TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. A I OSA - TOM. 421

ASTRONOMICA - CHEMICA - PHYSICA - MATHEMATICA

FUNCTIONAL STUDIES ON BACTERIAL NUCLEOTIDE-REGULATED INORGANIC PYROPHOSPHATASES

by

Joonas Jämsén

TURUN YLIOPISTO
UNIVERSITY OF TURKU
Turku 2011

Department of Biochemistry and Food Chemistry
University of Turku
Turku, Finland

Supervised by
Professor Reijo Lahti
Department of Biochemistry and Food Chemistry
University of Turku
Turku, Finland

AND

Professor Alexander Baykov
A.N. Belozersky Institute of Physico-Chemical Biology
Moscow State University
Moscow, Russia

Reviewed by
Professor Ari Hinkkanen
A.I. Virtanen Institute
Department of Biotechnology and Molecular Medicine
University of Eastern Finland
Kuopio, Finland

AND

Professor Simo Laakso
Department of Biotechnology and Chemical Technology
Aalto University
Helsinki, Finland

Opponent
Professor Markku Kulomaa
Institute of Medical Technology
University of Tampere
Tampere, Finland

ISBN 978-951-29-4659-4 (PRINT)
ISBN 978-951-29-4660-0 (PDF)
ISSN 0082-7002
Painosalama Oy – Turku, Finland 2011

ABSTRACT

CBS domains are ~60 amino acid tandemly repeated regulatory modules forming a widely distributed domain superfamily. Found in thousands of proteins from all kingdoms of life, CBS domains have adopted a variety of functions during evolution, one of which is regulation of enzyme activity through binding of adenylate-containing compounds in a hydrophobic cavity. Mutations in human CBS domain-containing proteins cause hereditary diseases.

Inorganic pyrophosphatases (PPases) are ubiquitous enzymes, which pull pyrophosphate (PP_i) producing reactions forward by hydrolyzing PP_i into phosphate. Of the two nonhomologous soluble PPases, dimeric family II PPases, belonging to the DHH family of phosphoesterases, require a transition metal and magnesium for maximal activity. A quarter of the almost 500 family II PPases, found in bacteria and archaea, contain a 120-250 amino acid N-terminal insertion, comprised of two CBS domains separated in sequence by a DRTGG domain. These enzymes are thus named CBS-PPases. The function of the DRTGG domain in proteins is unknown.

The aim of this PhD thesis was to elucidate the structural and functional differences of CBS-PPases in comparison to family II PPases lacking the regulatory insert. To this end, we expressed, purified and characterized the CBS-PPases from *Clostridium perfringens* (*cp*CBS-PPase) and *Moorella thermoacetica* (*mt*CBS-PPase), the latter lacking a DRTGG domain. Both enzymes are homodimers in solution and display maximal activity against PP_i in the presence of Co^{2+} and Mg^{2+}. Uniquely, the DRTGG domain was found to enable tripolyphosphate hydrolysis at rates similar to that of PP_i. Additionally, we found that AMP and ADP inhibit, while ATP and AP_4A activate CBS-PPases, thus enabling regulation in response to changes in cellular energy status.

We then observed substrate- and nucleotide-induced conformational transitions in *mt*CBS-PPase and found that the enzyme exists in two differentially active conformations, interconverted through substrate binding and resulting in a 2.5-fold enzyme activation. AMP binding was shown to produce an alternate conformation, which is reached through a different pathway than the substrate-induced conformation. We solved the structure of the regulatory insert from *cp*CBS-PPase in complex with AMP and AP_4A and proposed that conformational changes in the loops connecting the catalytic and regulatory domains enable activity regulation. We examined the effects of mutations in the CBS domains of *mt*CBS-PPase on catalytic activity, as well as, nucleotide binding and inhibition.

TABLE OF CONTENTS

ABBREVIATIONS

SAM	S-adenosyl methionine
AMP	Adenosine 5'-monophosphate
ADP	Adenosine 5'-diphosphate
ATP	Adenosine 5'-triphosphate
GMP	Guanosine 5'-monophosphate
GDP	Guanosine 5'-diphosphate
GTP	Guanosine 5'-triphosphate
NAD^+	Nicotinamide adenine dinucleotide, reduced
NADH	Nicotinamide adenine dinucleotide
AP_nA	Diadenosine polyphosphate (n = chain length)
AP_2A	Diadenosine diphosphate
AP_4A	Diadenosine tetraphosphate
CBS	Cystathionine-β-synthase
IMPDH	5'-Inosine monophosphate dehydrogenase
AMPK	5'-Adenosine monophosphate-activated protein kinase
CLC	Voltage-gated chloride channel
P_i	Orthophosphate
PP_i	Inorganic pyrophosphate
PPase	Inorganic pyrophosphatase
*ec*PPase	PPase from *Escherichia coli* (family I PPase)
*sc*PPase	PPase from *Saccharomyces cerevisiae* (family I PPase)
*bs*PPase	PPase from *Bacillus subtilis* (family II PPase)
*sg*PPase	PPase from *Streptoccus gordonii* (family II PPase)
smPPase	PPase from *Streptoccus mutans* (family II PPase)
CBS-PPase	CBS domain-containing PPase *(subfamily of family II PPases)*
*mt*CBS-PPase	CBS-PPase from *Moorella thermoacetica*
*cp*CBS-PPase	CBS-PPase from *Clostrium perfringens*
*fn*CBS-PPase	CBS-PPase from *Fusobacterium nucleatum*
Tris	Tris(hydroxymethyl)aminomethane
MOPS	3-(N-Morpholino)propanesulfonic acid
EDTA	Ethylenediaminetetraacetic acid
DTPA	Diethylenetriaminepentaacetic acid

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred in the text by the Roman numerals I-V. The original communications have been reproduced with the permission of the copyright holders. Unpublished data are included.

I. **Jämsen J***, Tuominen H*, Salminen A, Belogurov GA, Magretova NN, Baykov AA, Lahti R. (2007) A CBS domain-containing pyrophosphatase of *Moorella thermoacetica* is regulated by adenine nucleotides. *Biochem J* **408**, 327-33.
*Equal contribution

II. **Jämsen J**, Baykov AA, Lahti R. (2010) Nucleotide- and substrate-induced conformational transitions in the CBS domain-containing pyrophosphatase of *Moorella thermoacetica. Biochemistry* **49**, 1005-1013.

III. Tuominen H, Salminen A, Oksanen E, **Jämsen J**, Heikkilä O, Lehtiö L., Magretova N.N., Goldman A, Baykov AA, Lahti R. (2010) Crystal structures of the CBS and DRTGG domains of the regulatory region of *Clostridium perfringens* pyrophosphatase complexed with the inhibitor, AMP, and activator, diadenosine tetraphosphate. *J. Mol. Biol*. **398**, 400–413.

IV. **Jämsen J**, Tuominen H, Baykov AA, Lahti R. (2011) Mutational analysis of residues in the regulatory CBS domains of *Moorella thermoacetica* pyrophosphatase corresponding to disease-related residues of human proteins. *Biochem J* **433**, 497-504.

V. **Jämsen J**, Baykov AA, Lahti R. (2011) The CBS-domain containing pyrophosphatase of *Clostridium perfringens*: novel substrate and regulator specificities. *Manuscript.*

1. INTRODUCTION

1.1 Protein domains

Proteins evolve by gene duplication, fusion, genetic recombination and/or divergence, which modify the properties of the protein in order to increase the fitness of the organism in its environment. Protein domains are evolutionarily conserved, independently folding and functioning units of protein and organismal evolution with shared structural, sequence and functional properties. Structural domains, usually 50-200 amino acid residues in length, can be combined in different arrangements with each other and with other domains of different functions to create multidomain proteins with novel or modified properties. These domain combinations can, in turn, evolve to generate other proteins and properties.

Domain families and superfamilies, which have evolved from a common ancestor, are usually composed of domains shared by only a few other superfamilies, occurring in a defined order in sequence (Apic *et al.* 2001, Kummerfeld & Teichmann 2009). The most widely distributed domain families are usually also the most ancient (Orengo & Thornton 2005). Multiple domains adjacent to each other are often the result of single duplication events, which are then combined with other domains and evolve as a single unit (Gough 2005, Bashton & Chothia 2002). This is due to either selection of the functional advantage conferred by the resulting combination of domains or because the function is created by the fusion of the domains in question (Orengo & Thornton 2005). Recently duplicated domains usually occur at the ends of proteins (Weiner & Moore 2008) and are commonly involved in protein-protein interactions (Basu *et al.* 2008). In general, eukaryotes contain proteins composed of more complex combinations of multiple domains than bacteria or archaea (Apic *et al.* 2001, Kummerfeld & Teichmann 2009). Approximately 90% of the domains in eukaryotes and 70% in bacteria have undergone a duplication event (Chothia *et al.* 2003). Some families have undergone massive expansion during evolution, whereas others are highly conserved, reflecting their conserved function throughout life (Woese 2002). Other families have expanded through an increase in organismal complexity, requiring an increased level of regulation (Ranea 2004). Domain families share sequence and

structural similarity, whereas domain superfamilies share a defined 3D structure due to significant divergence of sequence, especially if the domain has evolved specialized functions. Domains found in a wide range of different architectures are able to tolerate substitution of residues structurally better than other folds, allowing sequence to diverge more readily (Murzin 1998, Shakhnovich *et al.* 2003). The 3D structures of domains have been shown to be more conserved than sequence during evolution (Chothia & Lesk 1986), and can be used in the identification of distant relatives (Orengo 1999).

1.2 CBS domains

Intramolecular association of two tandem ~60 amino acid CBS repeats (CBS1 and CBS2) at the N- or C-termini of mainly metabolic enzymes, transcription factors, proteases, and various transporters and channels, forms a so-called CBS pair or Bateman domain. The first members of the CBS domain superfamily were first identified by Alexander Bateman in the genome of *Methanococcus jannaschii* and were named after human cystathionine-β-synthase (Bateman 1997). CBS domains have now been observed in over 15000 proteins (Superfamily Database, Gough *et al.* 2001). CBS domains have a conserved secondary and tertiary structure but a low overall sequence identity, may exist as standalone proteins or as parts of larger multidomain proteins, and seem to have adopted a range of functions during evolution.

