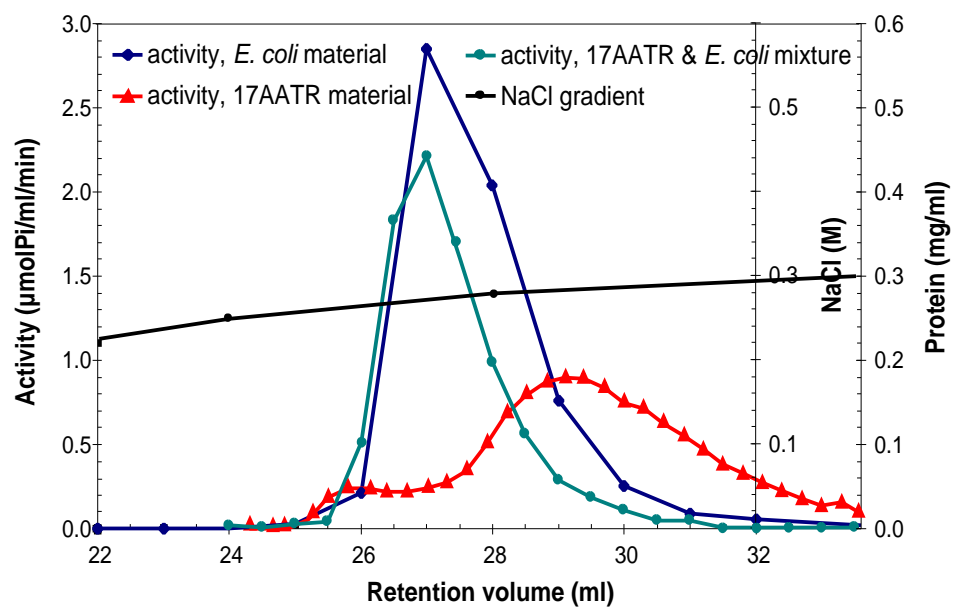


Figure 3.10: Comparison of elution profiles of pyrophosphatase activity detected in the *E. coli* and 17AATR mutant materials following MonoQ ion-exchange chromatography. Samples were run individually or in combination (17AATR & *E. coli*)

Chapter 4

Five Non-Conserved C-terminal Residues Are Crucial for Overall Enzyme Stability

4.1. Introduction

One of the most prominent sequence conservations in the representatives of Family II inorganic pyrophosphatases is found near the C-terminus (section 3.1, Fig. 5.1, p. 157). However, the residues closest to the C-terminus are themselves not conserved. Construction of the *Bacillus subtilis* variant lacking the last five C-terminal amino acids (5AATR) showed that though not conserved these residues were important for activity (Chapter 2.22.2). To elucidate the possible role of these residues in enzyme activity, 5AATR was further analysed.

The most characteristic features of the *Bacillus* PPase activity are Mn^{2+} activation (Kuhn and Ward 1998), inhibition by EDTA (Young et al. 1998) and insensitivity to fluoride (Young et al. 1998, Kuhn et al. 2000). Kuhn and Ward (1998) reported that separate activation by low concentrations of Mn^{2+} ions is an absolute requirement for the hydrolysis of the substrate, MgPP_i . The activation would convert the enzyme from a catalytically

inactive to a catalytically active form. Moreover, they noticed that Mn^{2+} activation was affected by enzyme concentration. This unusual Mn^{2+} dependence was one of the major enzyme characteristics that distinguished the *Bacillus* protein from other known PPases and eventually led to a new classification of PPases (Young et al. 1998). Prior activation by low concentrations of Mn^{2+} was adopted as a standard condition for a routine activity assay of *Bacillus subtilis* inorganic pyrophosphatase (see section 3.3).

In order to determine to what extent the truncation affected the intrinsic characteristics of the *Bacillus* PPase, activity of 5AATR was assayed following preincubation with EDTA and Mn^{2+} and in the presence of fluoride. Different parameters of the reactions were tested, such as metal and protein concentration dependence of the activation and the time-courses of activation and deactivation. The stability of truncated mutants, in terms of a protein and enzyme activity, was also monitored. The results obtained were analysed in the light of computer predictions of the secondary structure for the C-terminal region of *Bacillus* PPase. Activity assays on the longer chain phosphate substrate, tri-polyphosphate, shed some extra light onto the possible nature of the changes induced by the 5 amino acid C-terminal truncation.

4.2. The 5AATR Mutant Retains the Main Characteristics of the Untruncated Enzyme

4.2.1. Activation by Mn^{2+}

Preliminary Assays

To find the best method to assay the enzyme in the presence of Mn^{2+} two different ways of diluting the *Bacillus* PPase as well as different concentrations of EDTA and Mn^{2+} were

used. The protein was diluted (up to a required fold) ‘on incubation’, that is with the buffer containing EDTA, and then a standard, 5-fold dilution with MnCl_2 followed. The alternative method ‘on assay’ involved a standard 2-fold dilution with EDTA, followed by a standard 5-fold dilution with MnCl_2 , and only additional dilution (up to a required fold) preceding the assay. The activity was measured both following the EDTA inactivation and after the Mn^{2+} treatment.

Dilution ‘on assay’ gave more EDTA inhibition but also less Mn^{2+} activation. EDTA treatment was less effective when dilution ‘on incubation’ was used but consequently, the Mn^{2+} activation was higher. Effectively, both dilution methods proved to give comparable results as the ratios of activation/inactivation¹ were similar: 116 for dilution ‘on incubation’ and 109 for dilution ‘on assay’. Similarly, the combinations of EDTA and Mn concentrations that were checked, 1 mM EDTA/2 mM Mn^{2+2} and 2 mM EDTA/2 mM Mn^{2+} , gave very similar results: ratios of 115 and 119, respectively.

Taking everything into account, the dilution ‘on assay’ as well as the combination 2 mM EDTA/2 mM Mn^{2+} was chosen for a routine activity measurements. The dilution ‘on assay’ was better to help standardize the assay conditions. 2-fold dilution with EDTA and 5-fold dilution with Mn^{2+} could be used regardless of the concentration of the protein sample. However, to be able to activate the protein always at the same protein concentration (see section 4.2.1/Protein Concentration-Dependent Mode of Activation, p. 141) an additional dilution preceding the EDTA treatment (when required) was introduced.

¹activity values following the activation and inactivation were compared

²EDTA treatment followed by Mn^{2+} incubation was adopted for a standard assay routine except cases when standard buffer contained 1 mM EDTA

Mn²⁺ Concentration Dependence of Enzyme Activation

The pyrophosphatase activity of *B. subtilis* untruncated enzyme as well as the 5AATR mutant was measured as a function of an increasing concentration of Mn²⁺ ions (concentration of MnCl₂) used for the enzyme activation prior to activity assay. Protein concentration of both enzymes used for the activation step was 2 μg/ml (57 nM). The metal concentration range was 0.05 to 16 mM.

Both enzymes were sensitive to Mn²⁺ and low concentrations of MnCl₂ were sufficient to fully activate them (Fig. 4.1, p. 139). The maximum activity value was reached at about 1 – 2 mM MnCl₂, with 50% activation at about 0.3 mM MnCl₂. However, excess of Mn²⁺ ions inhibited the mutant protein's activity (Fig. 4.1 b, p. 139). This is discussed in section 4.3.3.

Enzymes were also assayed at lower protein concentrations (used for activation with manganese), namely 0.25 μg/ml (7 nM) for full-length PPase and 0.42 μg/ml (12 nM) for 5AATR. In case of 5AATR the effect of different manganese concentrations on enzyme activity was very similar to the one at 2 μg/ml, with 1 – 2 mM MnCl₂ needed for a maximum activation (Fig. 4.2 b, p. 140). Again, at higher MnCl₂ concentrations the inactivation of the mutant protein was observed (Fig. 4.2 b, p. 140) and this is discussed in section 4.3.3. As for PPase, the enzyme was not as sensitive to manganese. At 1 mM MnCl₂ it reached only 50% activation, while to get its maximum value over 6 mM MnCl₂ was necessary (Fig. 4.2 a, p. 140).

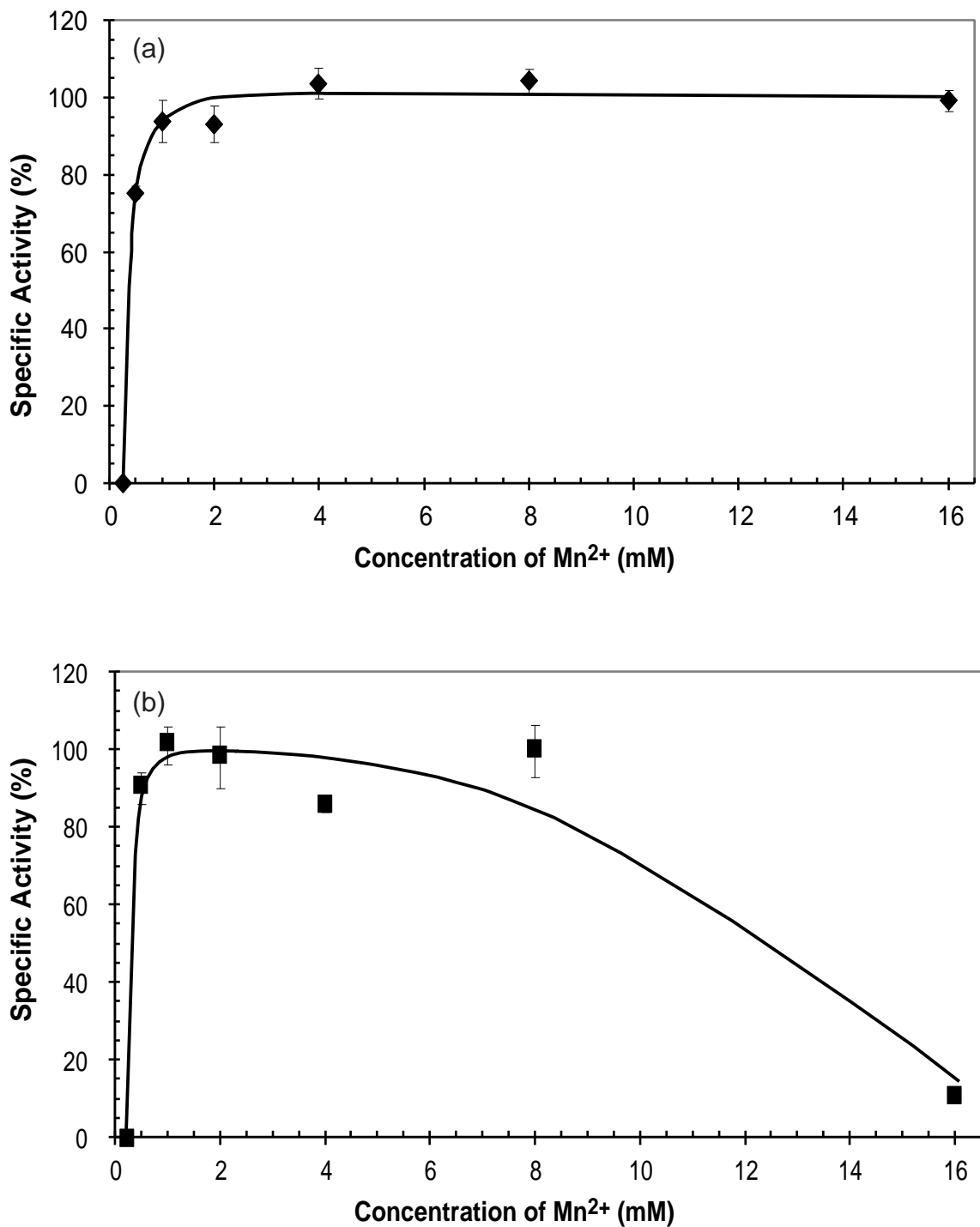


Figure 4.1: Effect of Mn^{2+} concentration on activation of *B. subtilis* PPase and 5AATR at 57 nM protein concentration: (a) the full-length *B. subtilis* PPase, (b) 5AATR. EDTA-treated enzymes (standard assay condition indicated in Materials and Methods, see section 2.21) were diluted into basic medium (25 mM BES buffer, pH 7.0) containing an appropriate concentration of $MnCl_2$ (0.25 – 16 mM), room temperature. Each result represents the mean \pm SD from one experiment performed in triplicate

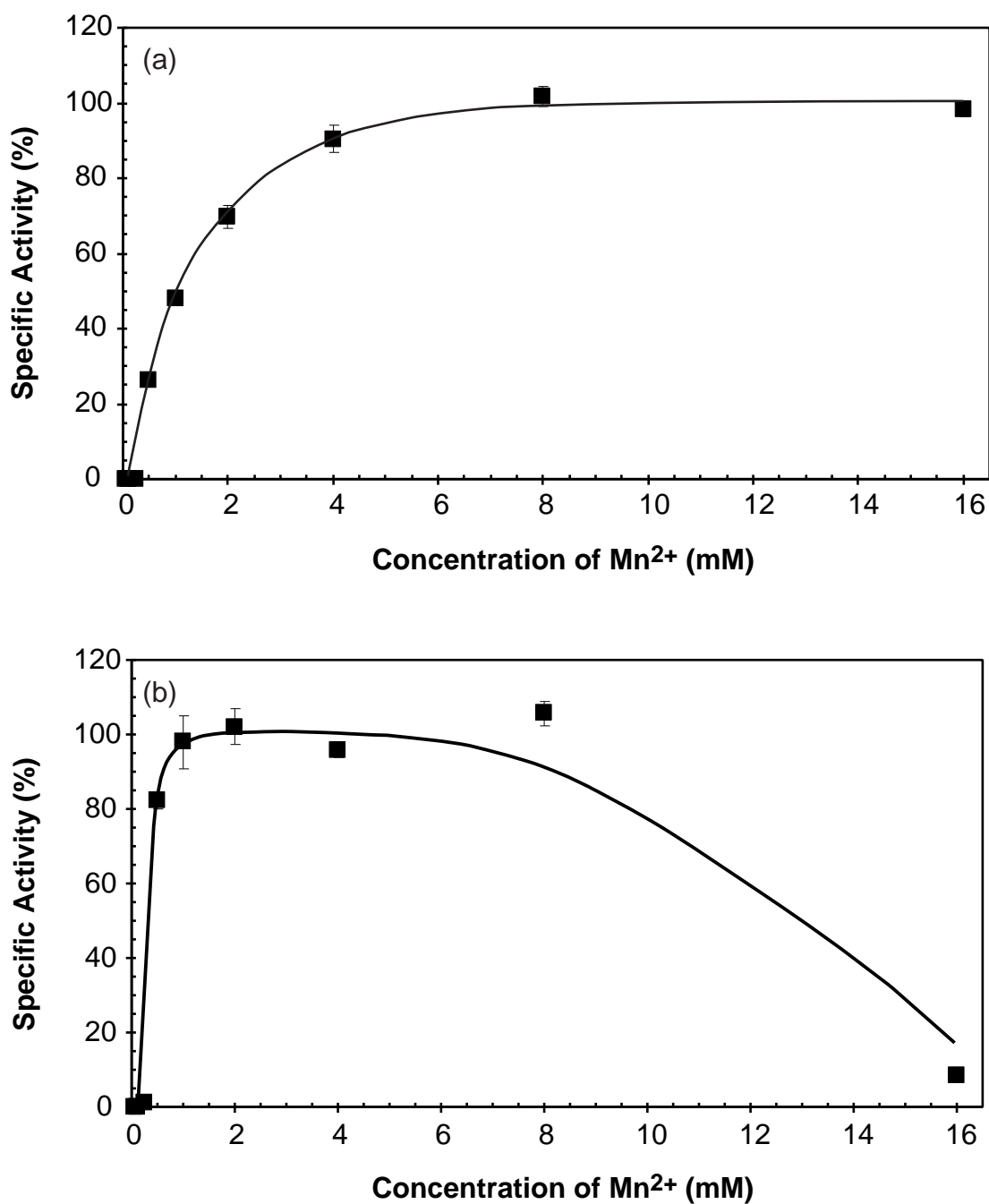


Figure 4.2: Effect of Mn^{2+} concentration on activation of *B. subtilis* PPase and 5AATR mutant at 7 nM and 12 nM protein concentration, respectively: (a) the full-length *B. subtilis* PPase, (b) 5AATR. EDTA-treated enzymes (standard assay condition indicated in Materials and Methods, see section 2.21) were diluted into basic medium (25 mM BES buffer, pH 7.0) containing an appropriate concentration of $MnCl_2$ (0.05 – 16 mM), room temperature. Each result represents the mean \pm SD from one experiment performed in triplicate

Protein Concentration-Dependent Mode of Activation

Two different protein concentrations, 0.25 $\mu\text{g}/\text{ml}$ (7 nM, Relative Concentration 1 or RC1) and 2 $\mu\text{g}/\text{ml}$ (57 nM, Relative Concentration 8 or RC 8), were used for the activation of enzymes with Mn^{2+} . Both untruncated and mutant protein were diluted appropriately to reach the above concentrations while incubated with 2 mM MnCl_2 .

Evidently, the specific activity of both enzymes was dependent on the protein concentration used for the activation step (Table 4.1, p. 142). It increased with increasing enzyme concentration in the incubation media. The ratio of activity RC8/RC1 was essentially of the same order for both untruncated and mutant PPase: 1.8 vs. 2.6, respectively. The slightly higher value for 5AATR might be explained by the general instability of its activity values (see section 4.3.2). 2 $\mu\text{g}/\text{ml}$ (57 nM) was the concentration used most often for the activation step in further studies.

Time-course of Activation

The process of activation of both untruncated and mutant PPase with Mn^{2+} was assayed as a function of time. The enzyme activity was monitored for three hours, starting from the very moment when 2 mM MnCl_2 was added. The protein concentration used for the incubation with MnCl_2 was 2 $\mu\text{g}/\text{ml}$ (57 nM).

Essentially, the rate of activation of both the proteins was very similar (Fig. 4.3, p. 143) with the half-times of 4 min. 90% of activity was obtained after 20 min while full activation – after 35 – 40 min of incubation with Mn^{2+} .

| Enzyme variant | Protein concentration* | | PPase activity | |
|----------------|------------------------|--|---------------------------|---------------------------|
| | (mg/ml) | Specific activity (μ mole Pi/mg protein) | Relative activity† (%) | Relative activity† (%) |
| PPase | 0.25 | 6019 \pm 201 | 100 | 100 |
| | 2 | 11 105 \pm 1723 | 100 | 100 |
| 5AATR | 0.25 | 533 \pm 274 | 8.9 \pm 4.6 | |
| | 2 | 1377 \pm 96 | 12.4 \pm 0.9 | |
| 17AATR | 5 | 1.5 \pm 0.35 | 0.01 \pm 0.003 | |

Table 4.1: The activity of the full-length *B. subtilis* PPase and truncated mutants. Inorganic pyrophosphatase activity of purified enzymes was assayed as described by Kuhn and Ward (1998) following treatment with 2 mM EDTA (30 min, 37 ° C) and activation with 2 mM MnCl₂ (60 min, room temperature). Each result represents the mean \pm SD from two separate experiments, with each experiment performed in triplicate; *, enzymes at two different concentrations, 0.25 and 2 μ g/ml, were used for activation with Mn²⁺ prior to the activity assay, †, specific activity of the full-length *B. subtilis* PPase (corresponding to an appropriate concentration) was taken as 100%

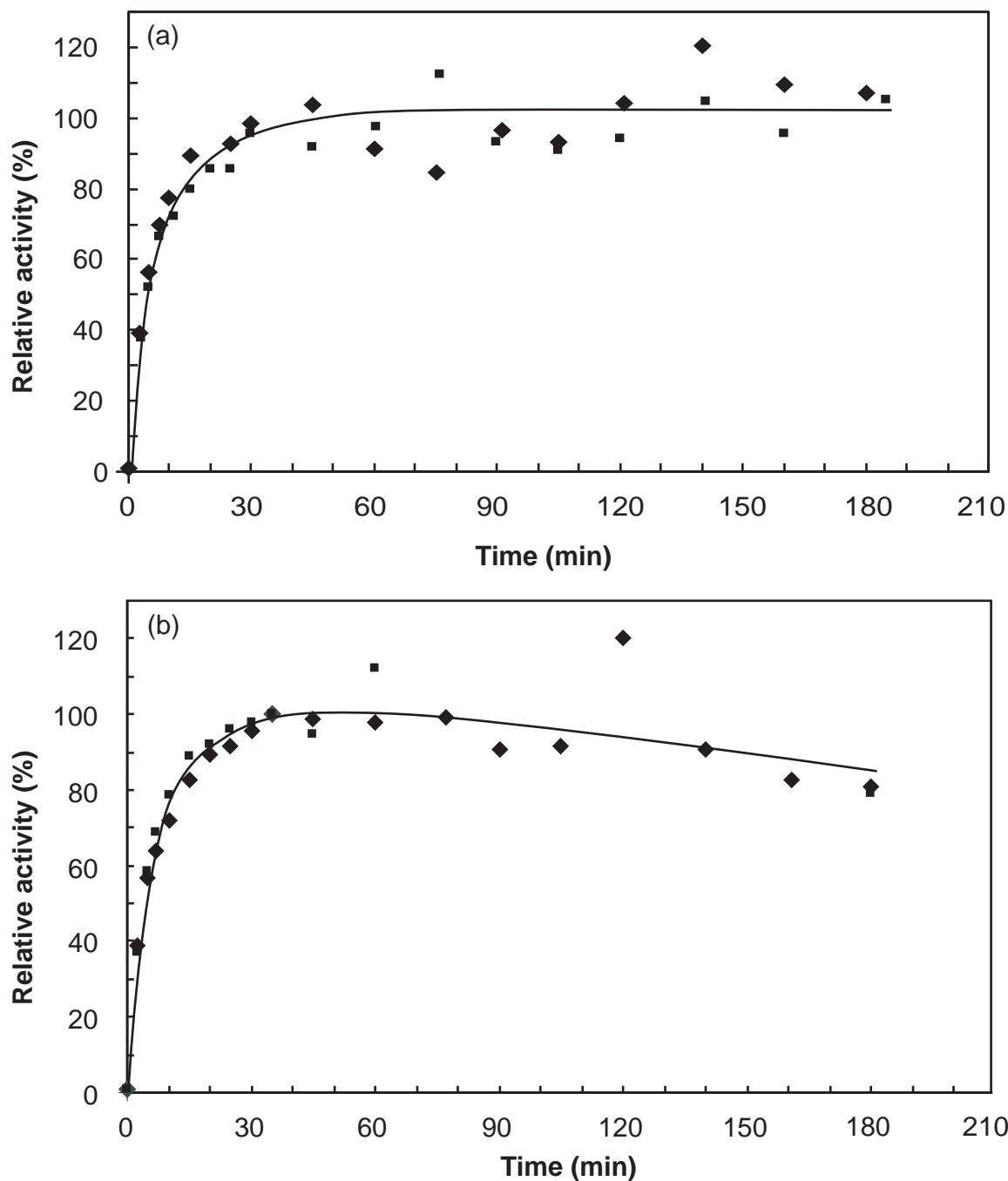


Figure 4.3: Time course of activation by Mn^{2+} : (a) full-length *B. subtilis* PPase, (b) 5AATR. EDTA-treated enzyme (standard assay conditions indicated in Materials and Methods, see section 2.21) was diluted into basic medium (25 mM BES buffer, pH 7.0) containing 2 mM $MnCl_2$, room temperature. The amount of enzyme used for the experiment was in the range of 0.006 – 0.1 μg for the untruncated PPase and 0.03 – 0.3 μg for 5AATR. (■) and (◆) are two independent assays

4.2.2. Insensitivity to Fluoride and Inhibition by EDTA

The activity assays of the untruncated PPase and 5AATR (standard assay conditions) were done following the preincubation with 2 mM EDTA as well as in the presence of 1 mM fluoride. The protein concentration used for the incubation with MnCl₂ (standard step in the activity assay, see section 3.3) was 2 µg/ml (57 nM).