As shown in Table 1, the largest number of CBS domains in a given genome is found in green plants, where a single genome contains on average 43 proteins with CBS domains. Chordates contain 26 domains and prokaryotes only 9.4 domains. Specifically, the human genome contains 70 proteins with CBS domains, while e.g. the citrus tree (*Citrus sinensis*) contains 106 CBS proteins or 136 CBS domains. The CBS domain superfamily is functionally widespread, occurring primarily in purine metabolism (39%), active transport of solutes (12%), selenoamino acid metabolism (6%), and methionine metabolism (6%), with the rest distributed sporadically in various functional classes (SMART database, Letunic I *et al.* 2009). CBS domains are found in a number of different architectures. The CBS domain-containing proteins in humans include cystathionine-β-synthase (Mudd *et al.* 1964), 5'-IMP dehydrogenase (Natsumeda *et al.* 1990), 5'-AMP activated protein kinase (Ferrer *et al.* 1985), CLC

voltage-gated chloride channels and transporters (CLC1-7, -Ka, -Kb) (Jentsch & Günther 1997), and the CNNM1-4 metal transporters (Wang *et al.* 2003a). The number of CBS domains in a genome probably reflects their functional specialization during evolution, where the more complex the organism, the greater are the demands placed for regulation of enzymatic function.

Table 1. Species distribution of CBS domains among different phylae. Generated from data in the SUPERFAMILY database (Gough et al. 2001). Values in bold are totals and the names in italics are species.

Species	No. of genomes	CBS domains	Proteins with CBS	Domains per genome
Prokarya	**1157**	**10863**	**10450**	**9,4**
Archaea	86	1418	1089	16,5
Eubacteria	1071	9445	9361	8,8
Eukarya	**305**	**6031**	**4799**	**19,8**
Green plants	**33**	**1796**	**1408**	**54**
Arabidopsis thaliana		55	42	
Chordata	**51**	**1339**	**1133**	**26,2**
Homo sapiens		78	70	
Arthropoda	**32**	**357**	**288**	**11,2**
Drosophila melanogaster		47	31	
Nematoda	**9**	**219**	**179**	**24**
Caenorhabditis elegans		32	26	
Fungi	**118**	**1536**	**1153**	**13**
Saccharomyces cerevisae		13	10	
Other metazoa	**9**	**145**	**126**	**16**
Other eukaryotes	**50**	**639**	**512**	**12,8**
Total	**1462**	**16894**	**15249**	**11,6**

The phyletic, architectural and functional distribution of CBS domains indicates that these domains were present and acquired specialized functions very early in evolution. The conservation of secondary and tertiary structure in time, despite significant changes in the function of individual proteins, supports evolution from a single unit. The identification of the ancient origins of these domains might help in understanding the functional, structural, mechanistic and evolutionary features shared by CBS domains. The aim of Chapter 1.2 of this thesis is to briefly summarize the properties of CBS domain-containing proteins, concentrating primarily on the role of CBS domains in structure, function and regulation.

1.2.1 Structure and function of CBS domains

The first crystal structure of a CBS domain was solved as part of the structure of the 5'-IMPDH of the Chinese hamster, *Cricetulus griseus* (Sintchak *et al.* 1996). Since then, at least 64 crystal structures containing CBS domains have appeared in the PDB (PDB IDs 1O50, 1PBJ, 1PVM, 1VR9, 1XKF, 1Y5H, 1YAV, 1ZFJ, 2D4Z, 2EF7, 2J9L, 2JA3, 2OUX, 2PFI, 2QH1, 2QLV, 2QR1, 2QRC, 2QRD, 2QRE, 2RC3, 2RIF, 2RIH, 2UV4, 2UV5, 2UV6, 2UV7, 2V8Q, 2V92, 2V9J, 2YVZ, 2YVY, 2YVX, 2ZY9, 3B4R, 3DDJ, 3FHM, 3FNA, 3FV6, 3FWR, 3FWS, 3GBY, 3GHD, 3HF7, 3JTF, 3K2V, 3K6E, 3KH5, 3KPB, 3KPC, 3KPD, 3KXR, 3L2B, 3L31, 3LFR, 3LFZ, 3LHH, 3LQN, 3LV9, 3NQR, 3OCO, 3PC2, 3PC3, 3PC4) and 21 publications describing some of these structures in the literature (Gómez García *et al.* 2011, Gómez-García *et al.* 2010, Koutmos *et al.* 2010, Lucas *et al.* 2010, Tuominen *et al.* 2010, Hattori *et al.* 2009, King *et al.* 2008, Proudfoot *et al.* 2008, Ragunathan *et al.* 2008, Sharpe *et al.* 2008, Amodeo *et al.* 2007, Day *et al.* 2007, Feng *et al.* 2007, Hattori *et al.* 2009, Jin *et al.* 2007, Markovic *et al.* 2007, Meyer *et al.* 2007, Xiao *et al.* 2007, Meyer & Dutzler 2006, Miller *et al.* 2004, Zhang *et al.* 1999).

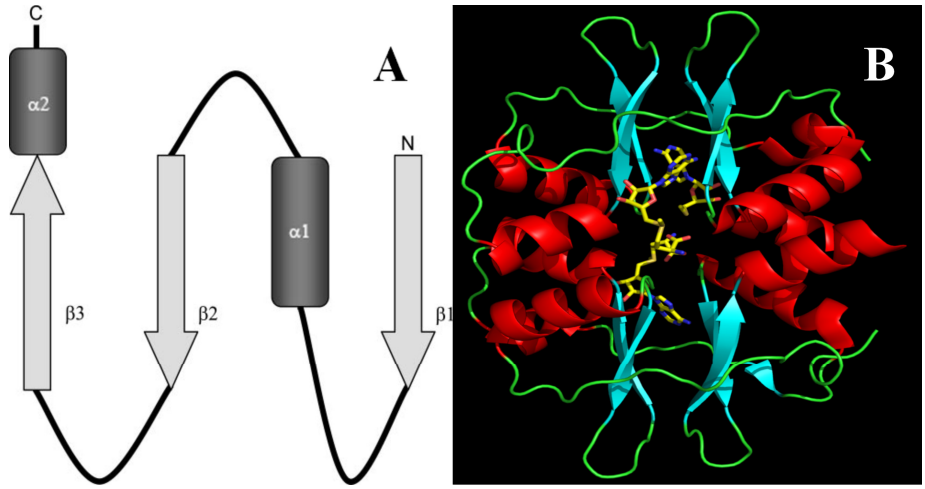


Figure 1. A) Secondary structure of CBS domains (Ignoul & Eggermont 2005). Helices are shown as dark grey rectangles and sheets as light grey arrows. B) 3D structure of the homodimeric CBS dimers of protein MJ0100 in complex with 5-methylthioadenosine and S-adenosyl-L-methionine (PDB ID 3KPD, Lucas *et al.* 2010). Helices are shown in red and sheets in cyan.

CBS domains exhibit low overall sequence similarity (15-25% depending on alignment) but are characterized by the general topology β1-α1-β2-β3-α2 (figure 1A-B) (Bateman 1997). The structures solved to date reveal a hydrophobic cleft at the β2-β3 interface of the CBS1 and CBS2 domains between two three-stranded antiparallel β-sheets lined by two α-helices on one side of a disk-like structure, forming the binding site for various ligands thus far observed to be bound there (AMP, ADP, ATP, AP_4A, SAM, MTA, ZMP). Interestingly, recent crystal structures contain either one or two ligands per CBS pair, where two symmetric cavities on either side of the central cleft have been shown to bind ligands. CBS domains may thus bind either one or two ligands per domain pair (Lucas *et al.* 2010). CBS domain-containing proteins mostly have either two or four CBS domains in sequence, which form dimers or tetramers in solution, while multimers of up to 16 pairs of domains have been observed (Frank *et al.* 2006). Mostly, the interactions between CBS domains are either head-to-head or head-to-tail (Rudolph *et al.* 2007). Interestingly, crystal structures of the CBS domains of voltage-gated chloride channels CLC-0, -5 and –Ka, showed that the monomers interact at an angle of 90° in contrast to other CBS domain-containing proteins. (Meyer & Dutzler 2006, Markovic & Dutzler 2007, Meyer *et al.* 2007). Despite the low sequence similarity between the two CBS domains of a pair, the 3D structures seem to be well conserved.

The functions of CBS domains have been reported for only a few proteins, and the true *in vivo* functions, including their role in the regulation of the functional properties of CBS domain-containing proteins in most cases remains to be elucidated (Ignoul & Eggermont, 2005). To date, CBS domains have been shown to function in a number of processes, probably more than one of these in most cases being biological functions for this domain.

1. **Mediating regulation in response to environmental changes *via* binding of cytosolic AdoMet, ATP, ADP and/or AMP.** Examples include AdoMet binding by the CBS domains of cystathionine-β-synthase, where a mutation (D444N) causing homocystinuria increased the K_d for SAM. The isolated CBS domains of IMPDH-II were found to bind AMP and ATP in a positively cooperative manner, while the full-length protein also similarly bound and was regulated by ATP (Scott *et al.* 2004). The γ1-3 subunits of AMPK, both full-length and isolated pairs (CBS1-2 and CBS3-4), bind AMP and ATP with positive cooperativity. Disease-causing mutations (R302Q,

T400N, H383R and R531G) increased dissociation constants for both nucleotides (Hardie and Hawley 2001, Scott *et al.* 2004). The CBS pair of ClC-2 binds ATP, which is again weakened by disease-causing mutations (G715E and G826D) (Scott *et al.* 2004). The CBS pair of ClC-5 binds ATP (Meyer *et al.* 2007). ATP binding to the CBS domains of CLC-1 regulates gating (Bennetts *et al.* 2005, Bennetts *et al.* 2007).

2. **Oligomerization scaffolds.** CBS domain-mediated oligomerization has been demonstrated in human cystathionine-β-synthase (Frank *et al.* 2006) compared to the truncated enzyme (Taoka *et al.* 2002). In contrast, bacterial CLCs lack CBS domains, and are dimers, while mammalian CLCs have CBS domains and are also dimeric, so CBS domains may not be required for CLC dimerization (Jentsch 2008).

3. **Facilitators of conformational changes resulting in increased transporter activity.** The MgtE transporter transports Mg^{2+} ions across the plasma membrane, by "sensing" intracellular [Mg^{2+}]. The CBS domain interface thus opens, causing a conformational change in a linker domain, which affects gating of the pore (Hattori *et al.* 2007, 2009). The cytoplasmic regions of CLCs contain a CBS pair and regulate gating of these chloride channels. Their long cytoplasmic region is connected to a helix R, which forms the channel pore and selectivity filter (Dutzler *et al.* 2002). Conformational changes in the cytoplasmic regions are thought to be transmitted to the pore via helix R, where disease mutations in the CBS domains affect channel gating through conformational changes (Meyer & Dutzler, 2006).

4. **Sensors of ionic strength.** The CBS domains in a bacterial glycine betaine transporter OpuA detect changes in intracellular ionic strength and thus control transport activity. This occurs through ion binding to the CBS2 domain and its subsequent interaction with the plasma membrane (Biemans-Oldehinkel *et al.* 2006).

5. **Modulators of receptor trafficking.** CBS domains regulate trafficking of CLC chloride channels. Truncation mutations of the CBS2 domain of CLC5 prevent normal endosome trafficking (Carr *et al.* 2003). Also, di-leucine motifs in CBS2 of CLC2 target the channel to basolateral instead of apical membranes (Pena-Munzenmayer *et al.* 2005).

6. **Binding to nucleic acid.** The IMPDH isoforms from humans, *E. coli* and *Trichomonas foetus* bind around 100 nucleotides of single stranded nucleic acids *in vivo* and *in vitro* (Maclean *et al.* 2004). Retinitis pigmentosa P10-linked mutations decreased affinity for nucleic acid *in vitro* (Mortimer *et al.* 2008). The CBS domains of protein MJ0729 from *Methanocaldococcus jannaschii* bind double-stranded calf

thymus DNA and E-box sequences characteristically bound by HLH transcription factors through conformational changes (Aguado-Llera *et al.* 2010).

Mutations in the CBS domains of human proteins impair the catalytic activity or regulatory function of the CBS domains, sometimes resulting in serious diseases (Ignoul & Eggermont, 2005). These mutations include, among others, D444N and V456P in cystathionine-β-synthase causing homocystinuria (Kluijtmans *et al.* 1996, Shan and Kruger 1998, Maclean *et al.* 2002, Urreizti *et al.* 2003, Oliveira *et al.* 2003), R224P and D226N in IMPDH causing retinitis pigementosa (Kennan *et al.* 2002, Bowne *et al.* 2002), R302Q and H383R in AMPKγ2 causing Wolf-Parkinson-White syndrome (Gollob *et al.* 2001) and hypertrophic cardiomyopathy (Blair *et al.* 2001), respectively. Mutations L766P and R767Q/W in CLC7 cause osteopetrosis (Cleiren *et al.* 2001). Mutations thus affect oligomerization of cysthathionine-β-synthase (Jhee *et al.* 2000a), affect ligand binding in many proteins (Scott *et al.* 2004, Jin *et al.* 2007, Day *et al.* 2007, Meyer *et al.* 2007, Tanaka *et al.* 2007), affect energy metabolism (Scott *et al.* 2004), alter ion channel gating (Estevez *et al.* 2004), cause muscle disorders (de Diego *et al.* 1999), and alter protein trafficking (Pena-Munzenmayer *et al.* 2005). Thus CBS domains have a variety of important functions in humans, but detailed knowledge of their structure-function relations are still lacking.

1.2.2 Human CBS domain-containing proteins

As shown in figure 2, CBS domains are found in variety of human proteins, such as cystathionine-β-synthase, IMP dehydrogenase, CLC chloride channels, AMP-activated protein kinase, and others (Ignoul & Eggermont 2005).

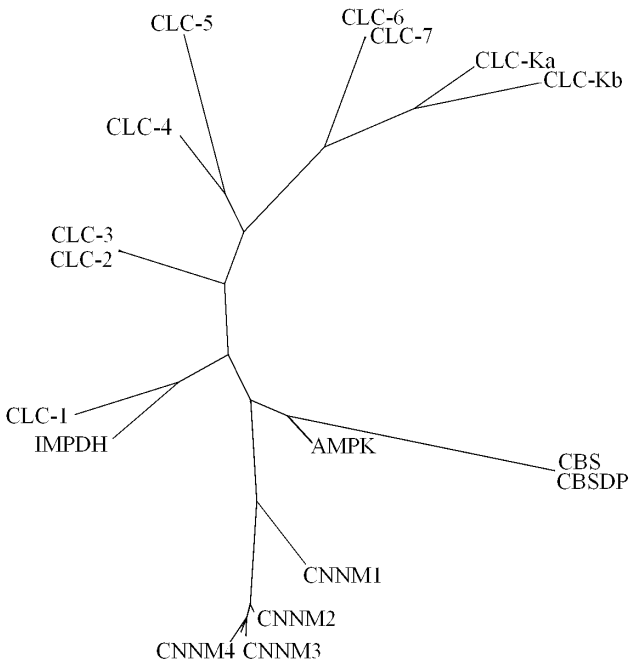


Figure 2. Neighbor-joining phylogenetic tree of the CBS domains found in the human genome. Sequences were retrieved by blast searches of the NR database at the NCBI (www.ncbi.nlm.nih.gov). CBSDP is a CBS domain protein with unknown function composed of a CBS domain pair.

Cystathionine-β-synthase (CBS, EC 4.2.1.22) catalyzes a pyridoxal-phosphate (PLP)-dependent β-replacement condensation of homocysteine with serine to form cystathionine. In mammals, this reaction is a point of feedback control in the transulfration pathway, in which cysteine is synthesized from methionine (Banerjee *et al.* 2005, Miles & Kraus 2004, Banerjee *et al.* 2003). In eukaryotes, CBS is allosterically regulated by S-adenosyl-L-methionine (adoMet or SAM), where adoMet is the product of the methionine biosynthetic cycle. AdoMet binds to the regulatory CBS domains found at the C-terminus of cystathionine-β-synthase (Scott *et al.* 2004) and donates methyl groups to a variety of acceptors. It is produced by methylation of L-homocysteine to methionine (Finkelstein 1984a) and controls its own synthesis. Thus low adoMet concentrations result in low CBS activity, directing homocysteine into methionine and adoMet synthesis (Finkelstein *et al.* 1975). In contrast, high adoMet concentrations result in high CBS activity, directing homocysteine into cysteine biosynthesis (Finkelstein *et al.* 1975, Finkelstein *et al.* 1984a, Bukovska *et al.* 1994).

5'-IMP dehydrogenase (IMP dehydrogenase, IMPDH; E.C. 1.2.1.14) catalyzes the rate-limiting step in the *de novo* biosynthetic pathway of GMP, or more specifically, the oxidation of IMP into XMP in a reaction which is NAD^+-dependent (Shu & Nair 2008). IMPDH contains two CBS domains in its C-terminus. Guanine nucleotides are critical cellular metabolites, involved in RNA, DNA and protein synthesis, including various signaling pathways. IMP is a common precursor for both adenylate and guanylate biosynthesis and is a substrate for adenylosuccinate synthetase (Pimkin *et al.* 2009). Inhibition of IMPDH results in a reduction in the amount of guanylate nucleotide synthesis, which among other things, has an antiproliferative effect in mammals (Pankiewicz *et al.* 2003, Pankiewicz *et al.* 2004). Human IMPDH activity is upregulated during increased cell proliferation (Collart & Huberman 1988, Natsumeda *et al.* 1988). It is thus a target for antiviral, antibacterial, anticancer and immunosuppressive therapy (Shu & Nair 2008).

5'-Adenosine monophosphate-activated protein kinase (5'-AMP-activated protein kinase, AMPK; E.C. 2.7.11.1) is a heterotrimeric serine/threonine protein kinase, which regulates cellular energy homeostasis by attaching a phosphate to downstream proteins in the cells of nearly every tissue. It is composed of the α, β, and γ-subunits, of which there are several isoforms in humans (Stapleton *et al.* 1996). The γ-subunit contains four CBS domains and binds adenylate nucleotides in a cooperative manner. AMPK activation results in stimulation of energy-producing and inhibition of energy-consuming pathways and processes due to metabolic stress that inhibits ATP synthesis or accelerates ATP consumption (Hardie & Hawley 2001). Activation results from an increase in the ([ADP]+[AMP])/[ATP]-ratio as commonly occurs when ATP is consumed by cellular processes (Hardie 2007).

Voltage-gated chloride channels and Cl^-/H^+ antiporters, both types being CLCs, are poorly understood ion channels with 9 members expressed in humans (CLCN1-7, Ka, and Kb) (Zifarelli & Pusch 2007). All members have two CBS domains in their C-terminal cytoplasmic domain, separated by long loops. CLCs are involved in a number of physiologically important functions, such as regulation of pH, volume homeostasis, organic solute transport, with effects on cellular migration, proliferation and differentiation (Jentsch 2008). Some are plasma membrane Cl^- channels, which function in transport of Cl^- through epithelial membranes (Uchida 2000), while others have roles in modulating neuronal excitability (Wang *et al.* 2006). Other CLC proteins

are mainly found in the endosomal-lysosomal system where they transport protons or chloride into various membrane-enclosed structures (Hara-Chikuma *et al.* 2005, Jentsch 2008). The CLCs 4, 5 and 6 are vesicular transporters, which carry two chloride ions and one proton through a membrane. The other CLCs only transport chloride (Accardi & Miller 2004, Picollo & Pusch 2005, Zifarelli & Pusch 2007, 2009).