Activity of *B. subtilis* PPase was unaffected by 1 mM fluoride and inhibited by EDTA. These properties were also shown by 5AATR (Table 4.2). When the protein concentration of 0.25 µg/ml (7 nM) was used for the Mn²⁺ activation step, the same effect of fluoride on both enzymes activity was detected, while activity following the EDTA treatment was not detectable under the assay conditions (not shown).

| Enzyme variant | Relative activity (%) [*] | | |
|-------------------------|------------------------------------|-------------|-----------|
| | none | 2 mM EDTA | 1 mM NaF |
| PPase | 100 | 0.3 ± 0.1 | 98 ± 14 |
| 5AATR | 100 | 1.4 ± 0.3 | 86 ± 10 |
| 17AATR† | 100 | 119.6 ± 9.3 | 8.1 ± 1.1 |
| <i>E. coli</i> control† | 100 | 105.7 ± 4.8 | 7.3 ± 0.1 |

Table 4.2: Effect of EDTA and fluoride on the full-length PPase, truncated mutants and *E. coli* control; *, activity of purified enzymes was assayed as described by Kuhn and Ward (1998), following standard assay routine indicated in Materials and Methods (see section 2.21). Enzymes were assayed without any additional treatment (none), after an additional exposure to 2 mM EDTA (25 mM BES buffer, pH 7.0, 30 min, 37 ° C) or in the presence of 1 mM NaF. Each result represents the mean ± SD from two separate experiments, with each experiment performed in triplicate. Relative activity of full-length *B. subtilis* PPase (1655 ± 257 µmol Pi/ml/min), 5AATR (73 ± 5), 17AATR (3.4 ± 0.1) or *E. coli* control (6.3 ± 0.04), where appropriate, was taken as 100%, †, activity of 17AATR and *E. coli* control was assayed in crude extracts

4.3. The 5AATR Mutant is Unstable

4.3.1. Protein Degradation

Proteins in whole-cell extracts after overexpression were analyzed by SDS-PAGE. Samples were regularly removed from the extract stocks kept on ice (or 4°C overnight) and frozen. When a sufficient number of the time points were collected, samples were loaded onto an SDS-polyacrylamide gel.

Analysis of 5AATR samples showed protein breakdown as a function of storage at 4°C. The pattern of bands changed gradually and the major one, clear at start, was successively accompanied by three newly appearing bands (Fig. 4.4 c and d, p. 146). Within the first few hours the process was slow. However, after about 20 hours a 4-band pattern was evident and during the additional 10 hours (1.5 day in total) the major protein band mostly disappeared. In contrast, the same experiment with the untruncated *Bacillus* PPase showed a stable protein up to 5 days (over 120 hours; Fig. 4.4 a and b, p. 146). The first traces of a breakdown appeared only after about 6.5 days (150 hours).

4.3.2. Loss of Activity with Time

Activity of 5AATR mutant was unstable. In the overexpressed material the mutant's specific activity was 5-fold lower than the one of the full-length enzyme. It decreased during purification at least by an additional 5% and in the purified material there was about a 7-fold difference between the mutant and the untruncated protein. However, it seemed that inclusion of 1 mM EDTA in the buffers used for the purification had a stabilizing effect

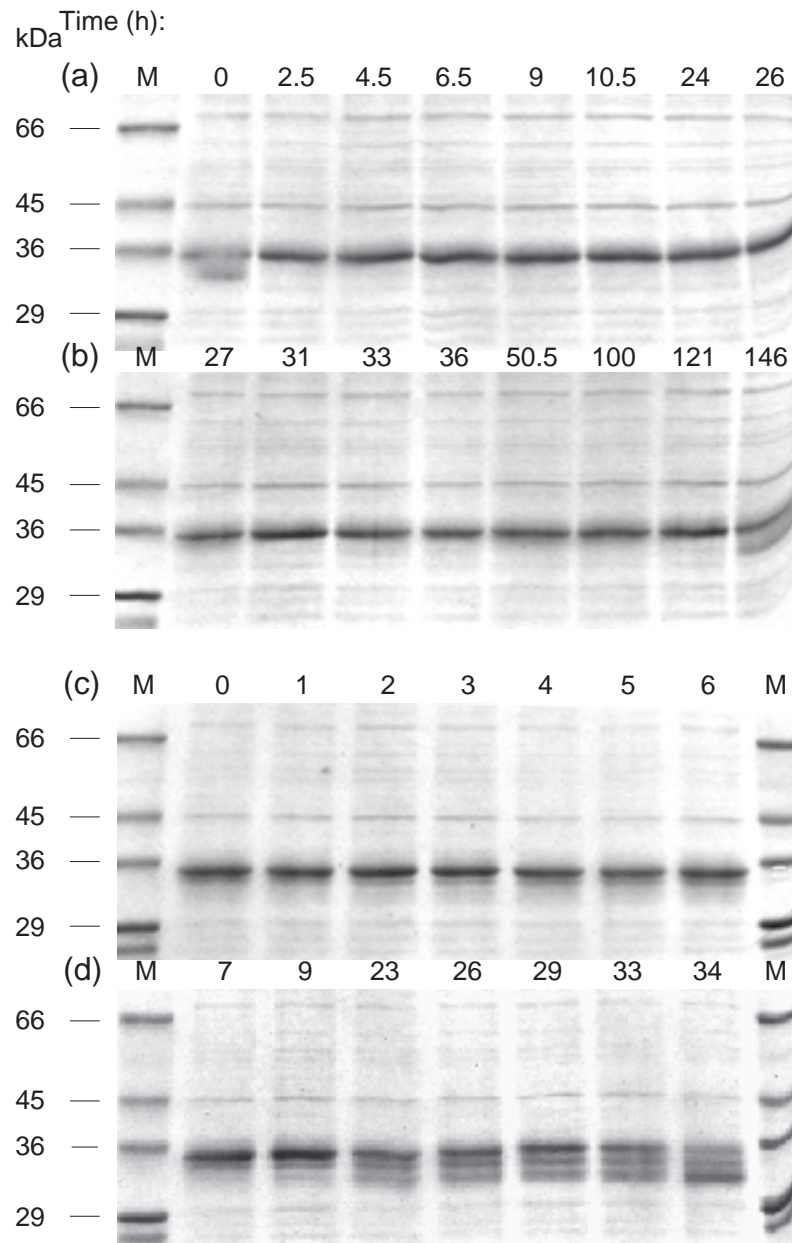


Figure 4.4: Effect of incubation at 4°C on the full-length *B. subtilis* PPase (a, b) and 5AATR (c, d). M – Sigma protein marker. Samples were regularly withdrawn from the whole-cell extract stocks of untruncated PPase and 5AATR kept on ice (or 4°C overnight) and diluted 6-fold for untruncated PPase or 4-fold for 5AATR. 20 μ l of each were mixed with an equal volume of 2 \times Laemmli sample buffer and frozen at -20°C. After a sufficient number of the time points has been collected, samples were heated at 95°C for 10 min and 10 μ l loaded per lane of the SDS-polyacrylamide gel. They were run using standard conditions (section 2.15.2)

and reduced the specific activity drop to 2% (overall difference in the purified material less than 6-fold).

A clear loss of activity with time could be observed during the analysis of the time-course of Mn^{2+} activation (Fig. 4.3 b, p. 143). In contrast to the full-length protein, the mutant activity started to decline immediately after reaching its maximum. Within the next two hours up to 20% of the activity was lost.

Finally, the activity of both enzymes untreated with manganese³ and stored at 4°C was monitored over several days. Interestingly, within the first 10 minutes activity of 5AATR was about the same as of the full-length *Bacillus* PPase. Both activities declined with time, however the mutant one at the much higher rate (Fig. 4.5, p. 148). After about 9 h the truncated protein lost 50% of its activity and nearly all of its activity (98%) within the next 100 hours (4 days). In contrast, the activity of the full-length enzyme was much more stable, reached the value of some 55% – 60% after about 35 hours of incubation and then remained constant.

4.3.3. Loss of Activity with Increasing Concentration of MnCl_2

When the pyrophosphatase activity of *B. subtilis* full-length enzyme and 5AATR were monitored upon addition of increasing amounts of MnCl_2 (section 4.2.1/ Mn^{2+} Concentration Dependence of Enzyme Activation, p. 138), excess of Mn^{2+} ions inhibited the mutant protein's activity. The mutant protein remained fully active up to 3 – 4 mM MnCl_2 concentration and only about 15% of activity was left when it was treated with 16 mM MnCl_2 (Fig. 4.1 b, p. 139 and 4.2 b, p. 140). This was unlike the untruncated enzyme in which the

³not the standard conditions used for an activity assay; according to Kuhn and Ward (1998) the enzyme is fully active only when treated with Mn^{2+}

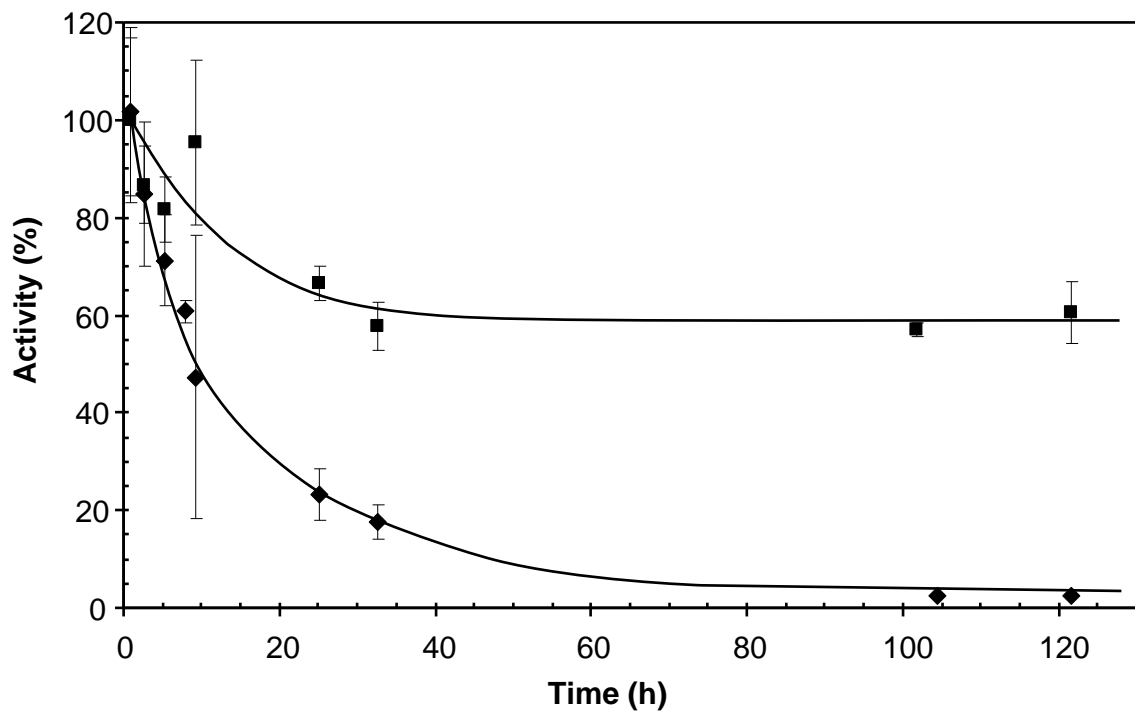


Figure 4.5: Effect of incubation at 4°C on the proteins untreated with manganese: (■) the full-length *B. subtilis* PPase, (◆) 5AATR. Not activated (not pretreated with manganese) proteins were diluted appropriately with basic medium (25 mM BES buffer, pH 7.0) and assayed for PPase activity. Each result represents the mean from one (the full-length PPase) or two (5AATR) separate experiments, with each experiment performed at least in triplicate

activity seemed not to be affected by increasing concentrations of MnCl_2 and retained its maximum value over the whole range of the manganese concentrations tested (Fig. 4.1 a, p. 139 and 4.2 a, p. 140).

4.3.4. Time-course of EDTA Inhibition

As with the activation by MnCl_2 (see section 4.2.1/Time-course of Activation, p. 141), the process of inactivation of both the full-length and mutant PPase with EDTA was assayed as a function of time. The enzyme activity was monitored for up to two hours, starting from the moment when the activated enzyme was diluted into basic medium containing 2 mM EDTA. The protein concentration used for the activation with manganese was 2 $\mu\text{g}/\text{ml}$ (57 nM).

Clearly, the rate of 5AATR inactivation was greater, approximately 2-fold (Fig. 4.6 b, p. 150) than the one for the full-length enzyme. The half-time of inactivation for PPase was 1 min and 0.5 min for the mutant. The full-length enzyme lost 85% of its activity after about 4 min while the same degree of inactivation 5AATR reached after only 1 min. 7 min were enough to get the mutant almost fully inactivated, while the full-length PPase needed 15 min for the 99% loss of its original activity.

4.4. Computer Predictions

The group of sixteen Family II representatives that have their complete genome sequences available (Fig. 5.1, p. 157) were taken for computer predictions of their protein secondary structures. The programme used was JPred² (Cuff et al. 1998). The analysis predicted

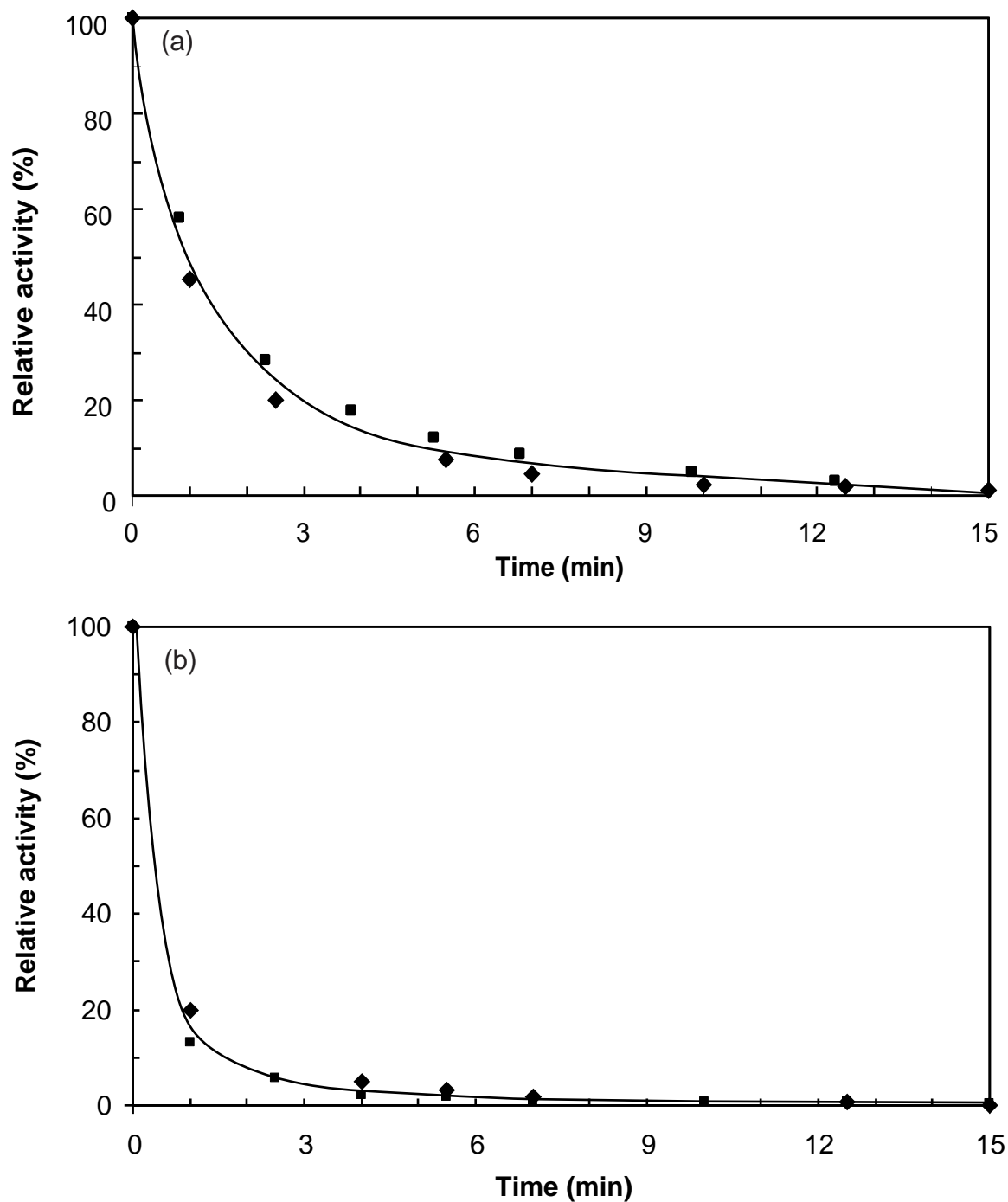


Figure 4.6: Time course of inactivation by EDTA: (a) full-length *B. subtilis* PPase, (b) 5AATR. EDTA and Mn^{2+} -treated enzyme (standard assay conditions indicated in Methods) was diluted into basic medium (25 mM BES buffer, pH 7.0) containing 2 mM EDTA, room temperature. The amount of enzyme used for the experiment was in the range of 0.006 – 0.05 g for the untruncated PPase and 0.03 – 0.05 g for 5AATR. (■) and (◆) are two independent assays

the presence of a six amino acid long α helix (V302-M307, *B. subtilis* numbering) at the C-terminus of each of the group members analysed (Fig. 4.7, p. 152).

4.5. Activity on Tri-Polyphosphate

Both the full-length protein and the 5AATR mutant were assayed for their activity on tri-polyphosphate. The activity was found to be inversely related to the dilution of the protein used for the activity assay. The lowest activity value was recorded for the enzyme samples that, following the manganese treatment, were not additionally diluted before the assay (dilution = 1). The value recorded for the untruncated PPase was about 4% of the activity against pyrophosphate while for the mutant it was almost three times higher, reaching 11% of the activity recorded on PP_i (Fig. 4.8 c and d, p. 153). The dilution range tested was from 1 (no dilution) to 16-fold. The activity of 5AATR increased about 3-fold within the dilution range used but, after reaching a maximum at about a 9-fold dilution, remained constant (Fig. 4.8 d, p. 153). The activity of the original PPase followed a similar path to that of 5AATR (Fig. 4.8 c, p. 153) but was always lower than the one recorded for the truncated mutant (e. g. 25% vs. 31% at 16-fold dilution).