1.2.3 Cystathionine-β-synthase

Examples of soluble CBS domain-containing proteins with two CBS domains, regulated by binding of small-molecule metabolites, and are well-studied include cystathionine-β-synthase and IMPDH. The structure of the truncated (without C-terminal CBS domains) dimeric form of the 551 amino acid human cystathionine-β-synthase (*h*CBS, PDB Id 1M54, Taoka S *et al.* 2002) revealed the composition of the CBS monomer: 11 α-helices, 2 β-sheets of four and six strands, and seven 3_{10} helices. The 140 residue C-terminus contains the regulatory CBS domains. Truncated *h*CBS is dimeric, which is much more active than the wild type but is also unresponsive to AdoMet (Shan *et al*. 1998, Kery *et al*. 1998, Meier *et al.* 2001, Taoka *et al.* 2002). Full-length *h*CBS can form complexes ranging from dimer to 16-mer (Frank *et al.* 2006) but is a functional homotetramer of ~252 kDa (Skovby *et al.* 1984). The 70 amino acid N-terminus coordinates a heme cofactor at the surface of the protein, on opposite ends of dimeric CBS and provides redox regulation (Jhee *et al.* 2000, Meier *et al.* 2001). The PLP is buried in a cleft between the N- and CBS domains bound to a conserved loop about 20 Å away from the hemes. It also forms a Schiff base with K119 forming an internal aldimine in the free enzyme, which plays a key role in catalysis. The active site is accessible to the substrates via a narrow channel (Evande *et al.* 2004, Meier *et al.* 2010, PDB IDs 1JBQ and 1M54). The recently solved full-length structure of *Drosophila melanogaster* CBS as free and in complex with serine and aminoacrylate-bound intermediates sheds further light on catalysis and regulation (Koutmos *et al.* 2010). The structure demonstrated similar conformations of N-terminal and PLP binding domains as observed previously for truncated *h*CBS, while interesting observations were made regarding the placement of the CBS domains. The CBS domains are connected to the catalytic domain by a flexible linker and were observed to facilitate formation of a tightly bound dimer, where the CBS domains cross-dimerize

with the CBS domains from the other subunit and are packed head-to-tail in the crystal (Koutmos *et al.* 2010, PDB IDs 3PC2, 3PC3, and 3PC4).

CBS catalyzes a β-replacement reaction via a ping-pong mechanism, in which the thiol nucleophile of L-homocysteine attacks the hydroxyl group of serine. Pre-steady and steady state kinetic methods have been applied to study the reaction mechanism of yeast CBS (Mudd *et al.* 1965, Suda *et al.* 1973, Aitken *et al.* 2003a, Aitken & Kirsch 2005). Thus, serine and homocysteine compete for binding to the free enzyme, where binding of homocysteine to the serine-E complex also inhibits it (Jhee *et al.* 2000). *h*CBS is not inhibited by either substrate (Kabil *et al.* 1999). The reaction proceeds via several PLP-bound intermediates. First, an external aldimine is formed by reaction of serine and the internal aldimine of the free enzyme. Serine binding results in large conformational changes (Kabil *et al.* 2001). Proton abstraction at the α-carbon yields an aminoacrylate intermediate through ternary complex formation in the presence of homocysteine (Borcsok & Abeles 1982). The abstraction step is not rate-limiting (Jhee *et al.* 2001, Woehl *et al*. 1996). Next, nucleophilic attack on the aminoacrylate and reprotonation at the α-carbon generates the external aldimine of cystathionine. Cystathionine is released from the enzyme after a final transaldimination reaction. The release of cystathione from the full-length enzyme is the rate-limiting step of the reaction and defines k_{cat} (Taoka *et al.* 2002), thought to be caused by slow conformational changes mediated by the C-terminal CBS domains in the product complex (Aitken & Kirsch 2005). In the truncated enzyme from yeast lacking the CBS domains, the rate-limiting step is the reaction of homocysteine with the aminoacrylate intermediate (Jhee *et al.* 2001), where slow conformational changes are thought to be the cause of the slowed reaction rate (Aitken & Kirsch 2003b).

The C-terminal part of cystathionine β-synthase containing the CBS domains thus regulates its activity through steric and allosteric mechanisms. The autoinhibition is alleviated by binding of adoMet or by deletion of the regulatory domain, where k_{cat} is 2.8 s^{-1} for the wild type, 5.2 s^{-1} for the adoMet activated enzyme (activated state), and 10 s^{-1} for the truncated enzyme (superactivated state, Evande *et al.* 2002). hCBS is activated 2.5-5-fold by adoMet with a dissociation constant of 15 µM (Janosik *et al.* 2001). AdoMet, with positively cooperative binding and a K_d of 7.4 µM, activates yeast CBS by increasing k_{cat}, while the K_m's for both substrates remain the same

(Taoka *et al*. 1999). The isolated CBS dimer can bind adoMet with a K_d of 34 μM (Scott *et al.* 2004).

1.2.4 IMPDH

IMPDH is a homotetramer composed of a 400 residue 8-stranded α/β-barrel core and a CBS pair at the C-terminus (Zhang *et al.* 1999, Colby *et al*. 1999). The core domains of the four subunits associate to form a central catalytic site, where the loops connecting the secondary structural elements, especially the catalytic cysteine (α_6-β_6) and a "flap" region (α_8-β_8), contain the catalytically important residues. The four CBS pairs are located at the outside of the core complex.

Mechanistic studies of IMPDH proteins have been ongoing for a long time (Brox & Hampton 1968, Heyde *et al.* 1976, Verham *et al.* 1987, Carr *et al.* 1993, Xiang *et al.* 1996, Wang & Hedstrom 1997, Xiang & Markham 1997, Digits & Hedstrom 1999). IMPDH catalyzes the conversion of IMP into XMP with the reduction of NAD^+ (Pimkin & Markham 2009). IMP binds first to the enzyme and then NAD^+ in a random addition mechanism (Heyde *et al.* 1976, Wang & Hedstrom, 1997). A bi-bi mechanism has also been suggested (Holmes *et al.* 1974, Carr *et al.* 1993, Verham *et al.* 1987, Anderson & Sartorelli 1968, Xiang *et al.* 1996, Nimmesgern *et al.* 1996). A conformational change occurs and Cys331 (IMPDH-II numbering) attacks position 2 of IMP. Transfer of the hydride to NAD^+ then yields a thioimidate (E-XMP*), which is hydrolyzed and a "flap" moves into the NADH site after its release to activate E-XMP* hydrolysis via a bound water molecule (Link & Straub 1996, Huete-Perez *et al.* 1995, Sintchak *et al.* 1996, Pimkin & Markham 2009). XMP release is the rate-limiting step. NAD^+ binding is uncompetitive with respect to E-XMP and E-XMP* (Hedstrom & Wang 1990).

Point mutations in the CBS domain of human IMPDH-I results in the autosomal dominant retinitis pigmentosa type 10 (RP10) and Leber Congenital Amaurosis (LCA), but have no effect on *in vitro* activity (Bowne *et al.* 2002, Kennan *et al.* 2002, Aherne *et al.* 2004, McLean *et al.* 2004, Mortimer *et al.* 2005, Xu *et al.* 2008, Mortimer *et al.* 2008, Gunter *et al.* 2008, Hedstrom *et al.* 2008). Scott *et al.* (2004) showed that the CBS pair of IMPDH-II binds and is activated by ATP *in vitro* in a positively

cooperative manner, while the effect of ATP was abolished by an RP10-causing substitution of human IMPDH-I (R224P) in the second CBS domain of IMPDH-II. This observation was not reproduced in later studies (Mortimer & Hedström 2005, Pimkin & Markham 2008). In the case of human IMPDH-II, CBS domain truncation results in fully active tetrameric enzyme, suggesting that the CBS domains are not involved in oligomerization (Nimmesgern *et al.* 1996).

Human IMPDH-I and -II are significantly different from bacterial and parasitic proteins (Franklin & Cook 1969, Hupe *et al.* 1986, Verham *et al.* 1987). Evidence for independent negative regulation of AMP synthesis by the CBS domains exists. Truncation mutants of *E. coli* IMPDH lacking CBS domains experience large increases in adenylate nucleotide concentrations and depletion of cellular PRPP, which leads to growth arrest (Pimkin *et al.* 2009). IMPDH may bind single-stranded nucleic acids, demonstrated *in vitro* and *in vivo* (McLean *et al.* 2004, Mortimer & Hedström 2005), or may bind DNA and be associate with proteins that are involved in RNA processing and splicing regulation in yeast (Matsuno *et al.* 1995, Cornuel *et al.* 2002). In fact, IMPDH associates with polyribosomes in tissue culture cells through the CBS domains (McLean *et al.* 2004, Mortimer & Hedström 2005, Mortimer *et al.* 2008). IMPDH thus may have a role in RNA metabolism and transcription regulation, mediated by the CBS domains. RP10-linked mutations (Mortimer *et al.* 2008) and the presence of an extra C-terminal domain (Xu *et al.* 2008) were shown to decrease nucleic acid binding *in vitro*. Therefore RP10-linked mutations may impair the RNA-related function of IMPDH and thus have an effect on retinal nucleotide pools (Mortimer *et al.* 2008).

1.2.5 Conclusions

CBS domains thus form a very large and widely distributed domain superfamily, which has been incorporated into a wide array of different architectures and has adapted to diverse functions. One of those functions is to serve as allosteric "internal inhibitors" of the functional domains of a protein. The fact that this inhibition can be relieved (in some cases) by the allosteric binding of adenosine nucleotides or derivatives and subsequent conformational changes makes it a very useful sensory module for cellular energy status. Slow conformational changes might be an important additional level of regulation afforded by CBS domains and are observed for some soluble CBS proteins, *e.g.* cystathionine-β-synthase (Aitken & Kirsch 2005). Regulatory enzymes commonly isomerize rapidly, where a factor which slows down a structural isomerization could enable more robust enzyme-ligand or –substrate interactions. A succession of slow isomerizations could create a timing mechanism allowing gradual changes in activity and may provide a broader spectrum of possible cellular responses (Purich 2010). Additionally, CBS domains seem to mediate oligomerization in some soluble enzymes, which affords an increase in catalytic efficiency and in turn allows an additional level of control.

Several unanswered questions remain: 1. What are the molecular mechanisms and structural determinants responsible for regulation and communication between the CBS domains and the functional core of the protein? 2. Can molecules be developed to target CBS domains to either upregulate or downregulate the function of the protein in which they reside, e.g. antidiabetic, anticancer, anti-obesity drugs, immunosuppresive therapy (Ignoul & Eggermont 2005).

1.3 Family II inorganic pyrophosphatase

1.3.1 Inorganic polyphosphates

Inorganic polyphosphates (PolyP) are composed of a succession of P_i units linked by a P-O bond. The functions of polyP, which has been found in all organisms so far tested (Brown & Kornberg 2004, Brown & Kornberg 2008, Docampo *et al.* 2005, Hooley *et al.* 2008, Kornberg *et al.* 1999, Kulaev *et al.* 2004, Schröder & Muller 1999)

include: substitute for ATP, metal chelation, phosphate reservoir, buffer, stress response (Kornberg 1995, Kulaev *et al.* 2004). PolyP also seems to be essential for the physiological adjustment of microbial growth to environmental conditions, and the virulence of some pathogens (Rao *et al.* 2009), while also stimulating mTOR, a protein kinase involved in mammalian cell proliferation (Wang *et al.* 2003b).

PolyP is mainly broken down by exo- (PPX) and endopolyphosphatases. Exopolyphosphatases degrade long polyphosphate chains processively from either terminus. There are two nonhomologous families of PPXs, the prototypes of each being the *Escherichia coli* type (*ec*PPX) (Akiyama *et al.* 1993) and the yeast type (*sc*PPX) (Wurst & Kornberg 1994). Yeast-type PPX is only found in eukaryotes, while *E. coli*-type PPX is found in bacteria and archaea. Yeast-type, or eukaryotic PPX, belongs to the DHH family of phosphoesterases, which includes family II PPases and RecJ exonucleases, among others (Aravind & Koonin,1998). *sc*PPX and human PPX (*h*PPX), also called *h*prune (Prunes are PPXs of multicellular animals; Tammenkoski *et al.* 2008), preferentially utilize tripolyphosphate (P_3) as substrate, but *sc*PPX is able to processively hydrolyze longer chain polyP. Both *sc*PPX and *h*prune display negligible activity against pyrophosphate, which actually inhibits PPX activity (Tammenkoski *et al.* 2007, 2008). Family II PPase, on the other hand, has negligible activity against tripolyphosphate and longer-chain polyP but displays maximal activity against pyrophosphate (Parfenyev *et al.* 2001). Both scPPX and family II PPase are very similar in structure, composed of two domains linked together by a flexible loop region with the active site in the domain interface interface (Tammenkoski *et al.* 2007, Ugochukwu *et al.* 2007). So far, no enzyme has been shown to hydrolyze both pyrophosphate and tripolyphosphate (or longer phosphate chains) directly into phosphate at similar rates, although a low PPase activity has been found in a PPX from *Corynebacterium glutamicum* (Lindner *et al.* 2009).

1.3.2 Family II pyrophosphatase

Inorganic pyrophosphatase (EC 3.6.1.1; subsequently referred to as PPase) catalyzes the synthesis and hydrolysis of pyrophosphate (PP_i), producing phosphate (P_i) and energy in the process. PP_i is the shortest linear inorganic polyP produced by 5'-triphosphate-dependent reactions, such as protein synthesis, RNA and DNA

synthesis, lipid metabolism, polysaccharide and nucleotide synthesis (Klemme 1976), in addition to basic metabolic reactions, like photophosphorylation (Baltscheffsky *et al.* 1966), oxidative phosphorylation (Baltrop *et al.* 1963) and glycolysis (Mansurova *et al.* 1976). The hydrolysis of PP_i pulls the equilibrium of these reactions forward (ΔG^0 = -17 kcal/mol *in vivo*; Cooperman 1982, Kornberg 1962, Peller 1976, Nelson *et al.* 2000, Heinonen 2001). Even so, the concentration of PP_i in e.g. rat blood and liver cells remains above the equilibrium value (Veech *et al.* 1980). PP_i has been shown to regulate calcification, cell proliferation and iron transport, thus also being linked to certain diseases (Heinonen 2001). Under specific conditions, ATP-dependent metabolic reactions may switch to use PP_i as an energy source (Huang *et al.* 2008), and further, certain classes of prokaryotes utilize PP_i instead of ATP as a central energy source in metabolism (e.g. Bielen *et al.* 2010). PP_i thus affects the fidelity and efficiency of RNA-, DNA-, and protein synthesis (Lahti 1983). PPases have been shown to be essential for the growth of a number of bacteria (Chen *et al.* 1990, Abu Kwaik 1998, Ogasawara 2000).

Three families of non-homologous PPases have been identified so far. The soluble family I enzymes, discovered in 1928 by Kay *et al.*, are the most well studied and are found in all kingdoms of life (Baykov *et al.* 1999). The enzymes from *E. coli* (*ec*PPase) and *Saccharomyces cerevisae* (*sc*PPase) are the most well-known members of this family. The soluble family II enzymes are restricted to some prokaryotic species, mostly found in the phylums firmicutes and proteobacteria. Variously called the Bs-family (Shintani *et al.* 1998) or the C Class (Young *et al.* 1998) of soluble PPases, family II PPase was initially characterized in *Bacillus subtilis* by Tono and Kornberg in 1967, who did not recognize that it belongs to an evolutionarily unrelated family. Soluble family I and II PPases have different primary and tertiary structures, catalytic efficiencies and cofactor selectivities (Shintani *et al.* 1998, Young *et al.* 1998, Merckel *et al.* 2001, Ahn *et al.* 2001, Parfenyev *et al.* 2001, Halonen *et al.* 2005). Although different in overall structure, the active site residues essential for catalysis are conserved between families (Ahn *et al.* 2001, Merckel *et al.* 2001). The third type of PP_i hydrolyzing enzyme is the membrane-bound PPase, found in plants and some bacteria, which uses the energy released upon pyrophosphate hydrolysis to drive proton or sodium transport across various cellular membranes (Serrano *et al.* 2007).

Soluble PPases are additionally regulated by transcriptional and post-translational mechanisms. Family I PPases are overexpressed in human cancer tissues (Chen *et al.* 2002, Lexander *et al.* 2005), upon decrease in phosphate concentration in certain bacteria (Gomez-Garcia *et al.* 2003), in the cold in wheat (Gulick *et al.* 2005), at increased salinity in barley (Ueda *et al.* 2006), and upon calorie restriction in cows (Kuhla *et al.* 2007). Phosphorylation has been shown to prevent self-fertilization in the field poppy flower (de Graaf *et al.* 2006). Family II PPases are also phosphorylated in *Streptococcus agalactiae* (Rajagopal *et al.* 2003). Interestingly, about a quarter of family II PPases contain a regulatory insert in the N-terminal DHH domain (Paper I). The insert is composed of two CBS domains and a DRTGG domain, and thus these enzymes are called CBS-PPases. CBS-PPases are inhibited by AMP and ADP, while ATP and AP_nA (*n* = the number of phosphate residues) activate several-fold (Papers I, V).

The following is intended to provide the reader an overview of the mechanism of family II pyrophosphatases, which will allow a better understanding of the properties of CBS-PPases and their regulation by CBS domains.

1.3.3 Structure and mechanism

Several structures of family II PPases have been solved (table 2). As shown in figure 3, family II PPases are composed of two domains, termed the DHH and DHHA2 domains. The active site is located in the interface of these domains. The N-terminal DHH domain binds two metal ions, that are crucial for catalysis. The C-terminal DHHA2 domain binds substrate. Movement of the C-terminal domain over the N-terminal domain creates a catalytically competent structure (Ahn *et al.* 2001, Merkel *et al.* 2001). A dimeric structure is crucial for activity (Parfenyev *et al.* 2001).

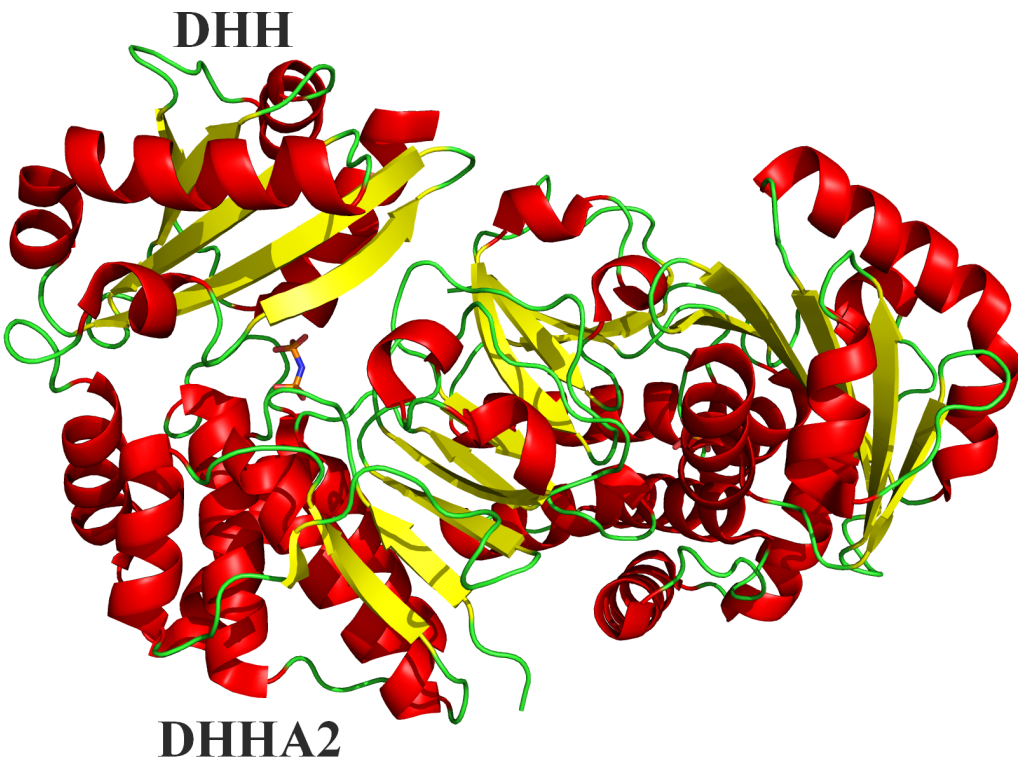


Figure 3. Cartoon representation of the structure of the family II PPase of *Streptococcus gordonii* (*sg*PPase) determined at 1.75 Å resolution with imidodiphosphate (PNP), F^- and Mg^{2+} bound to the active site located between the DHH and DHHA2 domains (PDB ID 2HAW, Fabrichniy *et al.* 2007).