Dilution following the Mn^{2+} activation step was a routine treatment for the assays of activity with pyrophosphate – otherwise readings were out of scale. However, dilution in these assays had an opposite effect to the one observed for the assays with tri-polyphosphate. Values of activity on PP_i were slowly but continuously decreasing up to about 10% activity loss for both the full-length and truncated protein (Fig. 4.8 a and b, p. 153).

| Organism | C-terminal sequence | No. of aa |
|---------------------------|----------------------------------|-----------|
| <i>B. subtilis</i> | KGVV SRKKQVVP VLTDAMAE-- | 305 |
| <i>A. fulgidus</i> | DGVM SRKKQVVP PLEKAF AEL- | 307 |
| <i>C. acetobutylicum</i> | PGIL SRKKQVIP PLTSAIEGK- | 304 |
| <i>D. radiodurans</i> | GNRI SRKKQIVP TLEKYFAPEA | 314 |
| <i>E. faecalis</i> | PGVV SRKKQVVP QLTEVFN--- | 302 |
| <i>Lis. innocua</i> | KGVV SRKKQVVP ILTDTFAK-- | 304 |
| <i>L. lactis</i> | AGAV SRKKQVVP QLTESFNK-- | 306 |
| <i>Lis. monocytogenes</i> | KGVV SRKKQVVP ILTDTFAK-- | 304 |
| <i>M. jannaschii</i> | EGVM SRKKQVVP PLERAYNG-- | 299 |
| <i>S. gordonii</i> | AGAV SRKKQVVP QLTESFNA-- | 304 |
| <i>S. mutans</i> | AGAV SRKKQVVP QLTESFNG-- | 304 |
| <i>S. pneumoniae</i> | AGAV SRKKQVVP QLTESFNA-- | 304 |
| <i>S. pyogenes</i> | AGAV SRKKQVVP QLTESFGA-- | 304 |
| <i>Staph. aureus</i> | SGVV SRKKQIVP QITEALTK-- | 304 |
| <i>T. maritima</i> | KGVM SRKKDFVP KIGEVLRRER | 297 |
| <i>V. cholerae</i> | ENTL SRKKQGW PWLQTELA--- | 301 |
| <hr/> | | |
| JPred ² | ---EE-----HHHHH---- | |

Figure 4.7: Computer predictions of the protein secondary structure at the C-termini (22 residues) of Family II PPases. Sequences shown include: *Bacillus subtilis* and *Archaeoglobus fulgidus* (Young et al. 1998), *Clostridium acetobutylicum* (Nolling et al. 2001), *Deinococcus radiodurans* (White et al. 1999), *Enterococcus faecalis* (TIGR 2002), *Listeria innocua* (Glaser et al. 2001), *Lactococcus lactis* (Bolotin et al. 2001), *Listeria monocytogenes* (Glaser et al. 2001), *Methanococcus jannaschii* and *Streptococcus gordonii* (Young et al. 1998), *Streptococcus mutans* (Shintani et al. 1998), *Streptococcus pneumoniae* (Tettelin et al. 2001), *Streptococcus pyogenes* (Ferretti et al. 2001), *Staphylococcus aureus* (Kuroda et al. 2001b), *Thermotoga maritima* (Nelson et al. 1999), and *Vibrio cholerae* (Heidelberg et al. 2000). The analysis was done using JPred² (Cuff et al. 1998). Letters *E* denote residues forming a predicted β -sheet while *H*, those forming a helix. The boxed region represents the postulated fingerprint (Shintani et al. 1998) or signature sequence with the invariant 'SRKK' motif in bold

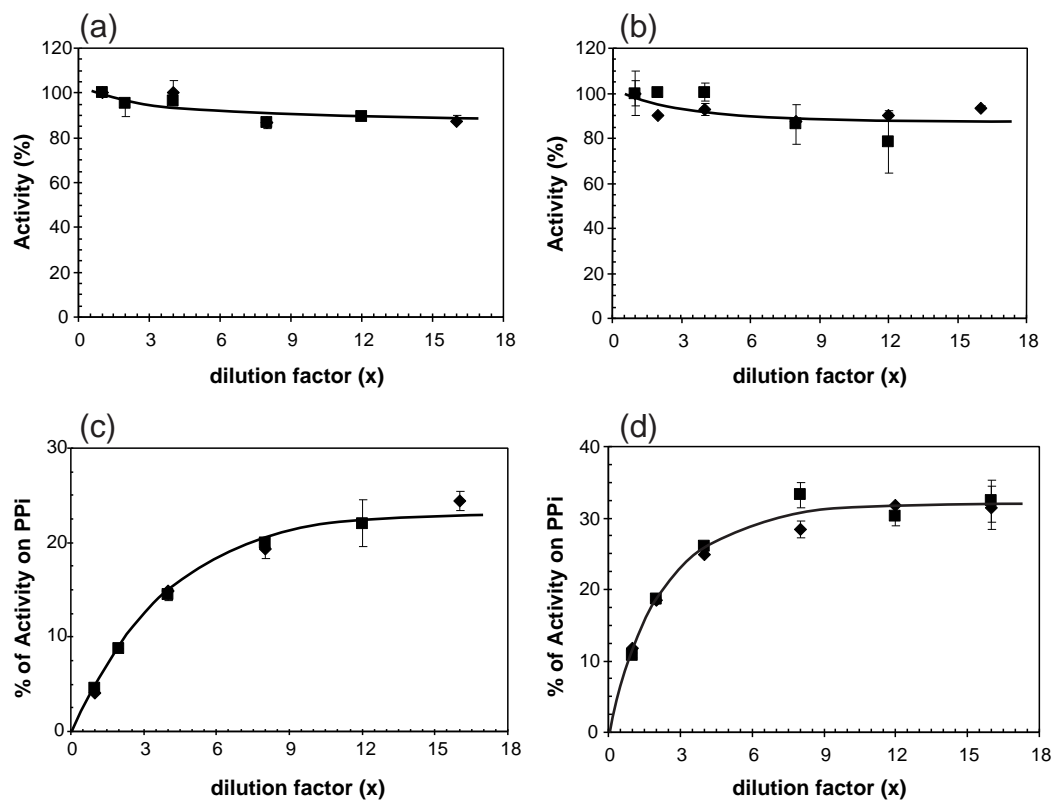


Figure 4.8: Effect of enzyme dilution on activity against tri-polyphosphate: (c) the full-length *B. subtilis* PPase and (d) 5AATR, compared with the activity on pyrophosphate: (a) untruncated *B. subtilis* PPase and (b) 5AATR. (■) and (◆) are two independent assays performed at least in triplicate (the full-length PPase) or in duplicate (5AATR)

4.6. Conclusions

Detailed studies on the 5AATR protein gave important insights into the nature of the observed significant loss of the mutant's activity.

5AATR retains all the main characteristics of the original *Bacillus subtilis* enzyme (Young et al. 1998, Kuhn and Ward 1998), such as the protein concentration-dependent mode of Mn^{2+} activation, its rate (Fig. 4.3, p. 143), insensitivity to fluoride and inhibition by EDTA (Table 4.2, p. 144). The protein concentration-dependent mode of activation and the rate of activation are very close to those reported by Kuhn and Ward (1998). Nonetheless, the mutant is very unstable. It is susceptible to proteolytic cleavage (Fig. 4.4, p. 146), its activity decreases during the purification process, is at least twice as sensitive to inhibition by EDTA as the activity of the full-length protein (Fig. 4.6, p. 150) and disappears, almost completely, when the protein is not supplied with Mn^{2+} ions (Fig. 4.5, p. 148). Evidently, C-terminal truncation of 5 aa appears to result in a thermodynamically unfavoured and weakened conformation of the enzyme. Also, the sensitivity of the enzyme to manganese changes at lower protein concentrations. The response of 5AATR to different manganese concentrations is similar at 12 nM and 57 nM protein, unlike that of the full-length protein, which at 7 nM needs much higher amounts of Mn^{2+} to get the maximum activation (compare Fig. 4.1 a, p. 139 with 4.2 a, p. 4.2, and Fig. 4.1 b, p. 139 with 4.2 b, p. 140). Clearly, the five C-terminal residues (D305-E309; Fig. 3.2, p. 117) are absolutely essential for the overall stability of the *B. subtilis* PPase.

Computer analysis of predicted secondary structures using the programme JPred² (Cuff et al. 1998) show a six residues long C-terminal α helix (V302-M307, *B. subtilis* numbering)

for all putative family II members examined so far (Fig. 5.1, p. 157). Three out of the five terminal residues are an intrinsic part of this helix and must be required to maintain this element of secondary structure. Deletion would most likely make the protein partially unfold, thus affecting the positioning of the helix and probably that of the adjacent 'SRKK' motif.

Finally, the increased catalytic activity of 5AATR towards tri-polyphosphate when compared to the activity of the original PPase, most likely indicates conformational changes the mutant undergoes that enable accommodation of larger substrates in its active site. The increase in activity with dilution of the enzyme may suggest a dilution-induced 'opening' of the structure, which allows the catalysis of the larger substrate. Whatever the mechanism, it is clear that in consequence of dilution-induced structural changes the catalysis of tri-polyphosphate is more favoured (Fig. 4.8 c and d, p. 153).

The study of the protein concentration dependence of activity (see section 4.2.1/ Protein Concentration-Dependent Mode of Activation, p. 141) revealed how important it is to work at known protein concentrations, especially when comparing different enzyme properties or working with the mutant variants.

Chapter 5

The Signature Sequence Preserves the Functionality of the Active Site

5.1. Introduction

Sequence alignment of the new Family II of inorganic pyrophosphatases reveals a number of extended regions of similarity. However, one of the most prominent sequence conservations in all the representatives is found near the C-terminus. Interestingly, despite the increasing number of the Family II members being identified, the number of conserved amino acids in this C-terminal fingerprint (Shintani et al. 1998) patch remains the same. It is eight residues long (294 – SRKKQVVP – 301; *Bacillus subtilis* numbering) in eleven of the group representatives identified so far (Fig. 5.1, p. 157).

Within this signature however, the ‘SRKK’ motif is invariant in all sixteen family members identified (Fig. 5.1, p. 157). Studies on the C-terminally truncated *Bacillus subtilis* enzyme revealed the importance of this C-terminal region for enzyme functioning. Further analysis of the unstable 5AATR suggested that enzyme function could be affected due to structural changes induced by truncation of as few as 5 terminal amino acids.

To learn more about the signature and the role of the individual amino acids within

| Organism | Acc. number | C-terminal sequence | No.of aa |
|--------------------------|-------------|--------------------------------|----------|
| <i>B.subtilis</i> † | NC_000964 | V SRKKQ VVP VLTDAMAE-- | 309 |
| <i>A.fulgidus</i> † | 029502 | M SRKKQ VVP PLEKAF AEL- | 322 |
| <i>C.acetobutylicum</i> | NC_003030 | L SRKKQ VIP PLTSAIEGK- | 310 |
| <i>D.radiodurans</i> | NC_001263 | I SRKKQ IVP TLEKYFAPEA | 314 |
| <i>E.faecalis</i> | - | V SRKKQ VVP QLTEVFN--- | 308 |
| <i>Lis.innocua</i> | AL596168 | V SRKKQ VVP ILTDTFAK-- | 308 |
| <i>L.lactis</i> | NC_002662 | V SRKKQ VVP QLTESFNK-- | 314 |
| <i>Lis.monocytogenes</i> | AL591979 | V SRKKQ VVP ILTDTFAK-- | 308 |
| <i>M.jannaschii</i> † | Q58025 | M SRKKQ VVP PLERAYNG-- | 307 |
| <i>S.gordonii</i> † | P95765 | V SRKKQ VVP QLTESFNA-- | 311 |
| <i>S.mutans</i> † | 068579 | V SRKKQ VVP QLTESFNG-- | 310 |
| <i>S.pneumoniae</i> | NC_003028 | V SRKKQ VVP QLTESFNA-- | 311 |
| <i>S.pyogenes</i> | NC_002737 | V SRKKQ VVP QLTESFGA-- | 311 |
| <i>Staph.aureus</i> | NC_002745 | V SRKKQ IVP QITEALTK-- | 309 |
| <i>T.maritima</i> | NC_000853 | M SRKKD FVP KIGEVLRER | 548 |
| <i>V.cholerae</i> | NC_002505 | L SRKKQ GWP WLQTELA--- | 308 |
| | | :****: :* : | |

Figure 5.1: Sequence alignment of the C-termini of Family II PPases (19 residues). Sequences shown include: *Bacillus subtilis* and *Archaeoglobus fulgidus* (Young et al. 1998), *Clostridium acetobutylicum* (Nolling et al. 2001), *Deinococcus radiodurans* (White et al. 1999), *Enterococcus faecalis* (TIGR 2002), *Listeria innocua* (Glaser et al. 2001), *Lactococcus lactis* (Bolotin et al. 2001), *Listeria monocytogenes* (Glaser et al. 2001), *Methanococcus jannaschii* and *Streptococcus gordonii* (Young et al. 1998), *Streptococcus mutans* (Shintani et al. 1998), *Streptococcus pneumoniae* (Tettelin et al. 2001), *Streptococcus pyogenes* (Ferretti et al. 2001), *Staphylococcus aureus* (Kuroda et al. 2001b), *Thermotoga maritima* (Nelson et al. 1999), and *Vibrio cholerae* (Heidelberg et al. 2000). The total number of amino acids in the sequence is given on the right. Symbol † denotes the very first family members, identified at the time I started the project. Sequences were aligned using ClustalW (Higgins et al. 1996). Residues fully conserved in all sequences are marked with *asterisks*. *Colons* denote residues of groups conserved including conservative replacements. The *boxed* region represents the postulated fingerprint (Shintani et al. 1998) or signature sequence (exceptions in conservation indicated with *italics*). The invariant ‘SRKK’ motif is *in bold*

this region for enzyme activity and structure, site-directed mutagenesis was employed and a series of site specific mutants produced. Three out of the eight conserved residues, namely Ser294, Arg295 and Lys296, were individually changed into an alanine, while Ser294 was additionally conservatively replaced with threonine as well as cysteine and valine. A double mutant Arg295Ala/Lys296Ala was also prepared. Val299 and Val300 were substituted with isoleucine (conservative replacement) but also a tryptophan mutant of Val300 as well as a double Val299Gly/Val300Trp were made. All the mutants were constructed and overexpressed in *E. coli*.

Arginine and lysine are the amino acids commonly found in the active sites of phosphate-related enzymes, being responsible for substrate binding (Kim et al. 1990, Ayabe et al. 1997). Both of them are also present in the active site of Family I PPases (Lahti et al. 1990). In our search for the active site these two residues, especially as part of the signature, were the most obvious targets for mutation. Additionally, preliminary X-ray crystallographic studies on the PPase from *Streptococcus gordonii* (White, personal communication) specifically suggested these two residues to interact with phosphate in the putative enzyme active site. Serine, on the other hand, is well known as a donor of a hydroxide ion in the catalytic mechanism of phospho-hydrolyses eg., alkaline phosphatases (Stec et al. 2000). However, no role for serine was found in the mechanism of activity of Family I enzymes (Heikinheimo et al. 1996a). During the final year of this study, X-ray crystallographic studies on *Streptococcus mutans* (Merckel et al. 2001), *Streptococcus gordonii* and *Bacillus subtilis* (Ahn et al. 2001) produced three-dimensional structures of these Family II members. According to these structures serine has no direct role in the catalytic mechanism of Family II PPases.

Investigation of the role of Ser294 as an invariant residue was of particular importance with regard to the catalysis of PP_i hydrolysis by Family II proteins.

5.2. Generation of Site-Specific Mutants

5.2.1. PCR and Transformation into Epicurian Coli Cells

Two single and one double alanine mutant of Arg295 and Lys296 (Arg295Ala, Lys296Ala, Arg295Ala/Lys296Ala) as well as four variants of Ser294 (Ser294Ala, Ser294Thr, Ser294Cys, Ser294Val) and three of Val299 and Val300 (Val299Ile, Val300Ile, Val300Trp, Val299Gly/Val300Trp) were constructed. All were produced from the *Bacillus subtilis* PPase coding sequence (gene *yybQ*) using the QuickChange Site-Directed Mutagenesis Kit (Stratagene). The procedure involved synthesis of the mutant sequences (6573 bp) by PCR (Fig. 5.2, p. 167) using high fidelity, nonstrand-displacing *Pfu Turbo* DNA polymerase, digestion of the methylated, non-mutated parental DNA template with *DpnI* and transformation of the mutated dsDNA into Epicurian Coli XL1 – Blue supercompetent cells. The oligonucleotide primers used for PCR are indicated in Materials and Methods (section 2.11.1/Site-Specific Mutants, p. 88). The yield of transformation into Epicurian Coli cells varied from 2 (Val299Gly/Val300Trp, 3.5 μ l of *DpnI* treated material used) to 300 colonies (Ser294Ala, 1 μ l of *DpnI* treated material used; only half of the normal amount of supercompetent cells used). Following the transformation into Epicurian Coli cells the transformant DNA was sequenced in one direction (see an example in Fig. 5.4, p. 169).

5.2.2. Re-cloning

Only after the desired mutations were confirmed the mutant inserts were re-cloned into a vector DNA of a confirmed sequence. A sequenced PPase construct (*B. subtilis* PPase clone, 6573 bp – vector DNA) as well as the transformant DNA (6573 bp) isolated from Epicurian cells (variants of the specifically mutated PPase clone) were treated, mostly separately, with *SacI* and *BamHI* (see Fig. 2.2, p. 93). Following digestion, DNA fragments, namely 339 bp variant mutant inserts and 6234 bp vector fragments (Fig. 5.3, p. 168) were separated and purified from an agarose gel by QIAquick Gel Extraction Kit (Qiagen). The vector DNA was additionally digested with alkaline phosphatase (section 2.7.2). Mutant fragments were ligated with the vector overnight at room temperature at the vector/insert ratios 1:10 except for Lys296Ala where 1:5 was also used. As before when generating truncated mutants (section 3.2.1) transformation into *E. coli* DH5 α and BL21(DE3) followed. Mutant DNA isolated from the DH5 α clones was sequenced in both directions (example shown in Fig. 5.4, p. 169).

5.2.3. Overexpression and Protein Isolation

Overexpression and protein extraction from the *E. coli* cultures was done essentially as before (section 3.2.2). The protein expression was induced with IPTG for 2.5 hours at 30°C. After 2.5-hour induction single major bands corresponding to a monomer of about 34 kDa were identified by SDS-PAGE in the crude extracts of *E. coli* BL21 (Fig. 5.5, p. 170). Sonication was in five 90-s bursts with 1 min rests in between. The highest yields of the protein were obtained only when the cells were washed before being frozen at -20°C, which proceeded sonication.

5.2.4. Relative Subunit Molecular Mass

The relative molecular masses of the specific mutants were determined in crude extracts as before using TOF MS ES. The M_r values coming from single major peaks (see an example in Fig. 3.6, p. 121) agreed, within experimental error, with those calculated from the sequence (Table 5.1).

| Enzyme variant | Relative Subunit Molecular Mass (M_r) | |
|---------------------|---|--------------|
| | Calculated* | Experimental |
| Ser294Ala | 33,972 | 33,978 |
| Ser294Cys | 34,004 | 34,010 |
| Ser294Thr | 34,002 | 34,004 |
| Ser294Val | 34,000 | 34,003 |
| Arg295Ala | 33,903 | 33,903 |
| Lys296Ala | 33,931 | 33,932 |
| Val299Ile | 34,002 | 34,008 |
| Val300Ile | 34,002 | 34,008 |
| Val300Trp | 34,075 | 34,080 |
| Val299Gly/Val300Trp | 34,033 | 34,034 |

Table 5.1: Molecular mass of *B. subtilis* PPase mutants measured by TOF MS ES; *, subunit molecular mass calculated from the amino acid sequence

5.3. Arg295 and Lys296. Direct Involvement in Catalysis

Alanine substitutions of both Arg295 (Arg295Ala) and Lys296 (Lys296Ala) resulted in proteins showing no detectable *B. subtilis* PPase activity: the low amounts of activity measured in crude extracts were of *E. coli* origin (not sensitive to EDTA and inhibited by fluoride, Table 5.2). The phenomenon of fluoride insensitivity appears unique to family

II PPases (Kuhn et al. 2000). As single alanine mutants were not active, the double one, Arg295Ala/Lys296Ala, was not considered for more detailed studies.

| Enzyme variant | Specific activity (%) | | |
|------------------------|-----------------------|--------------|------------|
| | None | 2 mM EDTA | 1 mM NaF |
| PPase | 100 | 6.4 ± 0.6 | 91.7 ± 2.8 |
| Arg295Ala | 100 | 107.7 ± 14.9 | 6.5 ± 1.8 |
| Lys296Ala | 100 | 112.4 ± 26.7 | 7.9 ± 0.2 |
| <i>E. coli</i> control | 100 | 105.7 ± 4.8 | 7.3 ± 0.1 |

Table 5.2: Effect of EDTA and fluoride on Arg295Ala and Lys296Ala *B. subtilis* PPase mutants. Activity was assayed in crude extracts as described by Kuhn and Ward (1998), following standard assay routine indicated in Materials and Methods (see section 2.21). Enzymes were assayed without any additional treatment (none), after an additional exposure to 2 mM EDTA (25 mM BES buffer, pH 7.0, 30 min, 37 ° C) and in the presence of 1 mM NaF. Each result represents the mean ± SD from at least two separate experiments, with each experiment performed at least in duplicate. Specific activity of *B. subtilis* PPase ($4125 \pm 28.6 \mu\text{mol Pi/ml/min}$), Arg295Ala (3.0 ± 0.2), Lys296Ala (3.1 ± 0.1) and *E. coli* control (4.7 ± 0.1), where appropriate, was taken as 100%

5.4. Ser294, Val299 and Val 300. Preservation of the Correct Folding Pattern

5.4.1. Ser294

Alanine substitution of Ser294 (Ser294Ala) resulted in 56% of the original enzyme activity, while conservative replacement Ser294Thr gave 27%. Variants in which Ser294 was replaced with a bulkier residue, Ser294Cys and Ser294Val, were almost completely inactive (0.4 and 0.33% of the original activity, respectively; Fig. 5.6, p. 171).