Table 2. Structures of family II PPases in the PDB.

Organism	Full-length or fragment/Ligand	PDB ID	Resolution (Å)	Reference
Streptococcus gordonii	Full/Mn^{2+}, SO_4^{2-}	1K20	1.50	Ahn *et al.* 2001
Bacillus subtilis	Full/Mn^{2+}	1K23	3.00	Ahn *et al.* 2001
Bacillus subtilis	Full/SO_4^{2-}	1WPM	2.05	Fabrichniy *et al.* 2004
Bacillus subtilis	N-term/Mn^{2+}, SO_4^{2-}	1WPN	1.30	Fabrichniy *et al.* 2004
Streptococcus gordonii	Full/PNP, Zn^{2+}, SO_4^{2-}	1WPP	2.05	Fabrichniy *et al.* 2004
Methanococcus jannaschii	Full/Mn^{2+}	2EB0	2.20	Unpublished
Streptococcus agalacticae	Full/PNP, Mg^{2+}, Mn^{2+}	2ENX	2.80	Rantanen *et al.* 2007
Bacillus subtilis	Full/PNP, Mg^{2+}, F^-, SO_4^{2-}	2HAW	1.75	Fabrichniy *et al.* 2007
Bacillus subtilis	Full/PNP, Mg^{2+}, Mn^{2+}, Fe^{2+}, SO_4^{2-}	2IW4	2.15	Fabrichniy *et al.* 2007
Clostridium perfringens	Regulatory domain/AP_4A	3L2B	2.30	Paper III
Clostridium perfringens	Regulatory domain/AMP	3L31	2.27	Paper III

The catalytic mechanism of family II enzymes (Parfenyev *et al.* 2001, Konopka *et al.* 2002, Zyryanov *et al.* 2002, Zyryanov *et al.* 2004a, 2004b, Fabrichniy *et al.* 2004, Halonen *et al.* 2005, Fabrichniy *et al.* 2007) are not as well-known as those of family I PPases (Heikinheimo *et al.* 1996, Baykov *et al.* 2000, Belogurov *et al.* 2000, Heikinheimo *et al.* 2001, Pohjanjoki *et al.* 2001, Halonen *et al.* 2002, Oksanen *et al.* 2007). Briefly, the mechanism is the following: Metal-activated PPase binds a metal-bound substrate ($MgPP_i$ or Mg_2PP_i, Cooperman *et al.* 1992), resulting in the breakdown of the P-O bond between the two phosphates of pyrophosphate. This occurs *via* nucleophilic attack of a transient hydroxide formed by abstraction of a proton from a water molecule activated by a two-metal cluster (Heikinheimo *et al.* 1996, 2001). The two product phosphates then leave the enzyme-product complex in a stepwise manner.

1.3.3.1 Metal binding

PPases are phosphohydrolases dependent on the presence of divalent metal cations in the active site for catalytic activity. As shown in figure 4, two metal ions bind to the N-terminal DHH domain of family II PPases to form a two-metal center, while the third and fourth metal ions bind with substrate (Merckel *et al.* 2001, Parfenyev *et al.* 2001, Fabrichniy *et al.* 2007). A water molecule is coordinated between the two metal ions (Ahn *et al.* 2001, Merckel *et al.* 2001). The metal cofactor of choice of family II enzymes is Mn^{2+} (Kuhn *et al.* 1998, Shintani *et al.* 1998, Young *et al.* 1998, Parfenyev *et al.* 2001, Halonen *et al.* 2005). Other divalent cations provide for lower catalytic efficiency (Zyryanov *et al.* 2004, Halonen *et al.* 2005). Family II PPases have ~20-fold higher activity with Mn^{2+} compared to the Mg^{2+} form of family I PPases (Parfenyev *et al.* 2001, Zyryanov *et al.* 2004b). However, the presence of Mg^{2+} is required for maximal activity (Parfenyev *et al.* 2001).

The metal binding sites M1 and M2 have similar structures in the family II PPases of *Streptococcus gordonii* (*sg*PPase), *Streptococcus mutans* (*sm*PPase) and *Bacillus subtilis* (*bs*PPase) (Merckel *et al.* 2001, Ahn *et al.* 2001, Fabrichniy *et al.* 2007). Sites M1 and M2 are highly integrated into the active site, and are formed by residues H9, D13, D75, H76 and D15, D75, H97, D149, respectively. When site M2 binds a metal, H97 moves away from contact with D75 and forms a trigonal bipyramidal site. Substrate hydrolysis apparently requires octahedral conformations of both sites, which can only form if site M2 is occupied by a metal (such as Mn^{2+} or Co^{2+}), which supports the adoption of this configuration (Fabrichniy *et al.* 2004). Mg^{2+} and Zn^{2+} typically fail to induce this conformation (Harding 2001).

The metal bound to site M2 is a determinant of quaternary structure and catalytic properties of family II PPases. In the absence of an activating metal in this site, family II PPases dissociate into monomers, and both monomer and dimer forms are inactive (Parfenyev *et al.* 2001). Co^{2+} and Mn^{2+} are able to support PP_i hydrolysis (Zyryanov *et al.* 2004b), where site M2 binds Mn^{2+} or Co^{2+} with nM affinity and has μM affinity for Mg^{2+} (Parfenyev *et al.* 2001, Zyryanov *et al.* 2004b). Mn^{2+}-activated family II *sg*PPase has a 15-fold higher k_{cat} than Mg^{2+}-activated family I *sc*PPase (Fabrichniy *et al.* 2004), while the k_{cat} for Mg^{2+}-activated *sg*PPase is the same as for *sc*PPase (Zyryanov *et al.* 2004b). Consequently, PP_i and P_i have lower affinity for Mn^{2+}-activated *sg*PPase,

while PP_i hydrolysis and P_i release are accelerated in comparison to *sc*PPase (Fabrichniy *et al.* 2004). Mg^{2+}-activated *sm*PPase has higher affinity for PP_i compared to Mn^{2+} (Zyryanov *et al.* 2002). However, Mn^{2+} and Mg^{2+} bind (Parfenyev *et al.* 2001) and activate *sg*PPase equally when bound to sites M1, M3 and M4 (Fabrichniy *et al.* 2004). Site M1 may have increased affinity for transition metals when substrate or product is bound to the active site (Fabrichniy *et al.* 2007). Substitution in any of the metal binding sites decreases k_{cat} and metal binding affinity by $>10^{5\text{-}6}$-fold (Halonen *et al.* 2005).

The histidines in the family II PPase active site are good ligands especially for Zn^{2+}, which binds with pM affinity to *sg*PPase and *bs*PPase, but fails to activate PP_i hydrolysis (Kuhn *et al.* 1998, Parfenyev *et al.* 2001, Fabrichniy *et al.* 2004, Zyryanov *et al.* 2004a, 2004b). Zn^{2+} bound to site M1 supports PP_i hydrolysis when Mn^{2+} or Mg^{2+} are bound to site M2, but slowly displaces them (Zyryanov *et al.* 2004a). Otherwise, Zn^{2+} is a poor activator of family II PPase (Zyryanov *et al.* 2004a, Fabrichniy *et al.* 2004). Mg^{2+} also displaces the Mn^{2+} bound to site M2, if not present in the reaction (Parfenyev *et al.* 2001). PP_i hydrolysis by *Methanococcus jannaschii* PPase is supported by Zn^{2+}, but inhibited when Co^{2+} is present (Kuhn *et al.* 2000).

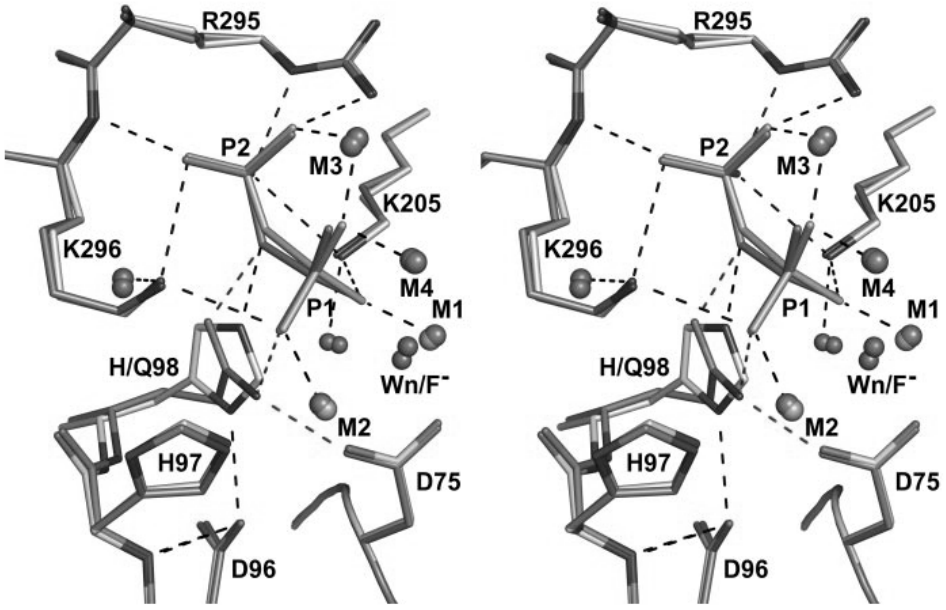


Figure 4. Stereoview of the superposition of the active sites of wild type and the H98Q variant of *bs*PPase in complex with the substrate analog imidodiphosphate (PNP) (adapted from Fabrichniy *et al.* 2007). The PNPs are shown as sticks, while metals, water and fluoride are shown as spheres. Hydrogen and coordination bonds to the P_is and metals are shown as broken lines.