Like the original *Bacillus* PPase, the activity of Ser294 variants was inhibited by EDTA and insensitive to fluoride (Table 5.3). The values for the alanine (Ser294Ala) and threonine (Ser294Thr) mutants were comparable to the ones obtained for the original PPase. However,

when the other two mutants, namely Ser294Cys and Ser294Val were tested, both EDTA inhibition and fluoride insensitivity were less pronounced than those of the original enzyme. The variants, especially Ser294Val, seemed less sensitive to EDTA and more sensitive to fluoride, at the same time (Table 5.3).

| Enzyme variant | Specific activity (%) | | |
|---------------------|-----------------------|-------------------|-------------------|
| | None | 2 mM EDTA | 1 mM NaF |
| PPase | 100 | 6.4 ± 0.6 | 91.7 ± 2.8 |
| Ser294Ala | 100 | 9.7 ± 0.6 | 90.5 ± 5.1 |
| Ser294Cys | 100 | 15.7 ± 3.8 | 73.0 ± 2.7 |
| Ser294Thr | 100 | 1.5 ± 0.5 | 83.2 ± 2.3 |
| Ser294Val | 100 | 28.0 ± 5.1 | 63.7 ± 1.7 |
| Val299Ile | 100 | 7.3 ± 0.3 | 93.1 ± 0.7 |
| Val300Ile | 100 | 1.5 ± 0.5 | 83.7 ± 0.1 |
| Val300Trp | 100 | 3.2 ± 1.3 | 85.4 ± 2.3 |
| Val299Gly/Val300Trp | 100 | 13.6 ± 0.6 | 84.2 ± 3.8 |

Table 5.3: Effect of EDTA and fluoride on site-specific *B. subtilis* PPase mutants. Activity of the full-length *B. subtilis* enzyme as well as the site-specific mutants was assayed in crude extracts as described by Kuhn and Ward (1998), following assay routine indicated in Materials and Methods (see section 2.21). Enzymes were assayed without any additional treatment (none), after an additional exposure to 2 mM EDTA (25 mM BES buffer, pH 7.0, 30 min, 37°C) and in the presence of 1 mM NaF. Each result represents the mean ± SD from at least two separate experiments, with each experiment performed in triplicate. In *bold* are the values that do not follow the expected pattern

5.4.2. Val299 and Val300

Conservative substitution of Val299 (Val299Ile) had a negligible effect on the enzyme activity (Fig. 5.6, p. 171). However, similar substitution at the position of Val300 (Val300Ile) resulted in a 19% loss of activity. Replacing Val300 with tryptophan showed even a greater effect and Val300Trp retained only 23% of the original enzyme activity. Finally, the double

mutation substituting both valines at the same time, Val299Gly/Val300Trp, left only 0.5% of activity of the original PPase (Fig. 5.6, p. 171).

Activity of the variants was inhibited by EDTA as well as insensitive to fluoride (Table 5.3). The effects were comparable to those obtained for the original PPase. Only Val299Gly/Val300Trp seemed slightly less sensitive to EDTA.

5.5. Conclusions

Site-directed mutagenesis studies on the C-terminus of *Bacillus subtilis* inorganic pyrophosphatase provided some interesting insights into the significance of the signature sequence for the new PPase family.

Arg295 and Lys296 located within the strictly conserved ‘SRKK’ motif were the primary targets for mutations. Apart from the common presence of arginine and/or lysine at the phosphate binding sites (Kim et al. 1990, Ayabe et al. 1997), preliminary X-ray crystallographic studies on the enzyme from *S. gordonii* (White, personal communication) suggested these two residues interact with phosphate in the putative enzyme active site. Using site-directed mutagenesis these two residues were replaced with alanine to get two variants, Arg295Ala and Lys296Ala. Both mutations produced proteins lacking detectable *B. subtilis* PPase activity. Clearly, both residues are crucial for activity, which supports their predicted role in interacting with the negatively charged phosphates of the substrate.

Ser294, another invariant member of the ‘SRKK’ motif, could be a good candidate to participate in the catalytic mechanism via phosphorylation of the enzyme. Yet, analysis of the structure (Ahn et al. 2001) does not envisage a role for this residue in the catalytic mechanism. In these studies, the alanine variant of Ser294 (Ser294Ala) retains more than

50% of the original enzyme activity. Consequently, the hydroxyl group is not required for the PPase to remain active and any mechanism of catalysis involving a phosphorylated enzyme intermediate can be definitely ruled out. Conservative mutation to threonine (Ser294Thr), but introducing a one carbon atom extension, leads to a further inactivation of the enzyme. Substitution with the bulkier hydrophobic group of valine (Ser294Val) almost completely inactivates the protein. A similar effect is given by the introduction of a thiol group (Ser294Cys), in spite of the fact that the volume occupied by the new residue is similar to that of threonine (where significant activity remained). Most likely, steric effects displace the adjacent and catalytically essential Arg295 and Lys296 from their optimal positions. As the residue volumes of serine and alanine are comparable, there must be an additional factor making serine the preferred amino acid in this position. It could be an involvement in local hydrogen bonding. The thiol group effect might be slightly more complex. There is residual activity of Ser294Cys and Ser294Val variants when exposed to fluoride (Table 5.3, p. 163; compare effect on *E. coli* PPase, Table 5.2, p. 162). Therefore, Family II activity is present.

The effect of EDTA and fluoride on Ser294Cys and Ser294Val mutants is less pronounced than on the original *B. subtilis* enzyme (Table 5.3, p. 163) and might be of importance in implicating the active site residues direct involvement in the interactions with fluoride compounds and, possibly, also their role in the EDTA inhibition. However, more detailed studies with purified proteins would be necessary to draw any final conclusions.

The 20% decrease in PPase activity caused by the conserved Val300Ile substitution and even more dramatic reductions found for Val300Trp (77%) and Val299Gly/ Val300Trp (99.5%) mutants support the previous observations that there are strict geometric require-

ments for activity. Introduction of a small flexible glycine would additionally destabilize the local fold, possibly also by a dramatic change to the hydrophobicity of the local environment. An apparent higher resistance of Val299Gly/ Val300Trp to EDTA when compared to the original *B. subtilis* enzyme might be of significance in the search for the mechanism of EDTA inhibition, but more studies are necessary before coming to final conclusions.

Taken together, the results show that residues in the invariant C-terminal region are needed to preserve the functionality of the active site both in terms of the residues directly involved in catalysis and the correct folding pattern of this part of the protein.

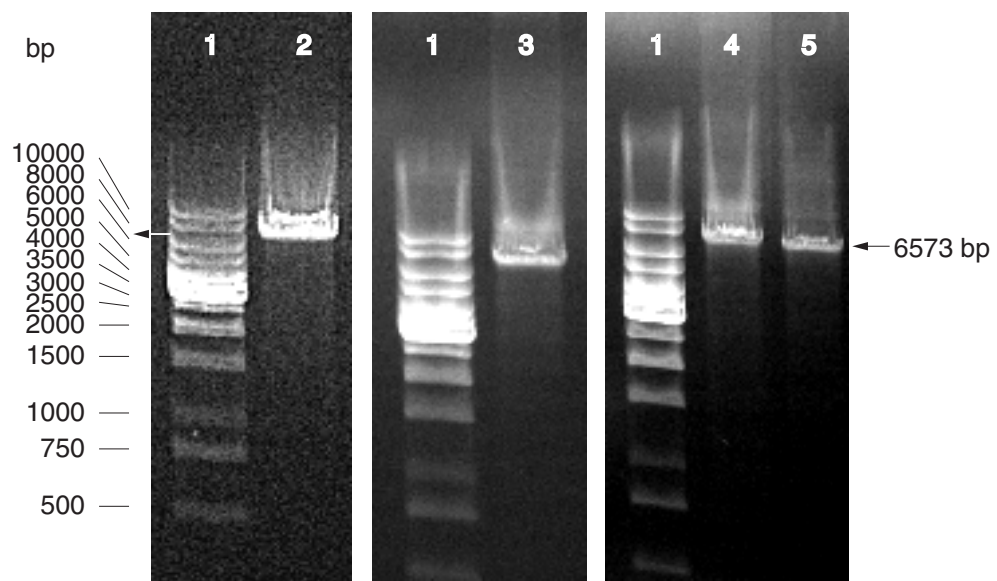


Figure 5.2: PCR of site specific *B. subtilis* PPase Mutants: 1 – Fermentas 1kb DNA ladder, 2 – Lys296Ala, 3 – Arg295Ala, 4 – Arg295Ala/Lys296Ala, 5 – Val299Gly/Val300Trp. Arrows indicate the size of bands produced

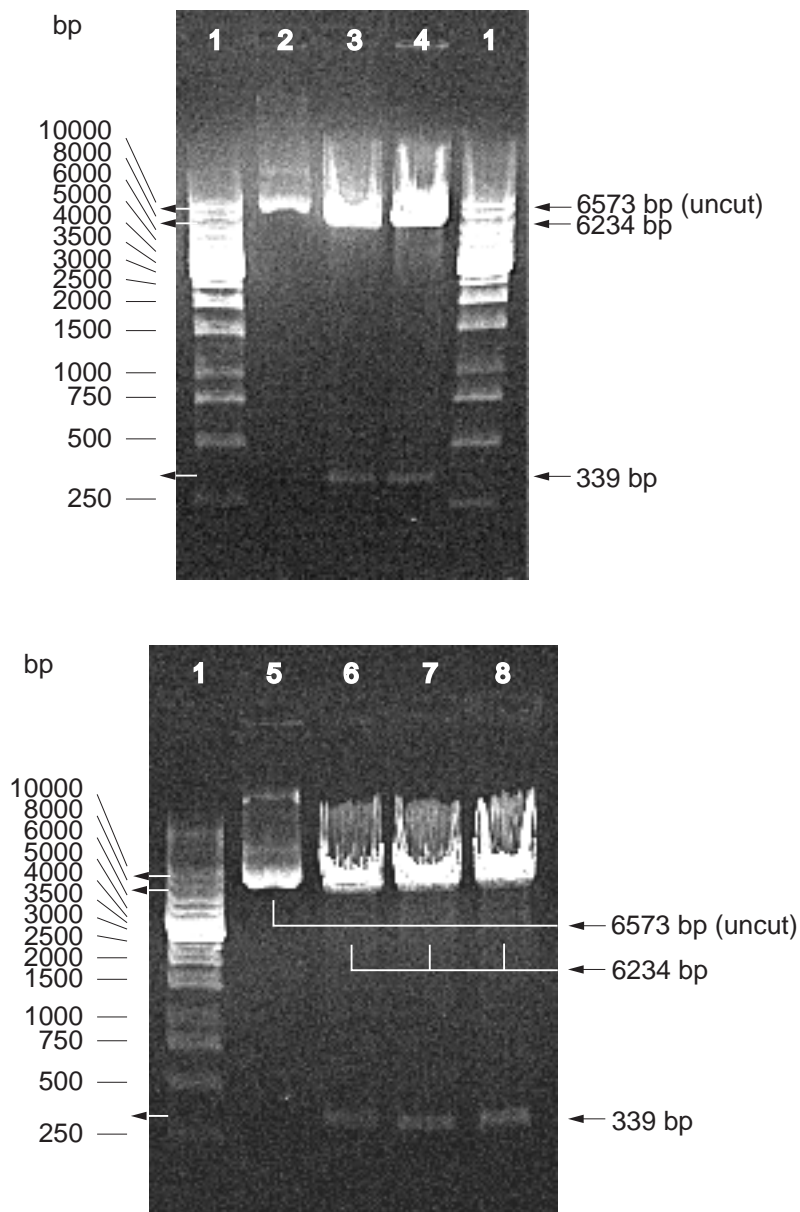


Figure 5.3: Re-cloning of site-specific mutant fragments: digestion with *SacI* and *BamHI* of a vector (sequenced *B. subtilis* PPase clone) and mutant PCR products: 1 – Fermentas 1kb DNA ladder, 2 – uncut Lys296Ala plasmid DNA, 3 – digested Lys296Ala plasmid DNA, 4 – digested vector DNA, 5 – uncut Arg295Ala/Lys296Ala plasmid DNA, 6 – digested Arg295Ala/Lys296Ala plasmid DNA, 7 – digested Val299Gly/Val300Trp plasmid DNA, 8 – digested Arg295Ala plasmid DNA. Arrows indicate the size of plasmid DNA and/or its fragments

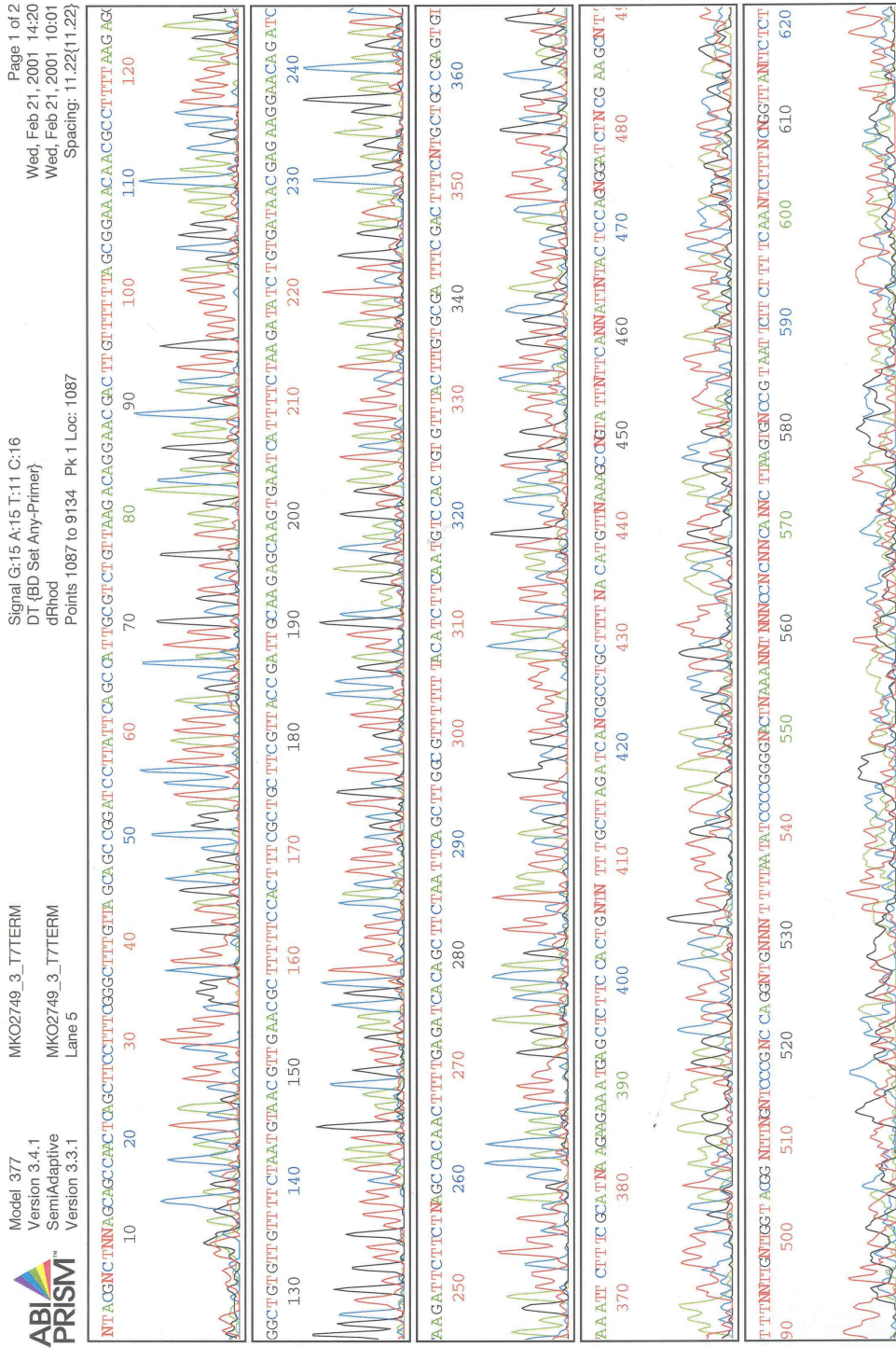


Figure 5.4: Sequence data obtained from an Arg295Ala clone using T7 terminator primer. Sequencing was done by Lark Technologies, Inc. The reaction was run on an ABI PRISM 377 DNA Sequencer and analysed on 5.75% and 4.75% Long Ranger Gels for ABI 377

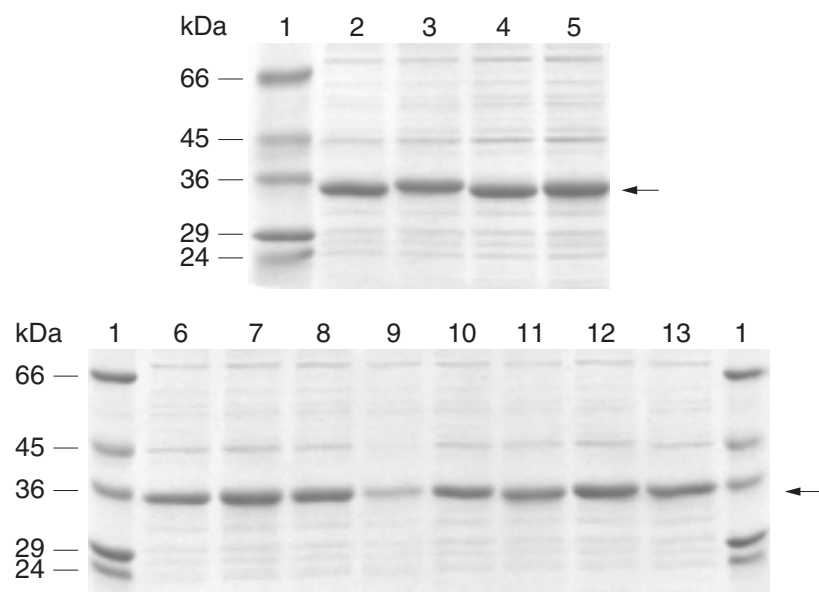
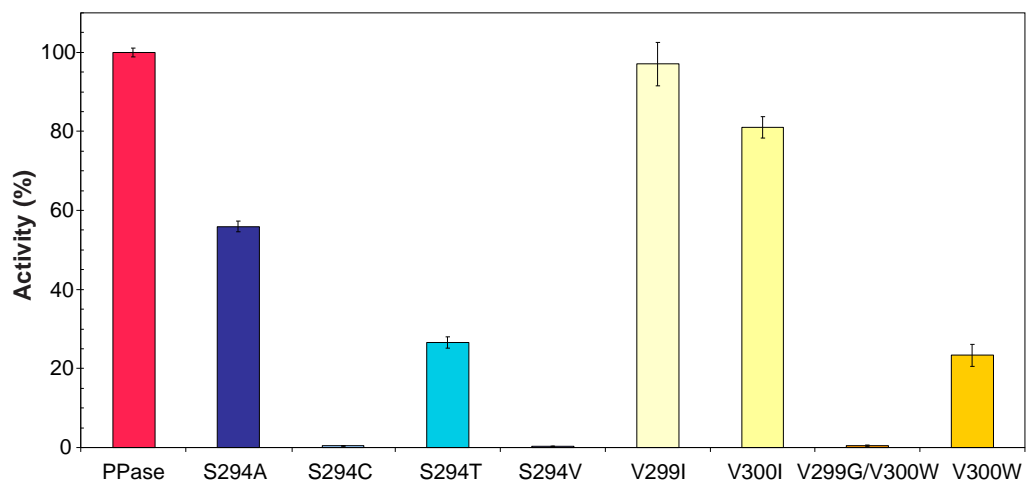


Figure 5.5: Overexpression of site-specific *B. subtilis* PPase mutants: 1 – Sigma protein marker, 2 – Lys296Ala, 3 – Arg295Ala, 4 – Arg295Ala/Lys296Ala, 5 – Val299Gly/Val300Trp, 6 – Ser294Ala, 7 – Ser294Cys, 8 – Ser294Thr, 9 – Ser294Val, 10 – PPase, 11 – Val300Trp, 12 – Val299Ile, 13 – Val300Ile. Arrows indicate overexpressed proteins. Proteins were stained with Coomassie Brilliant Blue



| Enzyme variant | Specific activity (%) |
|----------------|-----------------------|
| PPase | 100 |
| S294A | 55.95 ± 1.29 |
| S294C | 0.40 ± 0.05 |
| S294T | 26.55 ± 1.46 |
| S294V | 0.33 ± 0.01 |
| V299I | 97.04 ± 5.51 |
| V300I | 81.05 ± 2.69 |
| V300W | 23.35 ± 2.83 |
| V299G/V300W | 0.50 ± 0.14 |

Figure 5.6: Activity of site-specific *B. subtilis* PPase mutants. Activity of the full-length *B. subtilis* enzyme as well as the site-specific mutants was assayed in crude extracts as described by Kuhn and Ward (1998), following the assay routine indicated in Materials and Methods (see section 2.21). Each result represents the mean \pm SD from at least two separate experiments, with each experiment performed in triplicate

Chapter 6

Discussion

When in 1998 inorganic pyrophosphatase purified from *Bacillus subtilis* (Kuhn and Ward 1998) appeared different from all the other pyrophosphatases studied so far and representing a new Family II of these enzymes (Young et al. 1998, Shintani et al. 1998), a primary goal became to identify parts of this unique molecule crucial for enzyme activity. It seemed very likely that Family II enzymes shared a distinct active site together with a distinct mechanism of action. The sequence alignment, the only information about the new family available at the time, showed an eight-residue long sequence at the enzyme C-terminus as one of the most characteristic conservations among all the members. Studies on the *B. subtilis* PPase C-terminal region including this proposed fingerprint (Shintani et al. 1998) or signature, were the first attempt to learn more about this new family of proteins.

The loss of activity observed for the *B. subtilis* PPase mutant lacking 17 C-terminal residues (V293-E309; Fig. 3.2, p. 117) validated what sequence alignments implied and the three-dimensional structures (Ahn et al. 2001, Merckel et al. 2001) now confirm: the C-terminus with its signature sequence is absolutely essential for enzyme function. Inter-

estingly however, removal of the five non-conserved amino acids (D305-E309; Fig. 3.2, p. 117) from the C-terminus is sufficient to significantly impair the enzyme activity with at least a five-fold decrease in specific activity. 5AATR is very unstable, e.g. susceptible to proteolytic cleavage, losing activity on purification, more sensitive to EDTA inhibition and completely inactivated when not supplied with manganese ions, but it retains the most characteristic features of the full-length enzyme. These include its concentration-dependent mode of activation, the activation rate, insensitivity to fluoride and inhibition by EDTA. It is clear therefore that the five C-terminal residues are absolutely vital for the overall stability of the *B. subtilis* PPase.

Data from gel-filtration experiments are consistent, implying that the full-length and both truncated proteins are homotetramers. However, under physiological conditions (Mn^{2+} present) *B. subtilis* PPase is a dimer (Parfenyev et al. 2001) and it crystallises as a dimer (Ahn et al. 2001). We assume that gel-filtration produces anomalous results because of a non-spherical nature of the protein in solution compared to the reference standards used. In line with this hypothesis is an ‘open’, flexible conformation (Fig. 6.1, p. 175) proposed by Ahn et al. (2001) and the 20% increase in the axial ratio (measured by analytical ultracentrifugation) observed for the *S. gordonii* PPase in the presence of EDTA (M. Ilias, unpublished data). Analytical ultracentrifugation as well as gel-filtration results suggested a similar effect (increase in the molecule axial ratio when activating Co^{2+} ions are removed from the medium) for *Archaeoglobus fulgidus* enzyme (J. Osborne, unpublished data). Nevertheless, the fact that both 17AATR and 5AATR behave in the same manner as the full-length protein is consistent with the C-terminus having no major influence on the tertiary structure of the proteins.

Computer analysis of predicted secondary structures using the programme JPred² (Cuff et al. 1998) shows a six residues long C-terminal α helix (V302 – M307, *B. subtilis* numbering) for all putative Family II members examined so far (Fig. 5.1, p. 157). Structural studies now show that 14 amino acids at the C-terminii of *S. gordonii* and *B. subtilis* monomers (Ahn et al. 2001) form an α helix (helix K, R295 – A308, *B. subtilis* sequence; see Fig. 6.1, p. 175 and Fig. 6.2, p. 176). In *S. mutans* (Merckel et al. 2001) the helix is shorter (helix L, V300 – S306, *B. subtilis* numbering). Overall, at least two out of the five terminal residues (D305 – E309; Fig. 3.2, p. 117) are an intrinsic part of this helix and must be required to participate in this element of secondary structure and facilitate its interaction with other regions of the protein. Deletion clearly would affect the optimal positioning of the helix K and probably that of the adjacent ‘SRKK’ motif. It would affect the subunit conformation as well, as the helix K, part of the C-terminal domain, is found to be directly involved in making inter-domain contacts (Fig. 6.1, p. 175, also Fig. 6.2, p. 176) and to pack against the concave surface of the N-terminal domain (Ahn et al. 2001).

The induced-fit mechanism proposed by Ahn et al. (2001) pictures the active site, formed at the domain interface (Fig. 6.1, p. 175, 6.2, p. 176 and 6.3, p. 177), closing around pyrophosphate as it binds. Truncation of the 5 C-terminal amino acids, by affecting the inter-domain interactions, could possibly influence the efficiency of substrate binding as well. The results presented in this thesis show that 5AATR has increased activity towards triphosphate which would be consistent with the adoption of a less closed conformation of the C-terminal domain (Fig. 6.1, p. 175) permitting the binding of a larger substrate molecule (Ahn et al. 2001, Merckel et al. 2001). The observed dilution-induced increase in activity

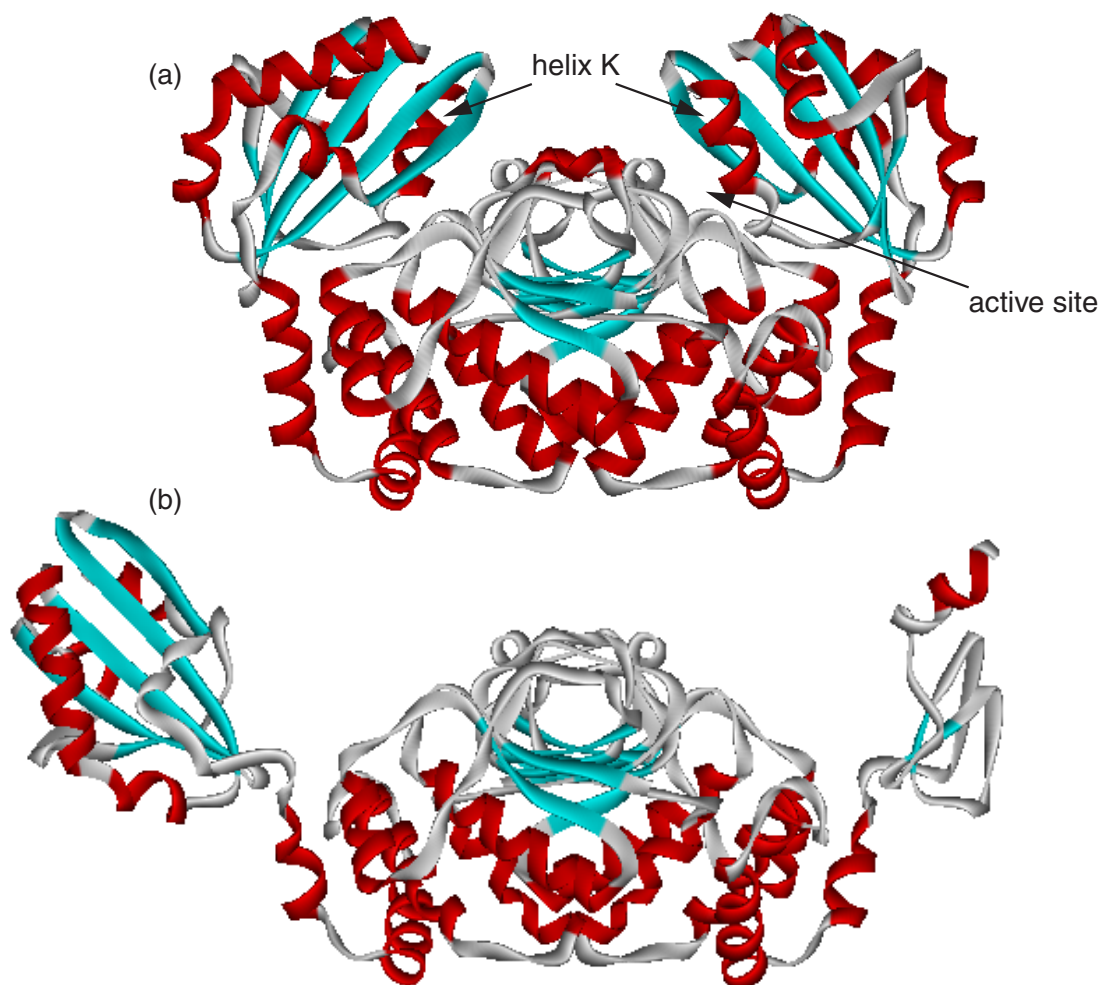


Figure 6.1: Overall fold of *S. gordonii* (a) and *B. subtilis* (b) PPase molecules. Helices are *red*, β -sheets are *cyan* and coils are *grey*. Helix K (Ahn et al. 2001) and the active site (at the domain interface) are marked with *arrows*. The dimer interface is hydrophobic (Ahn et al. 2001). In *B. subtilis* PPase dimer (Ahn et al. 2001) the C-terminal domain was found rotated away from the core structure by approximately $90 - 100^\circ$ ('open' conformation) compared to the *S. gordonii* PPase C-terminal domains ('closed' conformation). A weak electron density for the residues 191 – 307 (C-terminal domain, Ahn et al. 2001) of one of the *B. subtilis* PPase monomers (structural subunits) did not allow to trace the amino acid backbone confidently in this part of the protein. Thus, the backbone is truncated in the figure. Structural data of Ahn et al. (2001) was downloaded from OCA PDB (2002). Figure was generated using DS ViewerPro 5.0 (Accelrys Inc. 2002)

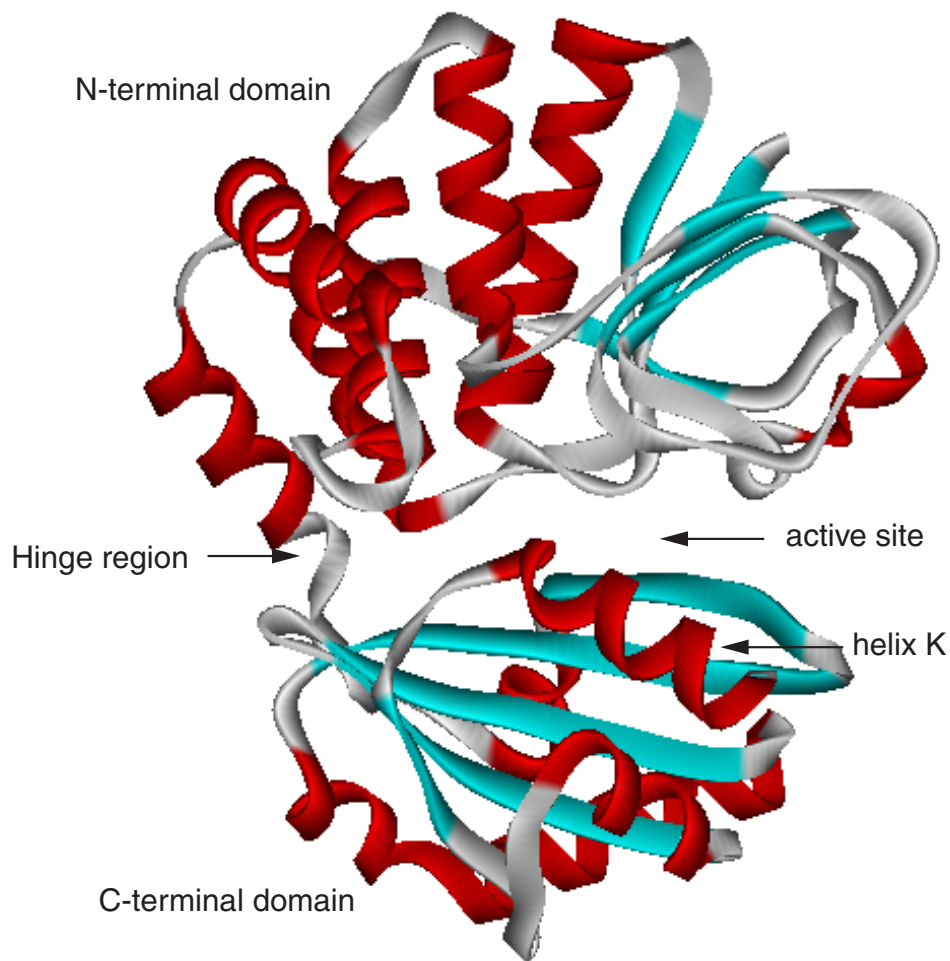


Figure 6.2: Overall fold of *S. gordonii* PPase monomer. Helices are *red*, β -sheets are *cyan* and coils are *grey*. It is proposed that Family II PPases fold into two domains joined by a flexible hinge at Gly190 (*S. gordonii* numbering, Ahn et al. 2001). The N-terminal domain (app. 190 aa) consists of a twisted six-stranded parallel β -sheet flanked by extended loops on one side and a helical-rich subdomain on the other, with first 130 aa reminiscent of the α/β , dinucleotide-binding or Rossmann fold topology, characteristic of many dehydrogenases (Ahn et al. 2001). The C-terminal domain (app. the last 120 aa) is a five-stranded mixed β -sheet with flanking α -helices on both sides, and its topology resembles that of an SH2 domain (phosphotyrosine binding Src homology 2), represented e.g. by human tyrosine phosphatase (Ahn et al. 2001). The hinge region enables large structural change between ‘open’ and ‘closed’ conformations (Ahn et al. 2001, Merckel et al. 2001). The active site is formed at the interface between the N- and C-terminal domains that are complementary in shape. The interface is extensive and polar in nature. The fold of the *B. subtilis* monomer was found to be almost identical, except the C-terminal domain that was found rotated away from the core structure by approximately 90 – 100° (Fig. 6.1, p. 175), compared to the *S. gordonii* PPase C-terminal domain position. Overall, the three-dimensional structures and the topological fold of a monomer of *S. gordonii* and *B. subtilis* enzymes are very different from those found for Family I PPases (see section 1.3.4 and Fig. 1.3, p. 42). Structural data of Ahn et al. (2001) was acquired from OCA PDB (2002). Figure was generated using DS ViewerPro 5.0 (Accelrys Inc. 2002)

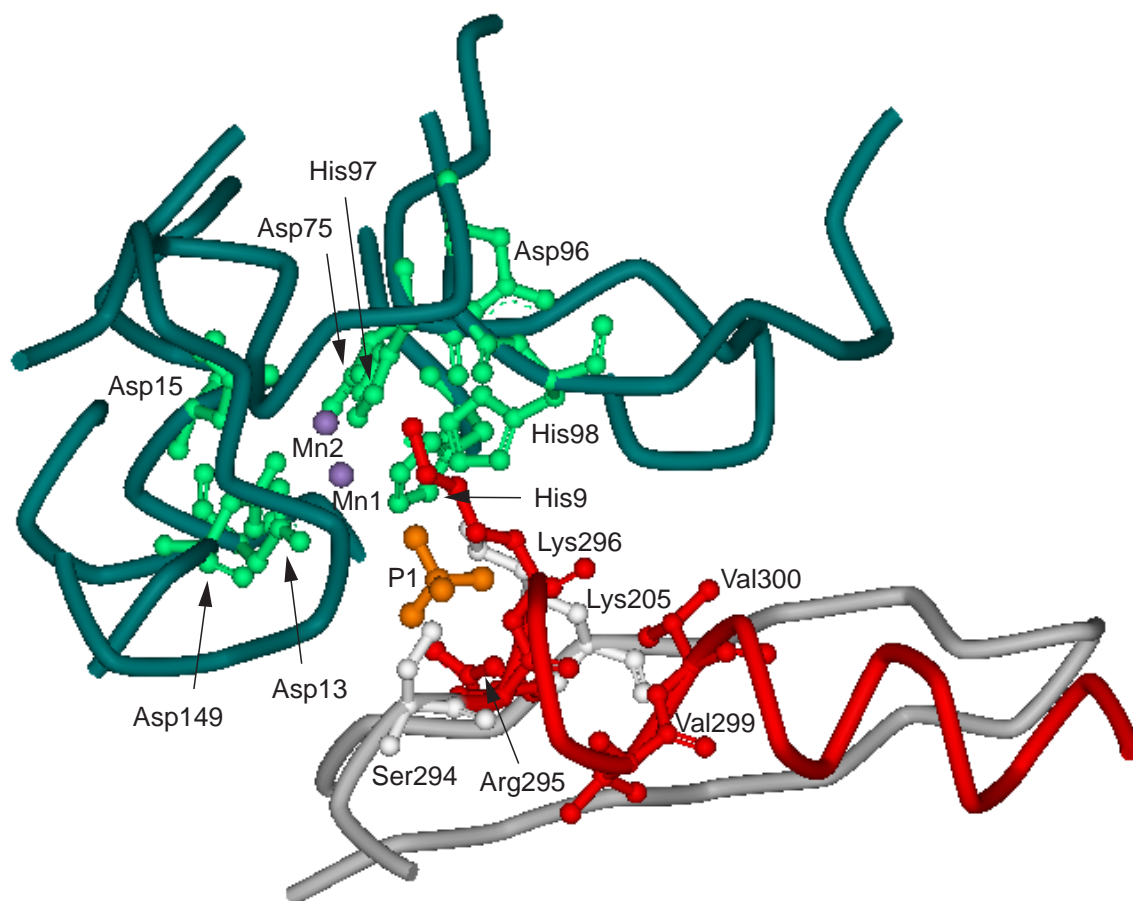


Figure 6.3: Detailed view of the mutated amino acids (Ser 294, Arg295, Lys296, Val299 and Val300) and other residues of the putative *B. subtilis* active site, proposed to be formed at the interface between the N- (*green*) and C-terminal (*grey*) domains. α -Helix K (Ahn et al. 2001), part of the C-terminal domain, is *red*. Metal binding sites, Mn1 and Mn2 (Mn bound in the structure) are marked, as well as phosphate group P1 (modelled to superimpose with the sulphate molecule bound to the active site in the crystal structure). Structural data of Ahn et al. (2001) was downloaded from OCA PDB (2002). Figure was generated using DS ViewerPro 5.0 (Accelrys Inc. 2002)

on tri-phosphate is almost twice as large for 5AATR as for the full-length enzyme: the original *B. subtilis* enzyme needed a 16-fold dilution to reach the maximum activity while for the mutant a 9-fold dilution was sufficient. Possibly, dilution promotes the ‘opening’ of the structure and the phenomenon, in line with the above hypothesis, is much more pronounced for the truncated mutant than the full-length protein. There are no reports so far about the dilution effect on the domain structure of the enzyme. Parfenyev et al. (2001) reported the effect of dilution (in the absence of Mn^{2+}) on the monomerisation of the Family II proteins. In fact, dimer stability was the lowest with *B. subtilis* PPase (Parfenyev et al. 2001), compared to the *S. gordonii* and *S. mutans* enzymes.

Arg295 and Lys296, located within the strictly conserved ‘SRKK’ motif, on the basis of the *S. gordonii* structure, are two of the three residues from the C-terminal domain (helix K, Fig. 6.3, p. 177 also Fig. 6.4, p. 179) that contribute to the putative active site (Ahn et al. 2001). The alanine variants of these two residues, Arg295Ala and Lys296Ala, both lacked detectable *B. subtilis* PPase activity. Clearly, both residues are crucial for the enzyme activity, which supports their predicted role in interacting with the negatively charged phosphates of the substrate.