1.3.3.2 PP_i binding and hydrolysis

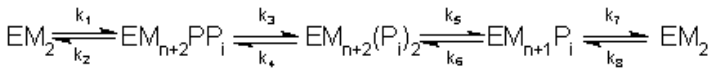
As shown in figure 4, substrate binds directly to a conserved section of the active site within the DHHA2 domain (Konopka *et al.* 2002, Ahn *et al.* 2001, Merckel *et al.* 2001), which induces a conformational change to the "closed" form of *sg*PPase (Ahn *et al.* 2001). This results in the movement of the N- over the C-terminal domain, and aligns residues in the active site for catalysis (Merckel *et al.* 2001, Ahn *et al.* 2001).

The P_i binding sites in the active site are designated P1 and P2. PP_i is bound first to site P2, resulting in the closure of the structure (Shizawa *et al.* 2001, Ahn *et al.* 2001), and requires a change in the coordination number of site M2 from 5- to 6-coordinated (Fabrichniy *et al.* 2004). Site P1 is indeed fully formed in the *bs*PPase. Mn_2 structure (Ahn *et al.* 2001). D203, R295 and K296 bind PP_i in *bs*PPase, the latter two residues being part of the SRKKQ motif in the DHHA2 domain (Ahn *et al.* 2001, Merckel *et al.* 2001, Shizawa *et al.* 2001, Parfenyev *et al.* 2001, Konopka *et al.* 2002, Fabrichniy *et*

al. 2007). R295 binds phosphate in site P2, while P1 interacts with the metals in sites M1 and M2. K205, the metal in site M3 and a water molecule interact with both sites P1 and P2, while D203 positions R295 for catalysis (Merckel *et al.* 2001, Fabrichniy *et al.* 2007). Thus, substitution of any of these residues increases K_m (Halonen *et al.* 2005). The DHHA2 domain restricts the active site for specific hydrolysis of PP_i instead of other inorganic or organic polyphosphates, as partial deletion of this domain increases activity against P_3 (Konopka *et al.* 2002). Family II PPases are thus unable to hydrolyze other substrates than PP_i and are not inhibited by polyphosphate (Zyryanov *et al.* 2002), but are weakly inhibited by fluoride (Kuhn *et al.* 2000, Fabrichniy *et al.* 2007).

Scheme 1 shows the reaction scheme of family II PPase catalyzed hydrolysis of PP_i (Zyryanov *et al.* 2004b). PPases transfer a phosphate from PP_i to water in a single step without formation of a covalent enzyme-bound intermediate (Gonzalez *et al.* 1984). PPases also catalyze PP_i synthesis (Baykov & Sheshtakov 1992) and P_i-H_2O oxygen exchange (Springs *et al.* 1981).

Scheme 1: PP_i-P_i equilibration by family II PPase (Zyryanov *et al.* 2004b)[a]



[a] E = enzyme, M = Mn^{2+} or Mg^{2+}, PP = PP_i, P = P_i, n = 1 or 2. PP_i binds as the Mg^{2+}-complex. The reaction involves PP_i binding (k_1-k_2), hydrolysis (k_3-k_4), and stepwise release of two P_is (k_5-k_6 and k_7-k_8).

A single E-PP_i intermediate has been observed in the catalytic cycle of family II PPase (Zyryanov *et al.* 2004b). Ahn *et al.* (2001) proposed that D98 is the general acid in the reaction, which protonates the leaving group phosphate in site P1. Even though substitution of this residue decreases k_{cat}, it has no effect on the pK_a of the general acid (Halonen *et al.* 2005). The DHH motif, however, has a critical role in catalysis, whereby D96 positions H97 and H98 for optimal interaction with PP_i and other residues, such as the metal binding D76 (Ahn *et al.* 2001, Fabrichniy *et al.* 2007). The rate of conversion of EPP_i into $E(P_i)_2$ for Mn^{2+}-activated *sg*PPase is 17000 s^{-1} (k_3), while the rate of the reverse reaction is 130 s^{-1} (k_4). When Mn^{2+} is replaced by Mg^{2+}, k_3 and k_4 are significantly lower by 5.2 and 3.7-fold, respectively (Zyryanov *et al.* 2004b).

1.3.3.3 P_i release

The structures of product complexes show that the C-terminal domain is closed over the N-terminal metal-binding site (Ahn *et al.* 2001, Merckel *et al.* 2001, Fabrichniy *et al.* 2004). P_i release in family II PPases follows concurrent structural changes from a "closed" catalytic to a less active "open" conformation. The rate of product release is governed by factors which affect these structural changes causing the opening of the interface between the DHH and DHHA2 domains (Ahn *et al.* 2001). Family II PPases release metal-bound P_i in sequential steps. The structures with sulfate bound to both P1 and P2 in complex with Mn^{2+}, Fe^{2+} (Fabrichniy *et al.* 2004) or Zn^{2+} (Merckel *et al.* 2001) show an opening and twisting of the C-terminus, where K205 moves out of the active site, facilitating product release from site P2. However, site P1 releases product first, as the first P_i released contains oxygen from the catalytic water molecule (Zyryanov *et al.* 2004b), and the *bs*PPase. S_i structure is in the closed conformation (Fabrichniy *et al.* 2004), while the *bs*PPase. Mn_2 structure has a fully formed P1 site (Ahn *et al.* 2001). The P_i from site P1 thus moves out of the active site through a channel formed by K205, Q80 and Q81 (Fabrichniy *et al.* 2004). The presence of Mn^{2+} promotes a 10-fold increase in the rate of P_i release in comparison to Mg^{2+} (Zyryanov *et al.* 2004b), while electrostatic repulsion between the P_is may also contribute to product release (Ahn *et al.* 2001).

2. AIMS

After the discovery and initial characterization of CBS-PPases, the research described in this thesis was undertaken with the primary goal being the expression, purification, and detailed kinetic characterization of this novel family of inorganic pyrophosphatases with an insertion composed of two CBS domains and a DRTGG in its N-terminal DHH domain. The primary hypothesis was that due to the presence of the adenosine nucleotide binding CBS domains, this subfamily would be regulated by adenosine nucleotides but otherwise be similar to other family II PPases lacking the N-terminal insert. This hypothesis expanded to characterization of a unique family of enzymes able to hydrolyze both tripolyphosphate and pyrophosphate at essentially the same rates. Moreover, we wanted to uncover the mechanistic details of how CBS domains regulate CBS-pyrophosphatase function and properties in relation to the previously characterized family II PPases lacking the additional domains, and to elucidate the structural determinants of ligand binding and regulation, including the effects of human hereditary disease-associated single residue substitutions in CBS domains. The 3D structure of CBS domains is highly conserved, thus the goal was to be able to apply the information gained to other proteins containing CBS domains and to utilize the uncovered mechanistic details to elucidate the consequences of point mutations in CBS domains in the generation of human hereditary diseases. In short, this work has broad significance related to fundamental mechanisms of CBS domain regulation of the function of CBS domain-containing proteins and to human pathology.

3. EXPERIMENTAL PROCEDURES

More detailed information on Materials and Methods used during this PhD work can be found in the section ORIGINAL PUBLICATIONS I-V. The following is a brief summary of these sections.

3.1. Molecular biology, expression and purification

The genes for the enzymes studied in this work were cloned into the pBluescript KS(-) or pUC19 cloning vectors using overlapping or inverse PCR (Stemmer & Morris 1992), transformed into XL-1 Blue or TOP10 cells, extracted and subsequently either subjected to mutagenesis using the QuickChange method, a 4-primer mutagenesis method and/or directly transferred to a pET15, 22, 23, 32 or 36 expression plasmid. This plasmid was then transformed into either BL21(DE3)ril, C41 or C43 cells (Belogurov *et al.* 2002) and grown for 1–4 hours followed by 2–6 hours of induction of protein expression using 0.4–1 mM IPTG. Cells were disrupted either by sonication or by passage through a French press chamber. Protein purification was performed by conventional anion-exchange and gel filtration in either Tris or MOPS buffers mainly at pH 7.2 including $MgCl_2$ and/or other metals, or by using Talon-Co^{2+} IMAC or HisTrap for His-tagged proteins in the same buffers. CBS-PPase could be well separated from the endogenous *E. coli* PPase. To estimate molecular masses, proteins were fractionated by SDS-PAGE and protein was visualized by staining with Coomassie Brilliant Blue. Protein quantitation was performed by measuring the absorbance at 280 nm, by the Bradford assay, by direct vacuum dehydration and weighing, and by using the NanoOrange protein quantitation kit.

3.2. Activity assays

Measurements of the rate of phosphate production of PPases were performed with an in-line phosphate analyzer utilizing change in absorption caused by phosphate binding to molybdate in acidic medium (Baykov & Avaeva, 1981) or the BIOMOL green phosphate assay system (BioMol Corp.). The concentrations of each component were determined after consideration of the dissociation constants of the Mg^{2+}, Na^+, K^+, and H^+ complexes of PP_i (Baykov *et al.* 1993). P_i release was additionally assayed

using an N-[2-(1-maleimidyl)ethyl]-7-(diethylamino)coumarin-3-carboxamide (MDCC)-labeled *E. coli* periplasmic phosphate binding protein (Brune *et al.* 1994) with an SX18.MV stopped flow reaction analyzer (Applied Photophysics, UK) or a KinTek RQF-3 quenched flow apparatus with $^{32}P_i$.

3.3. Binding assays

Equilibrium binding reactions were performed with ^{14}C-AMP and single-use DispoEquilibrium Dialyzers, a membrane filtration method utilizing 96-well filter-bottom plates or fluorescence titration by monitoring protein Trp or Mant-nucleotide fluorescence.

CBS domain nucleotide binding kinetics was characterized using fluorescently (Mant-) labeled nucleotides and an SX18.MV stopped flow reaction analyzer (Applied Photophysics, UK) with excitation at 295 nm or 355 nm and monitoring emission through a Schott GG filter (>400 nm).

3.4. Phylogenetics

Homology models were built using Modeller and the solved crystal structures of CBS domains from the PDB (mainly those of AMPK, IMPDH and CLC-5 in complex with different ligands, PDB-ID ZZZ, 1ZFJ, 2JA3). Multiple and pairwise structural alignments were performed using several online servers, but mainly the RAPIDO alignment software. Multiple alignments were generally built with the program Mafft v.6.3.

Phylogenetic trees were made using the programs in the Phylip package (SeqBoot, ProtDist, Neighbor/Fitch/Kitsch, Consense), RAxML and PhyML, as well as, MrBayes. Trees were visualized mainly using Dendroscope, and iTol was used for publication purposes.