Ser294 (Fig. 6.3, p. 177) is an invariant member of the ‘SRKK’ motif for which no direct role in catalysis has been proposed (Ahn et al. 2001, Merckel et al. 2001). We find that the alanine variant of Ser294 (Ser294Ala) retains more than 50% of the original enzyme activity. Accordingly, the hydroxyl group is not required for the PPase to remain active. As many phospho-hydrolases (eg. alkaline phosphatases) have a requirement for a hydroxyl group in the catalytic mechanism (Stec et al. 2000) and Ser294 is the only candidate amino acid with a hydroxyl group close to the putative active site (Ahn et al., 2001; Fig. 6.3,

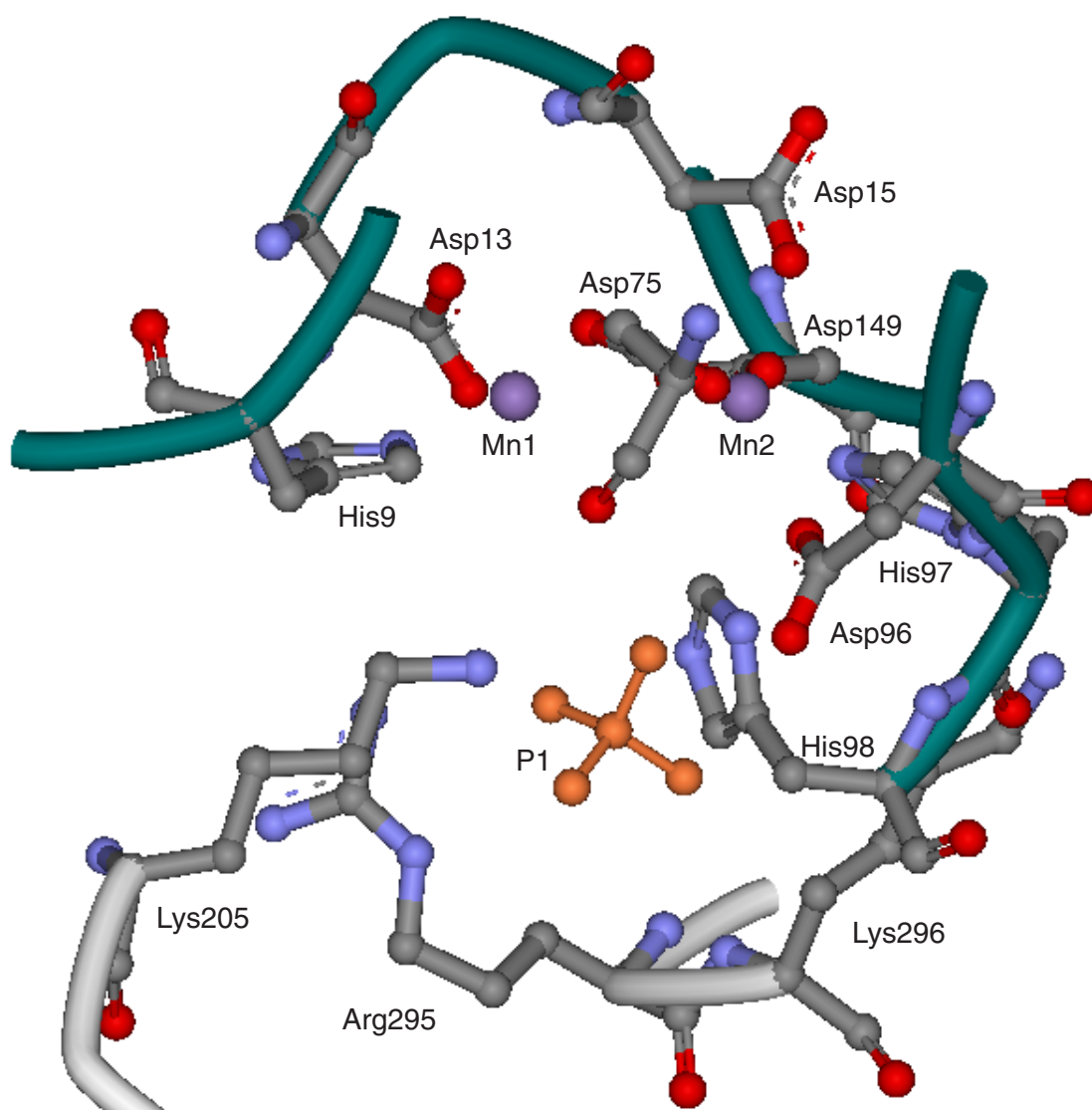


Figure 6.4: Structure of the *B. subtilis* PPase putative active site. The active site is formed at the domain interface and the polypeptide backbone of the N-terminal domain is *green*, while the one of the C-terminal domain – *white*. All the amino acid residues are invariant amongst the family homologues shown in Fig. 5.1, p. 157. Metal binding sites, Mn1 and Mn2 (Mn bound in the structure) are marked, as well as the phosphate binding site P1 (modelled to superimpose with the sulphate molecule bound to the active site in the crystal structure). Structural data of Ahn et al. (2001) was downloaded from OCA PDB (2002). Figure was generated using DS ViewerPro 5.0 (Accelrys Inc. 2002)

p. 177), any catalytic mechanism involving a phosphorylated Ser294 enzyme intermediate can be definitely ruled out. Mutation to threonine (Ser294Thr, an extra methyl group) leads to a further inactivation of the PPase while substitution with a bulkier group of valine (Ser294Val) almost completely inactivates the protein. Introducing a thiol group (Ser294Cys) gives a similar effect. The most likely interpretation is that steric effects displace the adjacent and catalytically essential Arg295 and Lys296 from their optimal position.

Evidently, the results presented imply the involvement of the C-terminal residues in the interactions with the substrate, possibly also their role in the substrate specificity, facilitated by the structural flexibility of the PPase molecule. The hypothesis of the structural flexibility is supported by the presence of the flexible hinge region in the structure of the *S. gordonii* PPase monomer (Fig. 6.2, p. 176) and, as discussed earlier, the ‘open’ conformation (Fig. 6.1, p. 175) proposed by Ahn et al. (2001). Further, the ‘SRKK’ motif, via Arg295 and Lys 296, is likely to participate in substrate binding, thus to be central to the catalytic activity of the enzyme. Interestingly, this would resemble the Family I active site organisation and their catalytic mechanism, where Arg78 (*S. cerevisiae* numbering) provides a bidentate coordination to P1 (high affinity phosphate group; Heikinheimo et al., 1996a; see Fig. 1.5, p. 47). Consequently, the Arg78Lys mutant shows a decrease in P1 affinity (Tuominen et al. 1998). Also, two Lys residues are involved in the Family I enzymes catalysis (Heikinheimo et al., 1996a; see Fig. 1.5, p. 47), interacting (hydrogen and ionic bonds) directly with P1 (Lys193) and P2 (Lys56). In line with the proposed role of Arg295 and Lys 296 in binding of phosphate, are the results of Shizawa et al. (2001), confirming the importance of these residues for the substrate binding and catalysis in *B.*

subtilis enzyme. The catalytic mechanism proposed by Ahn et al. (2001) shows Arg295 (*B. subtilis* numbering) interacting in a bidentate fashion (analogous to the one of Arg78 in yeast) with P1 (Fig. 6.5, p. 182). Lys298 but also Lys205 are proposed to be analogues of the yeast Lys193 and Lys56, respectively.

The 20% decrease in PPase activity caused by the conserved Val300Ile substitution and even more dramatic reductions found for Val300Trp (77%) and Val299Gly/ Val300Trp (99.5%) mutants confirm that there are strict geometric requirements for activity (Fig. 6.3, p. 177). As an example, from the modelling of the *S. gordonii* PPase three-dimensional structure it appears that bulky aromatic ring of a tryptophan in place of Val300 would cause steric clashes with the side chains of both conserved Phe207 and Arg99 (*B. subtilis* numbering), the residues that otherwise interact with invariant Val300 (S. White, personal communication). Clearly, preservation of the local fold must be important (helix K, Fig. 6.3, p. 177). Introduction of a flexible glycine would additionally destabilize the local structure, possibly also by a dramatic change to the hydrophobicity of the local environment.

An apparent decrease in sensitivity to EDTA of the two Ser294 mutants, Ser294Cys and Ser294Val, accompanied by the increase in sensitivity to fluoride, as well as decrease in sensitivity to EDTA of the double valine mutant, Val299Gly/Val300Trp (Table 5.3, p. 163), all might be of significance when the mechanism of both EDTA inhibition and fluoride insensitivity of *B. subtilis* PPase are discussed. All the above mutations cause quite drastic changes to the local environment (polar to hydrophobic or hydrophobic to hydrophilic). Most likely, the observed changes in the response to EDTA/fluoride are the result of the general disturbance to the function of the neighbouring residues, causing e. g. release of some ligands (making them free for new interactions) and displacement of

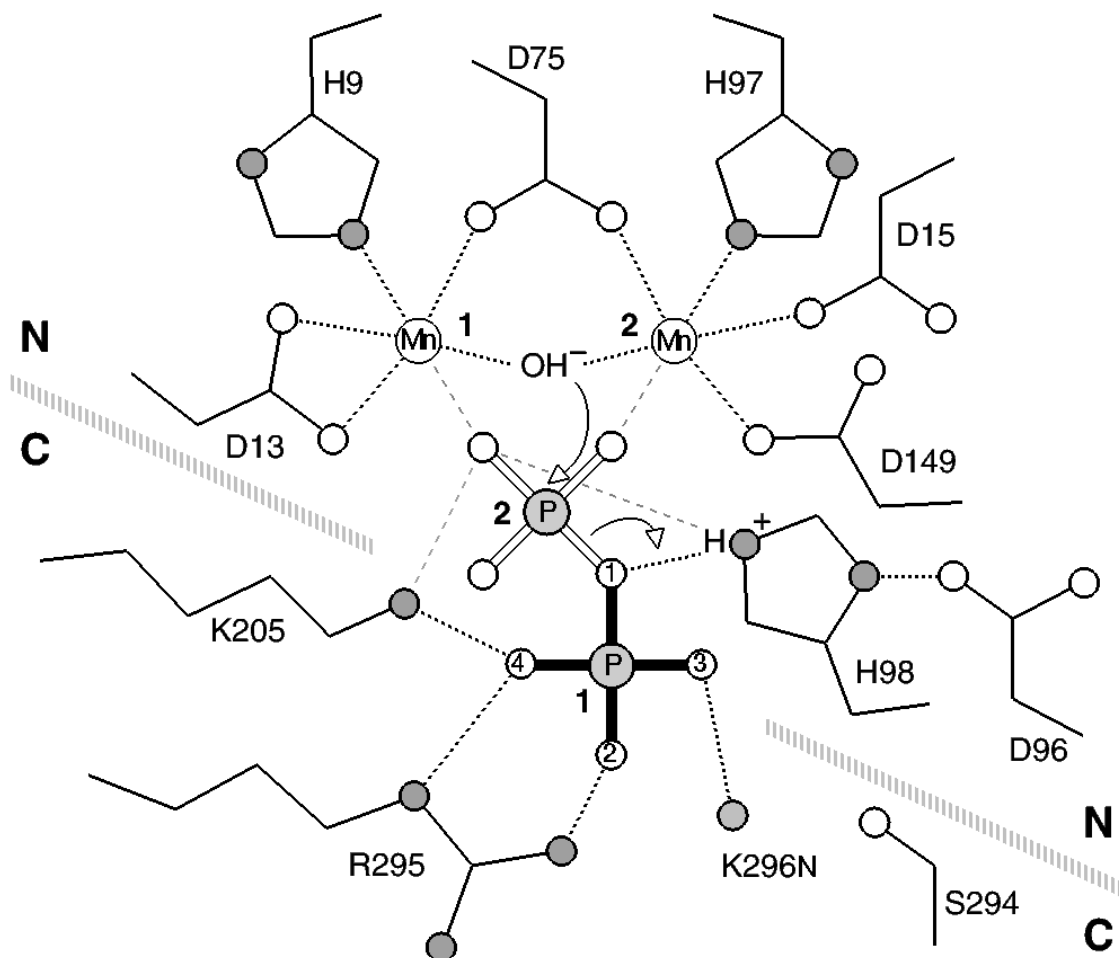


Figure 6.5: The proposed catalytic mechanism for Family II PPases (adopted from Konopka et al. 2002). The *hatched* line is the interface between the N- and C-terminal domains. The ‘SRKK’ motif (Ser294, Arg295, Lys296, Lys297) is located at the C-terminus. The pyrophosphate molecules have been fitted by modelling. P1 (*filled* bonds) is occupied by sulphate in the crystal structure. H-bonds and Lewis acid interactions determined from the structure of the *S. gordonii* enzyme are shown as *dashed* lines and potential H-bonds to the modelled P2 (*open* bonds) as *dotted* lines. The manganese ions act as Lewis acids, polarising the P-O bonds, the same making P2 more electrophilic. Hydrolysis precedes by an attack of the nucleophilic water-derived hydroxide ion (Wat1 bridged between manganese ions 1 and 2) on P2, which acquires its oxygen. Finally, His98 positioned ideally to act as a general acid and held in a fixed orientation by a salt-bridge formed to Asp96, protonates the ‘leaving group’ (P1), the same increasing the rate of hydrolysis. Only the latter stands in a sharp contrast with what was proposed for the Family I enzymes catalysis, where the water molecule (Wat6) positioned close to M3 was designated to act as a general acid (Fig. 1.5, p. 47). S294 is not proposed to be directly involved in pyrophosphate cleavage

others (accompanied by production of unfavourable forces). Combination of the two effects (higher sensitivity to fluoride and lower sensitivity to EDTA) for the same mutant variant could suggest that the two mechanisms, of EDTA inhibition and fluoride insensitivity, are interrelated.

Taken together, the results presented here show that residues in the invariant C-terminal region, the signature sequence, constitute part of the strictly conserved framework needed to preserve the functionality of the active site. This is achieved both in terms of the residues directly involved in catalysis and the correct folding pattern of this part of the protein.

A structural role for conserved regions in proteins is not unusual. A very good example of a high sequence conservation not purely due to the preservation of residues directly involved in enzyme catalysis is the carboxyl terminus of *Zymomonas mobilis* pyruvate decarboxylase (Chang et al. 2000). Many of the residues are conserved among the various pyruvate decarboxylase sequences, especially according to their source and in plant representatives there is a stretch of sixteen aa (His40 – Val55, *Z. mobilis* numbering) conserved in 10 out of 11 representatives. Part of the C-terminal residues form an α helix (Glu46 – Ser60, *Z. mobilis* sequence) that spans the subunit interface. Apart from being responsible for substrate specificity and enzyme affinity to its cofactors, the helix plays a vital structural role allowing closure of the active site. Another example of a protein where a sequence conservation is linked to a structural role is a cyclic AMP-dependent protein kinase. A conserved α helix as part of a core conserved throughout this large protein kinase family interacts with many different parts of the protein, playing a vital role in the active site dynamics, an absolute requirement for catalysis (Taylor et al. 1999).

Interestingly, stabilization of important structural features might occur due to the

preservation of the protein fold itself, as happens for the most populated protein folds (Mirny and Shakhnovich 1999). To this group belongs a Rossmann fold reminiscent of which is the topology of the first 130 amino acid residues of the N-terminal domain of the *B. subtilis* PPase (Ahn et al. 2001). However, SH2, another domain with an invariant Arg residue coordinating the phosphate oxygens and identified as the structural homologue of the C-terminal lobe of the *B. subtilis* PPase (Ahn et al. 2001), is also widely distributed among proteins.

This work in analysing the role of the C-terminal domain in catalysis and protein conformation of *Bacillus subtilis* PPase provides important insights to facilitate design of specific drugs. They would be effective against the significant number of pathogenic microorganisms (e. g. *S. mutans*, *S. gordonii*, *S. pneumoniae*, *Staph. aureus*), which depend upon a Family II inorganic pyrophosphatase for macromolecular synthesis.

Chapter 7

Future Perspectives

The research described in this thesis, into the role of the *Bacillus subtilis* PPase C-terminus, opens a wide range of possibilities for the continuation of the current study. The most relevant techniques are spectroscopic and kinetic analysis. Spectroscopic analysis of mutant proteins would enable the evaluation of the role of the C-terminus in protein secondary structure, but also metal and substrate binding. The last two properties could also be investigated by colorimetric techniques. Kinetic methods would allow an insight into the kinetic characterization of the pyrophosphate catalysis, the same enable deductions about the structural features of the enzyme active site and, possibly, the reaction mechanism.

7.1. Spectroscopic Techniques

7.1.1. Circular Dichroism (CD)

CD is a useful technique to gain structural information about the molecule of interest. It measures the optical activity of asymmetric molecules in solution and gives an information

about the unequal absorption of left- and right-handed circularly-polarized light as a function of wavelength. The CD spectrum in the far-UV region or amide region (175 – 250 nm) is dominated by contributions of the peptide bonds, and provides information about the secondary structure of a protein (structure of the amino acid backbone). In particular, the spectrum of an α helix is very strong and characteristic. The near-UV (250 – 300 nm) CD spectrum on the other hand, originating from the aromatic amino acids, is a highly sensitive measure of the native state of a protein. Consequently, it may be used as a fingerprint for the correctly folded conformation. CD spectrum therefore, especially in the near-UV region, would be a good method for a direct verification of the effect of various mutations on the protein structure.

7.1.2. Fourier Transform Infrared Spectroscopy (FTIR)

FTIR is a powerful technique for structural characterization of proteins. The infrared (IR) spectrum of an organic compound is like a molecular ‘fingerprint’ and provides specific information about chemical bonding (functional groups) and atomic structure. The general structure of compounds may be determined by the spectral locations of their IR absorptions as chemical bonds absorb infrared energy at specific wavelengths. FTIR, providing information unique to each molecule, could serve as an alternative to CD method to determine how much the structure of PPase is affected by specific mutations (especially important for tryptophan fluorescence studies).

7.1.3. Fluorescence Spectroscopy

Substrate specificity studies on the full-length *B. subtilis* PPase and the 5 aa C-terminally truncated mutant (5AATR), combined with the dilution experiments (section 4.5), suggested the possibility of the existence of an alternative, less closed conformation of the C-terminal domain (see also Fig. 6.1, p. 175). The phenomenon seems specifically pronounced for 5AATR as its activity towards larger substrate molecules, e.g. tri-phosphate increases compared to the full-length enzyme. However, under special circumstances the full-length protein appears capable of adopting the more ‘open’ conformation as well. It is able to bind and catalyze tri-phosphate at the low level, but the activity towards this substrate significantly increases with the dilution of the enzyme (Fig. 4.8, p. 153). Dilution seems to promote the ‘opening’ of the structure. These observations seem to support the hypothesis of an ‘open’, flexible conformation or the induced-fit mechanism, both proposed by Ahn et al. (2001). The ‘open’ conformation is proposed to close around pyrophosphate as it binds at the active site, which is formed at the domain interface. The results of this work, supporting the existence of the less closed/‘open’ conformation of the C-terminal domain, are also in line with the results from gel-filtration (section 3.4) indicating the possibility of a non-spherical shape for the full-length as well as both mutant proteins, 5AATR and 17AATR. Furthermore, in *Streptococcus gordonii* PPase, there is an observed increase in the molecule axial ratio when activating Co^{2+} ions are removed from the medium (M. Ilias unpublished results) and presumably this is also true for *B. subtilis* PPase lacking Mn^{2+} . Overall, it seems that *B. subtilis* PPase has its own dynamics and that the C-terminal part of the protein plays an important role in it. Fluorescence spectroscopy, as one of the spectroscopic techniques that provide structural information about dynamic changes in

proteins, seems to be an ideal tool to investigate the conformational properties of *B. subtilis* enzyme.

The fluorescence of proteins originates from the aromatic residues of Trp, Phe and Tyr. However, the spectrum is usually dominated by the contribution of Trp, as its absorbance at the wavelength of excitation and the quantum yield of emission are considerably greater than those of the other two residues. Besides, the intensity, quantum yield and wavelength of maximum fluorescence emission of Trp is very solvent (environment) dependent. The fluorescence spectrum shifts to shorter wavelength and the intensity increases as the polarity of the solvent surrounding the Trp residue decreases. Consequently, the Trp spectra might get shifted by 10 to 20 nm when the residues are transferred from the surface of the protein to its interior (hydrophobic core).

The original *B. subtilis* PPase is devoid of tryptophan. However, through the use of site-directed mutagenesis tryptophans can be selectively placed in various regions of the protein, providing optical probes of local dynamics and the protein-solvent microenvironment. Additionally, the fluorescent properties of Trp can be studied as a function of chemical denaturant and temperature to provide insights into the global structural and dynamic changes which occur during protein folding and unfolding. A wide range of fluorescence techniques might be used, such as steady-state and time-resolved fluorescence, steady-state and time-resolved polarization or fluorescence-intensity and lifetime quenching.

For the fluorescence studies the Trp mutant, Val300Trp could be used. However, as its specific activity is only about 25% of the native enzyme, the CD spectroscopy would be necessary to assess the overall conformation of the mutant relative to the original protein to confirm that they did not much differ in terms of the structural properties. In case

the structural changes were significant, new Trp mutants would be generated. Ideally, Trp should replace residues in the non-conserved regions so that the functional and structural properties of the enzyme are affected as little as possible. The closer they are to the original enzyme, the better probes of the C-terminal dynamics in the native, folded protein the mutant variants would become. It would be interesting to monitor the dynamic changes, if any, upon e.g. binding of substrate, metal ions, etc. Finally, comparison of the properties of the Trp variant the closest in terms of the structure and activity to the native protein to those of the significantly ‘affected’ mutant would give information on the role of the mutated residues (e.g. Val300).

7.2. Isothermal Titration Calorimetry (ITC)

Isothermal Titration Calorimetry allows studying the energy of interaction between two molecules. Upon the interaction, e.g. binding, the heat is released or absorbed. The measure of this process is directly proportional to the amount of binding that occurs. Consequently, the whole set of experimental parameters can be measured, such as binding affinity (K_a), binding stoichiometry (number of binding sites, n), as well as thermodynamic profile of the reaction, which includes enthalpy change (ΔH), heat capacity (ΔC_p) of binding, and the entropy change (ΔS). Large ΔC_p values, for instance, would indicate large scale conformational adaptation of the active site in substrate binding. ITC would be ideal to monitor the binding of phosphate by the *B. subtilis* PPase mutants and to compare their properties to those of the original enzyme.

7.3. Kinetic Studies

Kinetic studies allow determination of the kinetic constants (K_m , V_{max} , k_{cat}) of the analysed reaction. Relevant for these measurements would be continuous flow method monitoring some property change, e.g. absorbance (as a measure of increase in product concentration) in the reaction mixture. In this technique solutions of the enzyme and substrate are introduced into a small mixing chamber and then pumped through a narrow tube attached to a detector, e.g. spectrophotometer.

Measurement of the reaction progress (increase of product concentration as a function of time) allows the initial rate (v_o) to be calculated. By plotting different values of v_o versus varying substrate concentrations $[S]$, V_{max} and K_m ($V_{max}/2$) might be estimated. Alternatively, they may be estimated from the Lineweaver-Burke plot ($1/v_o$ versus $1/[S]$, where y intercept = $1/V_{max}$ and x intercept = $-1/K_m$). K_m (an apparent dissociation constant) is characteristic of an enzyme and related to the enzyme's affinity for the substrate. k_{cat} (the turnover number) is calculated as $V_{max}/[E_{tot}]$, with E_{tot} being an initial (total) concentration of the enzyme. It is a first-order rate constant for converting enzyme-substrate complex to product and represents the maximum number of substrate molecules which can be converted into product per enzyme molecule per unit time (only when enzyme is 'saturated' with substrate). Finally, k_{cat}/K_m may be calculated, a second-order rate constant known as the specificity constant, being a measure of the overall efficiency of the enzyme catalysis.

Further information on the active site architecture may be obtained by studying the influence of pH on the kinetic constants, since catalytic activity relies on a specific ionisation

state of the active site residues. By analysing the variation of $\log K_m$ and $\log V_{\max}$ (or $\log k_{\text{cat}}$) with pH, the pK_a values of the key amino acids involved in the catalysis may be obtained and consequently, these key residues are identified.

List of References

- Accelrys Inc. (2002). DS ViewerPro, http://www.accelrys.com/dstudio/ds_viewer.
- Ahn, S., Milner, A. J., Fütterer, K., Konopka, M., Ilias, M., Young, T. W. and White, S. A. (2001). The “open” and “closed” structures of the type-C inorganic pyrophosphatases from *Bacillus subtilis* and *Streptococcus gordonii*, *J. Mol. Biol.* **313**: 797–811.
- Anderson, R. S., Bollum, F. J. and Beattie, K. L. (1999). Pyrophosphorolytic dismutation of oligodeoxy-nucleotides by terminal deoxynucleotidyltransferase, *Nucleic Acids Res.* **27**: 3190–3196.
- Andreeva, N. A., Kulakovskaya, T. V. and Kulaev, I. S. (2001). Two exopolyphosphatases of the cytosol of the yeast *S. cerevisiae*: comparative characteristics, *Biochem. (Moscow)* **66**: 147–153.
- Avaeva, S., Ignatov, P., Kurilova, S., Nazarova, T., Rodina, E., Vorobyeva, N., Oganessyan, V. and Harutyunyan, E. (1996). *Escherichia coli* inorganic pyrophosphatase: site-directed mutagenesis of the metal binding sites, *FEBS Lett.* **399**: 99–102.
- Avaeva, S. M., Vorobyeva, N. N., Kurilova, S. A., Nazarova, T. I., Polyakov, K. M., Rodina, E. V. and Samygina, V. R. (2000). Mechanism of Ca²⁺-induced inhibition of *Escherichia coli* inorganic pyrophosphatase, *Biochem. (Moscow)* **65**: 373–387.
- Ayabe, T., Takenaka, H., Takenaka, O., Sumida, M., Maruyama, H., Onitsuka, T., Shibata, K., Uesugi, S. and Hamada, M. (1997). Essential lysine residues in the N-terminal and the C-terminal domain of human adenylate kinase interact with adenine nucleotides as found by site-directed random mutagenesis, *Biochemistry* **36**: 4027–4033.
- Baltscheffsky, H. and Baltscheffsky, M. (1986). Light and inorganic pyrophosphate as possible key components in the development of the earliest bioenergetic systems, *Origins Life Evol. Biosphere* **16**: 377–378.
- Baltscheffsky, H., Baltscheffsky, M., Nadanaciva, S., Persson, B. and Schultz, A. (1997a). Possible origin and evolution of inorganic pyrophosphatases, in R. Lahti (ed.), *Proceedings of the First International Meeting on Inorganic Pyrophosphatases*, University of Turku, Turku, Finland, pp. 1–3.
- Baltscheffsky, H., Lundin, M., Luxemburg, C., Nyrén, P. and Baltscheffsky, M. (1986). Inorganic pyrophosphate and the molecular evolution of biological energy coupling, *Chem. Scr.* **26B**: 259–262.

- Baltscheffsky, H., Schultz, A. and Baltscheffsky, M. (2001). On the molecular origin and early photophosphorylation and energy-rich phosphate metabolism, http://www.bio.usyd.edu.au/photoevol/EvolPS_Abstracts.doc.
- Baltscheffsky, M. (1967). Inorganic pyrophosphate and ATP as energy donors in chromatophores from *Rhodospirillum rubrum*, *Nature* **216**: 241–243.
- Baltscheffsky, M., Nadanaciva, S. and Schultz, A. (1998). A pyrophosphate synthase gene: molecular cloning and sequencing of the cDNA encoding the inorganic pyrophosphate synthase from *Rhodospirillum rubrum*, *Biochim. Biophys. Acta* **1364**: 301–306.
- Baltscheffsky, M., Schultz, A. and Baltscheffsky, H. (1999). H⁺-PPases: a tightly membrane-bound family, *FEBS Lett.* **457**: 527–533.
- Baltscheffsky, M., Schultz, A. and Nadanaciva, S. (1997b). The inorganic pyrophosphatase from *Rhodospirillum rubrum* and its gene, in R. Lahti (ed.), *Proceedings of the First International Meeting on Inorganic Pyrophosphatases*, University of Turku, Turku, Finland, pp. 16–18.
- Barkla, B. J. and Pantoja, O. (1996). Physiology of ion transport across the tonoplast of higher plants, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**: 159–184.
- Baykov, A. A., Cooperman, B. S., Goldman, A. and Lahti, R. (1999). Cytoplasmic inorganic pyrophosphatase, *Prog. Mol. Subcell. Biol.* **23**: 127–150.
- Baykov, A. A., Fabrichniy, I. P., Pohjanjoki, P., Zyryanov, A. B. and Lahti, R. (2000). Fluoride effects along the reaction pathway of pyrophosphatase: evidence for a second enzyme-pyrophosphate intermediate, *Biochemistry* **39**: 11939–11947.
- Baykov, A. A., Hyytiä, T., Volk, S. E., Kasho, V. N., Vener, A. V., Goldman, A., Lahti, R. and Cooperman, B. S. (1996). Catalysis by *Escherichia coli* inorganic pyrophosphatase: pH and Mg²⁺ dependence., *Biochemistry* **35**: 4655–4661.
- Baykov, A. A., Tam-Villoslado, J. J. and Avaeva, S. M. (1979). Fluoride inhibition of inorganic pyrophosphatase. IV. Evidence for metal participation in the active center and a four-site model of metal effect on catalysis, *Biochim. Biophys. Acta* **569**: 228–238.
- Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarne, K., Weissenbach, J., Ehrlich, S. D. and Sorokin, A. (2001). The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403, *Genome Res.* **11**: 731–753.
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* **72**: 248–254.
- Chang, A. K., Nixon, P. F. and Duggleby, R. G. (2000). Effects of deletions at the carboxyl terminus of *Zymomonas mobilis* pyruvate decarboxylase on the kinetic properties and substrate specificity, *Biochemistry* **39**: 9430–9437.

- Chen, J., Brevet, A., Fromant, M., L ev eque, F., Schmitter, J.-M., Blanquet, S. and Plateau, P. (1990). Pyrophosphatase is essential for growth of *Escherichia coli*, *J. Bacteriol.* **172**: 5686–5689.
- Chi, A. and Kemp, R. G. (2000). The primordial high energy compound: ATP or inorganic pyrophosphate?, *J. Biol. Chem.* **275**(46): 35677–35679.
- Cooperman, B. S. (1982). The mechanism of action of yeast inorganic pyrophosphatase, *Methods Enzymol.* **87**: 526–548.
- Cooperman, B. S., Baykov, A. A. and Lahti, R. (1992). Evolutionary conservation of the active site of soluble inorganic pyrophosphatase, *Trends Biochem. Sci.* **17**: 262–266.
- Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. and Barton, G. J. (1998). JPred: A consensus secondary structure prediction server, *Bioinformatics* **14**: 892–893.
- Davies, J. M. (1997). Vacuolar energization: pumps, shunts and stress, *J. Exp. Bot.* **48**: 633–641.
- Davies, J. M., Poole, R. J. and Sanders, D. (1993). The computed free energy change of hydrolysis of inorganic pyrophosphate and ATP: apparent significance for inorganic-pyrophosphate-driven reactions of intermediary metabolism, *Biochim. Biophys. Acta* **1141**: 29–36.
- Drozdowicz, Y. M. and Rea, P. A. (2001). Vacuolar H⁺ pyrophosphatases: from the evolutionary backwaters into the mainstream, *Trends Plant Sci.* **6**: 206–211.
- Drozdowicz, Y. M., Kissinger, J. C. and Rea, P. A. (2000). AVP2, a sequence-divergent, K⁺-insensitive H⁺-translocating inorganic pyrophosphatase from *Arabidopsis*, *Plant Physiol.* **123**: 353–362.
- Drozdowicz, Y. M., Lu, Y.-P., Patel, V., Fitz-Gibbon, S., Miller, J. H. and Rea, P. A. (1999). A thermostable vacuolar-type membrane pyrophosphatase from the archaeon *Pyrobaculum aerophilum*: implications for the origins of pyrophosphate-energized pumps, *FEBS Lett.* **460**: 505–512.
- du Jardin, P., Rojas-Beltran, J., Gebhardt, C. and Brasseur, R. (1995). Molecular cloning and characterization of a soluble inorganic pyrophosphatase in potato, *Plant Physiol.* **109**: 853–860.
- Fairchild, T. A. and Patejunas, G. (1999). Cloning and expression profile of human inorganic pyrophosphatase, *Biochim. Biophys. Acta* **1447**: 133–136.
- Farr e, E. M., Geigenberger, P., Willmitzer, L. and Trethewey, R. N. (2000). A possible role for pyrophosphate in the coordination of cytosolic and plastidial carbon metabolism within the potato tuber, *Plant Physiol.* **123**: 681–688.
- Ferretti, J. J., McShan, W. M., Ajdic, D., Savic, D. J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A. N., Kenton, S., Lai, H. S., Lin, S. P., Qian, Y., Jia, H. G., Najjar, F. Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S. W., Roe, B. A. and McLaughlin, R. (2001). Complete genome sequence of an M1 strain of *Streptococcus pyogenes*, *Proc. Natl. Acad. Sci. USA* **98**: 4658–4663.

- Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., Portillo, F. G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J. C., Purcell, R., Rimmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. and Cossart, P. (2001). Comparative genomics of *Listeria* species, *Science* **294**: 849–852.
- Glover, D., McEwen, R. K., Thomas, C. R. and Young, T. W. (1997). pH-regulated expression of the acid and alkaline extracellular proteases of *Yarrowia lipolytica*, *Microbiol.* **143**: 3045–3054.
- Gomez, R., Losada, M. and Serrano, A. (1997). Cyanobacterial and algal inorganic pyrophosphatases and the molecular phylogeny of the higher plant, in R. Lahti (ed.), *Proceedings of the First International Meeting on Inorganic Pyrophosphatases*, University of Turku, Turku, Finland, pp. 6–9.
- Grodberg, J. and Dunn, J. J. (1988). ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification, *J. Bacteriol.* **170**: 1245–1253.
- Hachimori, A., Fujii, T., Ohki, K. and Iizuka, E. (1983). Purification and properties of inorganic pyrophosphatase from porcine brain, *J. Biochem.* **93**: 257–264.
- Hanahan, D. (1983). Studies on transformation of *E. coli* with plasmids, *J. Mol. Biol.* **166**: 557–580.
- Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Umayam, L., Gill, S. R., Nelson, K. E., Read, T. D., Tettelin, H., Richardson, D., Ermolaeva, M. D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R. D., Nierman, W. C., White, O., Salzberg, S. L., Smith, H. O., Colwell, R. R., Mekalanos, J. J., Venter, J. C. and Fraser, C. M. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*, *Nature* **406**: 477–483.
- Heikinheimo, P., Lehtonen, J., Baykov, A., Lahti, R., Cooperman, B. S. and Goldman, A. (1996a). The structural basis for pyrophosphatase catalysis, *Structure* **4**: 1491–1508.
- Heikinheimo, P., Pohjanjoki, P., Helminen, A., Tasanen, M., Cooperman, B. S., Goldman, A., Baykov, A. and Lahti, R. (1996b). A site-directed mutagenesis study of *Saccharomyces cerevisiae* pyrophosphatase. Functional conservation of the active site of soluble inorganic pyrophosphatases, *Eur. J. Biochem* **239**: 138–143.
- Heikinheimo, P., Tuominen, V., Ahonen, A.-K., Teplyakov, A., Cooperman, B. S., Baykov, A. A., Lahti, R. and Goldman, A. (2001). Toward a quantum-mechanical description of metal-assisted phosphoryl transfer in pyrophosphatase, *Proc. Natl. Acad. Sci. USA* **98**: 3121–3126.

- Henin, O., Barbier, B. and Brack, A. (1997). Simple polyanionic peptides as artificial PPases, in R. Lahti (ed.), *Proceedings of the First International Meeting on Inorganic Pyrophosphatases*, University of Turku, Turku, Finland, pp. 3–5.
- Herschlag, D. and Jencks, W. P. (1990). Catalysis of the hydrolysis of phosphorylated pyridines by $\text{Mg}(\text{OH})^+$: a possible model for enzymatic phosphoryl transfer, *Biochemistry* **29**: 5172–5179.
- Higgins, D. G., Thompson, J. D. and Gibson, T. J. (1996). Using CLUSTAL for multiple sequence alignments, *Methods Enzymol.* **266**: 383–402.
- Hoe, H.-S., Kim, H.-K. and Kwon, S.-T. (2001). Expression in *Escherichia coli* of the thermostable inorganic pyrophosphatase from the *Aquifex aeolicus* and purification and characterization of the recombinant enzyme, *Prot. Expr. Pur.* **23**: 242–248.
- Hoe, H.-S., Kim, H.-K. and Kwon, S.-T. (2002). Cloning, analysis, and expression of the gene for inorganic pyrophosphatase of *Aquifex pyrophilus* and properties of the enzyme, *Mol. Cells* **13**: 296–301.
- Ilias, M., Konopka, M. A., Kuhn, N. J., Osborne, J. G., Sampayo, C. D. M., White, S. A. and Young, T. W. (2001). Family II (Class C) soluble inorganic pyrophosphatases, in A. Serrano (ed.), *New trends in inorganic pyrophosphatases research. Proceedings of the Second International Meeting on Inorganic Pyrophosphatases*, CSIC-Universidad de Sevilla, Seville, Spain, pp. 69–73.
- Jensen, C. L., Stephenson, K., Jørgensen, S. T. and Harwood, C. (2000). Cell-associated degradation affects the yield of secreted engineered and heterologous proteins in the *Bacillus subtilis* expression system, *Microbiol.* **146**: 2583–2594.
- Josse, J. (1966). Constitutive inorganic pyrophosphatase of *Escherichia coli*: I. Purification and catalytic properties, *J. Biol. Chem.* **241**: 1938–1947.
- Josse, J. and Wong, S. C. K. (1971). Inorganic pyrophosphatase of *Escherichia coli*, in P. D. Boyer (ed.), *The enzymes*, 3rd edn, Academic Press, New York, chapter 4, pp. 499–527.
- Kankare, J., Salminen, T., Lahti, R., Cooperman, B. S., Baykov, A. A. and Goldman, A. (1996). Crystallographic identification of metal-binding sites in *Escherichia coli* inorganic pyrophosphatase, *Biochemistry* **35**: 4670–4677.
- Kawasaki, I., Adachi, N. and Ikeda, H. (1990). Nucleotide sequence of *S. pombe* inorganic pyrophosphatase, *Nucleic Acids Res.* **18**: 5888.
- Keasling, J. D., Bertsch, L. R. and Kornberg, A. (1993). Guanosine pentaphosphate phosphohydrolase of *Escherichia coli* is a long-chain exopolyphosphatase, *Proc. Natl. Acad. Sci. USA* **90**: 7029–7033.
- Keasling, J. D., Dien, S. J. V. and Pramanik, J. (1998). Engineering polyphosphate metabolism in *Escherichia coli*: implications for bioremediation of inorganic contaminants, *Biotechnol. Bioeng.* **58**: 231–239.

- Khvorova, A., Motorin, Y. and Wolfson, A. D. (1999). Pyrophosphate mediates the effect of certain tRNA mutations on aminoacylation of yeast tRNA^{Phe}, *Nucleic Acids Res.* **27**: 4451–4456.
- Kieber, J. J. and Signer, E. R. (1991). Cloning and characterization of an inorganic pyrophosphatase gene from *Arabidopsis thaliana*, *Plant Mol. Biol.* **16**: 345–348.
- Kim, H. J., Nishikava, S., Tokutomi, Y., Takenaka, H., Hamada, M., Kuby, S. A. and Uesugi, S. (1990). *In vitro* mutagenesis studies at the arginine residues of adenylate kinase. A revised binding site for AMP in the X-ray-deduced model, *Biochemistry* **29**: 1107–1111.
- Knowles, J. R. (1980). Enzyme-catalyzed phosphoryl transfer reactions, *Ann. Rev. Biochem.* **49**: 877–919.
- Kolakowski Jr, L. F., Schloesser, M. and Cooperman, B. S. (1988). Cloning, molecular characterization and chromosome localization of the inorganic pyrophosphatase (PPA) gene from *S. cerevisiae*, *Nucleic Acids Res.* **16**: 10441–10452.
- Konopka, M. A., White, S. A. and Young, T. W. (2002). *Bacillus subtilis* inorganic pyrophosphatase: the C-terminal signature sequence is essential for enzyme activity and conformational integrity, *Biochem. Biophys. Res. Commun.* **290**: 806–812.
- Kornberg, A. (1993). ATP and inorganic pyro- and polyphosphate, *Prot. Sci.* **2**: 131–132.
- Kornberg, A. (1999). Arthur Kornberg, <http://cmgm.stanford.edu/biochem/faculty/kornberg.html>.
- Kuhn, N. J. and Ward, S. (1998). Purification, properties, and multiple forms of a manganese-activated inorganic pyrophosphatase from *Bacillus subtilis*, *Arch. Biochem. Biophys.* **354**: 47–56.
- Kuhn, N. J., Wadeson, A., Ward, S. and Young, T. W. (2000). *Methanococcus jannaschii* ORF mj0608 codes for a class C inorganic pyrophosphatase protected by Co²⁺ or Mn²⁺ ions against fluoride inhibition, *Arch. Biochem. Biophys.* **379**: 292–298.
- Kukko-Kalske, E. and Heinonen, J. (1985). Inorganic pyrophosphate and inorganic pyrophosphatase in *Escherichia coli*, *Int. J. Biochem.* **17**: 575–580.
- Kulaev, I. S. (2000). Biochemistry and biotechnology of inorganic polyphosphates, *Biochem. (Moscow)* **65**(3): 269–270.
- Kulaev, I. S., Kulakovskaya, T. V., Andreeva, N. A. and Lichko, L. P. (1997). Evolution of inorganic polyphosphate functions at different stages of phylogenetic development of living organisms, *J. Evol. Biochem. Physiol.* **33**: 63–71 (Russ.).
- Kulaev, I. S., Vagabov, V. M., Kulakovskaya, T. V., Lichko, L. P., Andreeva, N. A. and Trilisenko, L. V. (2000). The development of A. N. Belozersky's ideas in polyphosphate biochemistry, *Biochem. (Moscow)* **65**: 271–278.

- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessières, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S., Codani, J. J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Düsterhöft, A., Ehrlich, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S. and Galizzi, A. (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*, *Nature* **390**: 249–256.
- Kuranova, I. P., Obmolova, G. V. and Konareva, N. V. (1987). Inorganic pyrophosphatase from the extremely thermophilic bacterium *Thermus thermophilus* HB8, *Dokl. Akad. Nauk. SSSR* **295**: 1013–1016.
- Kuroda, A. and Kornberg, A. (1997). Polyphosphate kinase as a nucleoside diphosphate kinase in *Escherichia coli* and *Pseudomonas aeruginosa*, *Proc. Natl. Acad. Sci. USA* **94**: 439–442.
- Kuroda, A. and Ohtake, H. (2000). Molecular analysis of polyphosphate accumulation in bacteria, *Biochem. (Moscow)* **65**(3): 304–308.
- Kuroda, A., Nomura, K., Ohtomo, R., Kato, J., Ikeda, T., Takiguchi, N., Ohtake, H. and Kornberg, A. (2001a). Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in *E. coli*, *Science* **293**: 705–708.
- Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N. K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K. (2001b). Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*, *Lancet* **357**: 1225–1240.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* **227**: 680–685.
- Lahti, R. (1983). Microbial inorganic pyrophosphatases, *Microbiol. Rev.* **47**: 169–179.
- Lahti, R., Kolakowski, L. F., Heinonen, J., Vihinen, M., Pohjanoksa, K. and Cooperman, B. S. (1990). Conservation of functional residues between yeast and *E. coli* inorganic pyrophosphatases, *Biochim. Biophys. Acta* **1038**: 338–345.
- Lahti, R., Pitkaranta, T., Valve, E., Iltta, I., Kukko-Kalske, E. and Heinonen, J. (1988). Cloning and characterization of the gene encoding inorganic pyrophosphatase of *E. coli*, *J. Bacteriol.* **170**: 5901–5907.
- Leo, C. D., Coppola, R. C., Blasi, G., Eftimiadi, C., Salvarani, M. and Molina, A. M. (1990). Prevalence of *Streptococcus mutans* and dental decay in schoolchildren living in Genoa (Italy), *Eur. J. Epidemiol.* **6**: 166–174.

- Lundin, M., Baltscheffsky, H. and Ronne, H. (1991). Yeast PPA2 gene encodes a mitochondrial inorganic pyrophosphatase that is essential for mitochondrial function, *J. Biol. Chem.* **266**: 12168–12172.
- Maeshima, M. (2000). Vacuolar H⁺-pyrophosphatase, *Biochim. Biophys. Acta* **1465**: 37–51.
- Mansurova, S. E. (1989). Inorganic pyrophosphatase in mitochondrial metabolism, *Biochim. Biophys. Acta* **977**: 237–247.
- Marchesini, N., Luo, S., Rodrigues, C. O., Moreno, S. N. J. and Docampo, R. (2000). Acidocalcisomes and a vacuolar H⁺-pyrophosphatase in malaria parasites, *Biochem. J.* **347**: 243–253.
- Martín, F., Pintor, J., Rovira, J. M., Ripoll, C., Miras-Portugal, M. T. and Soria, B. (1998). Intracellular diadenosine polyphosphates: a novel second messenger in stimulus-secretion coupling, *FASEB J.* **12**: 1499–1506.
- Maruyama, S., Maeshima, M., Nishimura, M., Aoki, M., Ichiba, T., Sekiguchi, J. and Hachimori, A. (1996). Cloning and expression of the inorganic pyrophosphatase gene from thermophilic bacterium PS-3, *Biochem. Mol. Biol. Int.* **40**: 679–688.
- McIntosh, M. T. and Vaidya, A. B. (2002). Vacuolar type H⁺ pumping pyrophosphatases of parasitic protozoa, *Int. J. Parasitol.* **32**: 1–14.
- Merckel, M. C., Fabrichniy, I. P., Salminen, A., Kalkkinen, N., Baykov, A. A., Lahti, R. and Goldman, A. (2001). Crystal structure of *Streptococcus mutans* pyrophosphatase: a new fold for an old mechanism, *Structure* **9**: 289–297.
- Mertens, E. (1991). Pyrophosphate-dependent phosphofructokinase, an anaerobic glycolytic enzyme?, *FEBS Lett.* **285**: 1–5.
- Meyer, W., Moll, R., Kath, T. and Schäffer, G. (1995). Purification, cloning and sequencing of archaeobacterial pyrophosphatase from the extreme thermoacidophile *Sulfolobus acidocaldarius*, *Arch. Biochem. Biophys.* **319**: 149–156.
- Mirny, L. A. and Shakhnovich, E. I. (1999). Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function, *J. Mol. Biol.* **291**: 177–196.
- Mitchell, S. J. and Minnick, M. F. (1997). Cloning, functional expression, and complementation analysis of an inorganic pyrophosphatase from *Bartonella bacilliformis*, *Can. J. Microbiol.* **43**: 734–743.
- Moreno, B., Urbina, J. A., Oldfield, E., Bailey, B. N., Rodrigues, C. O. and Docampo, R. (2000). ³¹P NMR spectroscopy of *Trypanosoma brucei*, *Trypanosoma cruzi*, and *Leishmania major*, *J. Biol. Chem.* **275**(37): 28356–28362.
- Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann,

- R. D., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C. and Fraser, C. M. (1999). Evidence for lateral gene transfer between *Archaea* and *Bacteria* from genome sequence of *Thermotoga maritima*, *Nature* **399**: 323–329.
- Nesmeyanova, M. A. (2000). Polyphosphates and enzymes of polyphosphate metabolism in *Escherichia coli*, *Biochem. (Moscow)* **65**(3): 309–314.
- Nolling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., Lee, H. M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y. I., Tatusov, R. L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M. J., Bennett, G. N., Koonin, E. V. and Smith, D. R. (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*, *J. Bacteriol.* **183**: 4823–4838.
- Novagen (1998). Novagen vector tables, http://www.novagen.com/SharedImages/TechnicalLiterature/7_TB042.pdf.
- Nyrén, P. and Strid, A. (1991). Hypothesis: the physiological role of the membrane-bound proton-translocating pyrophosphatase in some phototrophic bacteria, *FEMS Microbiol. Lett.* **77**: 265–270.
- OCA PDB (2002). OCA Browser, <http://oca.ebi.ac.uk/oca-bin/ocamain>.
- Ogasawara, N. (2000). Systematic function analysis of *Bacillus subtilis* genes, *Res. Microbiol.* **151**: 129–134.
- Oliva, G., Romero, I., Ayala, G., Barrios-Jacobo, I. and Celis, H. (2000). Characterization of the inorganic pyrophosphatase from the pathogenic bacterium *Helicobacter pylori*, *Arch. Microbiol.* **174**: 104–110.
- Parfenyev, A. N., Salminen, A., Halonen, P., Hachimori, A., Baykov, A. A. and Lahti, R. (2001). Quaternary structure and metal ion requirements of family II pyrophosphatases from *Bacillus subtilis*, *Streptococcus gordonii*, and *Streptococcus mutans*, *J. Biol. Chem.* **276**: 24511–24518.
- Pohjanjoki, P., Fabrichniy, I. P., Kasho, V. N., Cooperman, B. S., Goldman, A., Baykov, A. A. and Lahti, R. (2001). Probing essential water in yeast pyrophosphatase by directed mutagenesis and fluoride inhibition measurements, *J. Biol. Chem.* **276**: 434–441.
- Pohjanjoki, P., Lahti, R., Goldman, A. and Cooperman, B. S. (1998). Evolutionary conservation of enzymatic catalysis: quantitative comparison of the effects of mutation of aligned residues in *Saccharomyces cerevisiae* and *Escherichia coli* inorganic pyrophosphatases on enzymatic activity, *Biochemistry* **37**: 1754–1761.
- Rashid, M. H. and Kornberg, A. (2000). Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of *Pseudomonas aeruginosa*, *Proc. Natl. Acad. Sci. USA* **97**: 4885–4890.
- Rashid, M. H., Rumbaugh, K., Passador, L., Davies, D. G., Hamood, A. N., Iglewski, B. H. and Kornberg, A. (2000). Polyphosphate kinase is essential for biofilm development,

- quorum sensing, and virulence of *Pseudomonas aeruginosa*, *Proc. Natl. Acad. Sci. USA* **97**: 9636–9641.
- Rea, P. A. and Poole, R. J. (1993). Vacuolar H⁺-translocating pyrophosphatase, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44**: 157–180.
- Rea, P. A. and Sanders, D. (1987). Tonoplast energization: two H⁺ pumps, one membrane, *Physiol. Plant.* **71**: 131–141.
- Rea, P. A., Kim, Y., Sarafian, V., Poole, R. J., Davies, J. M. and Sanders, D. (1992). Vacuolar H⁺-translocating pyrophosphatases: a new category of ion translocase, *Trends Biochem. Sci.* **71**: 348–353.
- Reusch, R. N. (2000). Transmembrane ion transport by polyphosphate/poly-(R)-3-hydroxybutyrate complexes, *Biochem. (Moscow)* **65**: 280–295.
- Richter, O. M. and Schäfer, G. (1992a). Cloning and sequencing of the gene for the cytoplasmic inorganic pyrophosphatase from the thermoacidophilic archaeobacterium *Thermoplasma acidophilum*, *Eur. J. Biochem.* **209**: 351–355.
- Richter, O. M. and Schäfer, G. (1992b). Purification and enzymic characterization of the cytoplasmic pyrophosphatase from the thermoacidophilic archaeobacterium *Thermoplasma acidophilum*, *Eur. J. Biochem.* **209**: 343–349.
- Rodrigues, C. O., Scott, D. A. and Docampo, R. (1999). Characterization of a vacuolar pyrophosphatase in *Trypanosoma brucei* and its localization to acidocalcisomes, *Mol. Cell. Biol.* **19**: 7712–7723.
- Ruiz, F. A., Marchesini, N., Seufferheld, M., Govindjee and Docampo, R. (2001). The polyphosphate bodies of *Chlamydomonas reinhardtii* possess a proton-pumping pyrophosphatase and are similar to acidocalcisomes, *J. Biol. Chem.* **276**: 46196–46203.
- Rusnak, F. (2000). Manganese-activated phosphatases, *Met. Ions Biol. Syst* **37**: 305–343.
- Ryan, L. M. and McCarty, D. J. (1995). Understanding inorganic pyrophosphate metabolism: toward prevention of calcium pyrophosphate dihydrate crystal deposition, *Ann. Rheum. Dis.* **54**(12): 939–941.
- Salminen, T., Käpylä, J., Heikinheimo, P., Kankare, J., Goldman, A., Heinonen, J., Baykov, A. A., Cooperman, B. S. and Lahti, R. (1995). Structure and function analysis of *Escherichia coli* inorganic pyrophosphatase: a hydroxide ion the key to catalysis?, *Biochemistry* **34**: 782–791.
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, second edn, Cold Spring Harbor Laboratory, NY.
- Samygina, V. R., Popov, A. N., Rodina, E. V., Vorobyeva, N. N., Lamzin, V. S., Polyakov, K. M., Kurilova, S. A., Nazarova, T. I. and Awaeva, S. M. (2001). The structures of *Escherichia coli* inorganic pyrophosphatase complexed with Ca²⁺ or CaPP_i at atomic resolution and their mechanistic implications, *J. Mol. Biol.* **314**: 633–645.

- Satoh, T., Samejima, T., Watanabe, M., Nogi, S., Takahashi, Y., Kaji, H., Teplyakov, A., Obmolova, G., Kuranova, I. and Ishii, K. (1998). Molecular cloning, expression, and site-directed mutagenesis of inorganic pyrophosphatase from *Thermus thermophilus* HB8, *J. Biochem (Tokyo)* **124**: 79–88.
- Satoh, T., Shinoda, H., Ishii, K., Koyama, M., Sakurai, N., Kaji, H., Hachimori, A., Irie, M. and Samejima, T. (1999). Primary structure, expression, and site-directed mutagenesis of inorganic pyrophosphatase from *Bacillus stearothermophilus*, *J. Biochem.* **125**: 48–57.
- Schröder, H. C., Kurz, L., Müller, W. E. G. and Lorenz, B. (2000). Polyphosphate in bone, *Biochem. (Moscow)* **65**(3): 296–303.
- Setlow, P. (1994). Mechanisms which contribute to the long-term survival of spores of *Bacillus* species, *J. Appl. Bacteriol.* **76**: S49–S60.
- Shiba, T., Tsutsumi, K., Ishige, K. and Noguchi, T. (2000). Inorganic polyphosphate and polyphosphate kinase: their novel biological functions and implications, *Biochem. (Moscow)* **65**(3): 315–323.
- Shiba, T., Tsutsumi, K., Yano, H., Ihara, Y., Kameda, A., Tanaka, K., Takahashi, H., Munekata, M., Rao, N. N. and Kornberg, A. (1997). Inorganic polyphosphate and the induction of *rpoS* expression, *Proc. Natl. Acad. Sci. USA* **94**: 11210–11215.
- Shimizu, T., Imai, M., Araki, S., Kishida, K., Terasawa, Y. and Hachimori, A. (1997). Some properties of inorganic pyrophosphatase from *Bacillus subtilis*, *Int. J. Biochem. Cell Biol.* **29**(2): 303–310.
- Shintani, T., Uchiumi, T., Yonezawa, T., Salminen, A., Baykov, A. A., Lahti, R. and Hachimori, A. (1998). Cloning and expression of a unique inorganic pyrophosphatase from *Bacillus subtilis*: evidence for a new family of enzymes, *FEBS Lett.* **439**: 263–266.
- Shizawa, N., Uchiumi, T., Taguchi, J., Kisseleva, N. A., Baykov, A. A., Lahti, R. and Hachimori, A. (2001). Directed mutagenesis studies of the C-terminal fingerprint region of *Bacillus subtilis* pyrophosphatase, *Eur. J. Biochem.* **268**: 5771–5775.
- Sivula, T., Salminen, A., Parfenyev, N., Pohjanjoki, P., Goldman, A., Cooperman, B. S., Baykov, A. A. and Lahti, R. (1999). Evolutionary aspects of inorganic pyrophosphatase, *FEBS Letters* **454**: 75–80.
- Solomon, J. M. and Grossman, A. D. (1996). Who's competent and when: regulation of natural competence in bacteria, *Trends Genet.* **12**(4): 150–155.
- Stark, M. J. and Milner, J. S. (1989). Cloning and analysis of the *Kluyveromyces lactis* TRP1 gene: a chromosomal locus flanked by genes encoding inorganic pyrophosphatase and histone H3, *Yeast* **5**: 35–50.
- Stec, B., Holtz, K. M. and Kantrowitz, E. R. (2000). A revised mechanism for the alkaline phosphatase reaction involving three metal ions, *J. Mol. Biol.* **299**: 1303–1311.

- Stitt, M. (1990). Fructose-2,6-bisphosphate as a regulatory molecule in plants, *Annu. Rev. Plant Mol. Biol.* **41**: 153–185.
- Studier, F. W. and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *J. Mol. Biol.* **189**: 113–130.
- Taylor, S. S., Radzio-Andzelm, E., Madhusudan, Cheng, X., Eyck, L. T. and Narayana, N. (1999). Catalytic subunit of cyclic AMP-dependent protein kinase: structure and dynamics of the active site cleft, *Pharmacol. Ther.* **82**: 133–141.
- Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H., Dodson, R. J., Durkin, A. S., Gwinn, M., Kolonay, J. F., Nelson, W. C., Peterson, J. D., Umayam, L. A., White, O., Salzberg, S. L., Lewis, M. R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A. M., Utterback, T. R., Hansen, C. L., McDonald, L. A., Feldblyum, T. V., Angiuoli, S., Dickinson, T., Hickey, E. K., Holt, I. E., Loftus, B. J., Yang, F., Smith, H. O., Venter, J. C., Dougherty, B. A., Morrison, D. A., Hollingshead, S. K. and Fraser, C. M. (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*, *Science* **293**: 498–506.
- TIGR (2002). TIGR Microbial Database: a listing of microbial genomes and chromosomes in progress, <http://www.tigr.org/tdb/mdb/mdbinprogress.html>.
- Tono, H. and Kornberg, A. (1967). Biochemical studies of bacterial sporulation. Inorganic pyrophosphatase of vegetative cells and spores of *Bacillus subtilis*, *J. Biol. Chem.* **242**: 2375–2382.
- Tsutsumi, K., Munekata, M. and Shiba, T. (2000). Involvement of inorganic polyphosphate in expression of SOS genes, *Biochim. Biophys. Acta* **1493**: 73–81.
- Tuominen, V., Heikinheimo, P., Kajander, T., Torkkel, T., Hyytiä, T., Käpylä, J., Lahti, R., Cooperman, B. S. and Goldman, A. (1998). The R78K and D117E active-site variants of *Saccharomyces cerevisiae* soluble inorganic pyrophosphatase: structural studies and mechanistic implications, *J. Mol. Biol.* **284**: 1565–1580.
- Urbina, J. A., Moreno, B., Vierkotter, S., Oldfield, E., Payares, G., Sanoja, C., Bailey, B. N., Yan, W., Scott, D. A., Moreno, S. N. J. and Docampo, R. (1999). *Trypanosoma cruzi* contains major pyrophosphate stores, and its growth *in vitro* and *in vivo* is blocked by pyrophosphate analogs, *J. Biol. Chem.* **274**: 33609–33615.
- Vagabov, V. M., Trilisenko, L. W. and Kulaev, I. S. (2000). Dependence of inorganic polyphosphate chain length on the orthophosphate content in the culture medium of the yeast *Saccharomyces cerevisiae*, *Biochem. (Moscow)* **65**: 349–354.
- van Alebeek, G.-J., Keltjens, J. T. and van der Drift, C. (1994). Purification and characterization of inorganic pyrophosphatase from *Methanobacterium thermoautotrophicum*, *Biochim. Biophys. Acta* **1206**: 231–239.
- Varma, A. K., Rigsby, W. and Jordan, D. C. (1983). A new inorganic pyrophosphate utilizing bacterium from a stagnant lake, *Can. J. Microbiol* **29**: 1470–1474.

- Vijay, K., Brody, M. S., Fredlund, E. and Price, C. W. (2000). A PP2C phosphatase containing a PAS domain is required to convey signals of energy stress to the σ^B transcription factor of *Bacillus subtilis*, *Mol. Microbiol.* **35**: 180–188.
- Visser, K., Heimovaara-Dijkstra, S., Kijne, J. W. and Wang, M. (1998). Molecular cloning and characterization of an inorganic pyrophosphatase from barley, *Plant Mol. Biol.* **37**: 131–140.
- Volk, S. E., Dudarenkov, V., Käpylä, J., Kasho, V. N., Voloshina, O. A., Salminen, T., Goldman, A., Lahti, R., Baykov, A. A. and Cooperman, B. S. (1996). Effect of E20 substitution in the active site of *Escherichia coli* inorganic pyrophosphatase on its quaternary structure and catalytic properties., *Biochemistry* **35**: 4662–4669.
- Wakagi, T., Oshima, T., Imamura, H. and Matsuzawa, H. (1998). Cloning of the gene for inorganic pyrophosphatase from a thermoacidophilic archaeon, *Sulfolobus* sp. strain 7, and overproduction of the enzyme by coexpression of tRNA for arginine rare codon, *Biosci. Biotechnol. Biochem.* **62**: 2408–2414.
- Weiner, H., Stitt, M. and Heldt, H. W. (1987). Subcellular compartmentation of pyrophosphate and alkaline pyrophosphatase in leaves, *Biochim. Biophys. Acta* **893**: 13–21.
- White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., Dodson, R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., Moffat, K. S., Qin, H., Jiang, L., Pampile, W., Crosby, M., Shen, M., Vamathevan, J. J., Lam, P., McDonald, L., Utterback, T., Zalewski, C., Makarova, K. S., Aravind, L., Daly, M. J., Fraser, C. M. and et al. (1999). Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1, *Science* **286**: 1571–1577.
- Wolfson, A. D. and Uhlenbeck, O. C. (2002). Modulation of tRNA^{Ala} identity by inorganic pyrophosphatase, *Proc. Natl. Acad. Sci. USA* **99**: 5965–5970.
- Wood, H. G. (1977). Some reactions in which inorganic pyrophosphate replaces ATP and serves as a source of energy., *Fed. Proc.* **36**: 2197–2205.
- Wood, H. G. (1985). Inorganic pyrophosphate and polyphosphates as sources of energy, *Curr. Topics Cell. Regul.* **26**: 355–369.
- Yamagata, Y., Watanabe, H., Saitoh, M. and Namba, T. (1991). Volcanic production of polyphosphates and its relevance to prebiotic evolution, *Nature* **352**: 516–519.
- Yang, Z. and Wensel, T. G. (1992). Molecular cloning and functional expression of cDNA encoding a mammalian inorganic pyrophosphatase, *J. Biol. Chem.* **267**: 24641–24647.
- Young, T. W., Kuhn, N. J., Wadeson, A., Ward, S., Burges, D. and Cooke, G. D. (1998). *Bacillus subtilis* ORF *yybQ* encodes a manganese-dependent inorganic pyrophosphatase with distinctive properties: the first of a new class of soluble pyrophosphatase?, *Microbiol.* **144**: 2563–2571.
- Zyranov, A. B., Pohjanjoki, P., Kasho, V. N., Shestakov, A. S., Goldman, A., Lahti, R. and Baykov, A. A. (2001). The electrophilic and leaving group phosphates in the catalytic mechanism of yeast pyrophosphatase, *J. Biol. Chem.* **276**: 17629–17634.

Chapter 8

Publications