3.5. Data analysis

Initial rates were obtained from linear fits to the timecourses of phosphate production (Papers I, III-V). When the timecourse was nonlinear, the initial and final

rates, as well as, the rate of transition between them was obtained by fitting the timecourse to a differential equation describing the transition and associated parameters. The latter equations generally take into account a two-state transition between populations of inactive (or partially active) and active enzyme with depletion of substrate during the assay (Papers II, IV). Rates as a function of substrate or ligand concentration were then fit to the Michaelis-Menten or a similar hyperbolic equation, which provided k_{cat}, K_m, and K_d values (Papers I-IV).

Transient kinetic timecourses were fit to a sum of exponentials, where the rates and amplitudes were further analyzed by fitting to simple hyperbolae (Papers II and IV). Fluorescence titration binding profiles were fit to differential equations describing the transitions from nucleotide-free enzyme to different nucleotide bound states taking into account the mass balance between different species in the reaction (Papers II, and VI-V).

The reported kinetic schemes are minimal, so that addition of a species to the models did not significantly improve the fit in terms of the sum of square of residuals, whereas elimination of a species did not make the fit significantly worse.

4. RESULTS AND DISCUSSION

4.1 Discovery of CBS-PPases (Papers I, V)

4.1.1 Novel subfamily of family II PPases

A recent flood of sequenced genomes prompted a re-evaluation of the complement of enzymes hydrolysing linear inorganic polyphosphates. A number of novel multidomain fusion sequences with significant similarity to family II PPases were observed in the genomes of certain Archaea and Eubacteria (table 3). As described in figure 5, these sequences are composed of a DHH and DHHA2 domain, as well as, two CBS domains. Accordingly, the novel enzymes were named CBS-PPases, referring to the ~120-250 amino acid insertion in the N-terminal DHH domain. Most of these sequences additionally contain a DRTGG domain. This domain is named for the conserved amino acids found in its sequence, is also found in certain phosphotransacetylases, but has an as yet unknown function (Papers I, V).

As shown in table 3, both types of family II PPases are found in the Eubacteria and Archaea, but not in eukaryotes. An exception is the intestinal protozoan parasite *Giardia lamblia*, which causes giardiasis. Interestingly, this early-diverging eukaryotic diplomonad, which reproduces by binary fission, has two nuclei with very similar gene content. Two encoded CBS-PPases are the only pyrophosphatases found in this organism and are very different from other CBS-PPases and thus provide an interesting future point of study regarding the evolution of CBS-PPases and of eukaryotes (Morrison *et al.* 2007).

Family I PPases are much more widespread, occurring in all three kingdoms of life. Most family II PPases are found in the Firmicutes in completely sequenced genomes, with more than two thirds of these in the Bacilli and a fifth in the Clostridia. Of the total 421 family II PPases, a quarter (101/421) are CBS-PPases, distributed predominantly among the Firmicutes (54/101) and (16/101) proteobacteria. The majority of these contain both CBS domains and a DRTGG domain, while interestingly, the enzymes from *Moorella thermoacetica*, *Synthrophomonas wolfeii*,

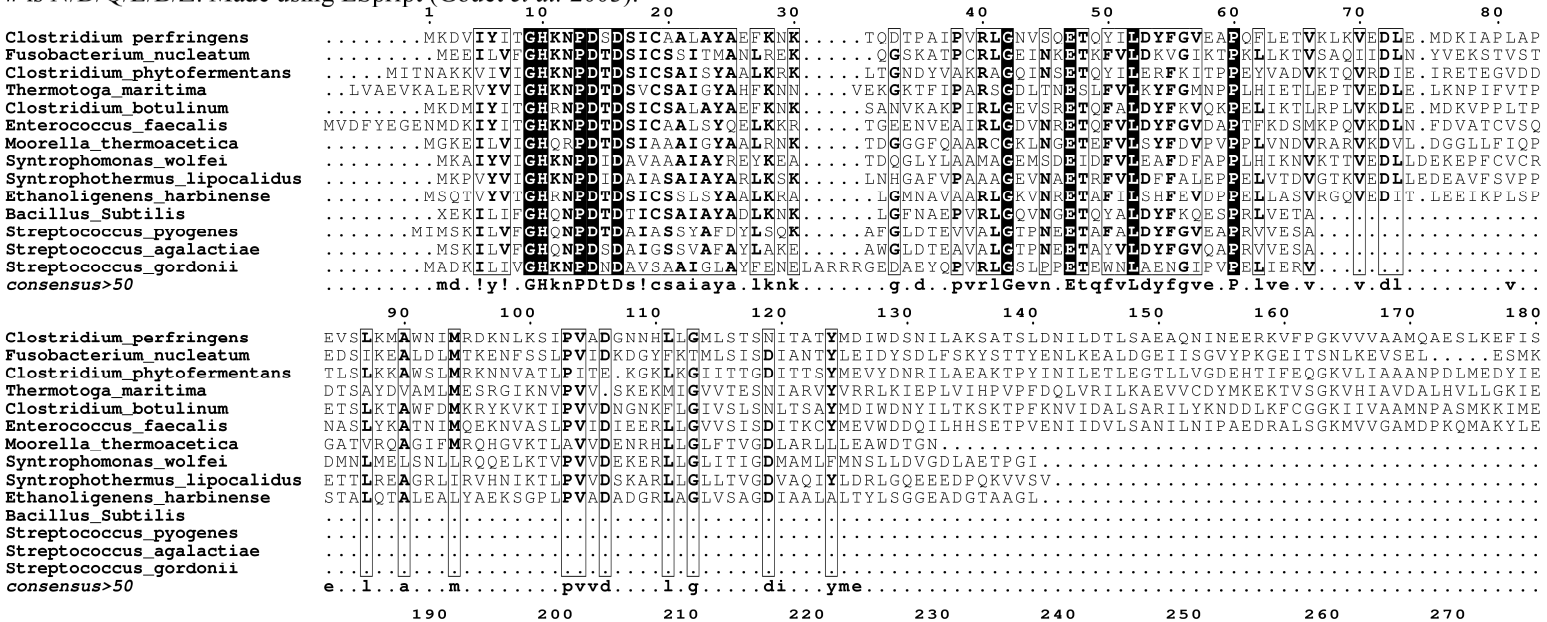
Syntrophothermus lipocalidus, *Eggerthella lenta* and *Ethanoligenens harbinense* (Paper V) lack the DRTGG domain but have retained the two CBS domains (Paper I).

Table 3. Distribution of family II PPase in completely sequenced microbial genomes from the NR database (www.ncbi.nlm.nih.gov). The presence of CBS-PPase is indicated by a star (*). The number of species containing family II PPases is shown in parentheses.

Archaea (24)		**Eubacteria**		**Eukaryotes**	
Crenarchaeota	-	Actinobacteria (7)	+	Alveolata	-
Euryarchaeota (24)*	+	Aquificae	-	Amoebozoa	-
Korarchaeota	-	Bacteroidetes/Chlorobi	-	Apusozoa	-
Nanoarchaeota	-	Chlamydiae/Verrucomicrobia (3)*	+	Centroheliozoa	-
		Chloroflexi	-	Cryptophyta	-
		Chrysiogenetes (1)	+	Diplomonadida*	+
		Cyanobacteria (1)	+	Euglenozoa	-
		Deferribacteres (3)	+	Fungi/Metazoa	-
		Deinococcus-Thermus (4)	+	Glaucocystophyceae	-
		Dictyoglomi	-	Haptophyceae	-
		Fibrobacteres/Acidobacteria (1)	+	Heterolobosea	-
		Firmicutes (251)*	+	Jakobida	-
		Fusobacteria (8)*	+	Katablepharidophyta	-
		Gemmatimonadetes	-	Malawimonadidae	-
		Nitrospirae (1)	+	Nucleariidae	-
		Planctomycetes (1)*	+	Oxymonadida	-
		Proteobacteria (109)*	+	Parabasalidea	-
		Spirochaetes (1)	+	Rhizaria	-
		Synergistetes (1)	+	Rhodophyta	-
		Tenericutes	-	Stramenopiles	-
		Thermodesulfobacteria	-	Viridiplantae	-
		Thermotogales (4)*	+		

Figure 5 shows a sequence alignment of selected family II PPases. The polar active site residues present in family II PPases (Merckel *et al.* 2001, Ahn *et al.* 2001) are conserved in the CBS-PPases identified to date. The regulatory region possesses 27.9% identical residues overall with similar identities in individual domains. The insert is thus not well conserved in sequence among CBS-PPases. However, some highly conserved residues can be found (occurring in >90% of sequences): D72 in CBS1, G113 and L147 in the DRTGG domain, and G292 in CBS2 (*cp*CBS-PPase numbering).

Figure 5. Sequence alignment of selected family II PPases. The locations of the domains found in the alignment are: DHH domain 1-66, CBS1 67-132, DRTGG 133-240, CBS2 241-300, DHHA2 301-. Identical residues are in white on black background, while >50% identical are in black on white background. Consensus characters in uppercase are identical, in lowercase identity is >50%. **!** indicates residues I/V, **$** is L/M, **%** is F/Y, and # is N/D/Q/E/B/Z. Made using ESpript (Gouet *et al.* 2003).



Clostridium perfringens
Fusobacterium_nucleatum
Clostridium_phytofermentans
Thermotoga_maritima
Clostridium_botulinum
Enterococcus_faecalis
Moorella_thermoacetica
Syntrophomonas_wolfei
Syntrophothermus_lipocalidus
Ethanoligenens_harbinense
Bacillus_Subtilis
Streptococcus_pyogenes
Streptococcus_agalactiae
Streptococcus_gordonii
consensus>50

4.1.2 DRTGG domains enable P_3 hydrolysis by CBS-PPase

CBS-PPases were purified to homogeneity (≥98% purity) in the presence of 0.1 mM Co^{2+} and 5-20 mM Mg^{2+} using a combination of anion-exchange and size-exclusion chromatography. Initial characterization of the CBS-PPases from *Moorella thermoacetica* (*mt*CBS-PPase) (Paper I), *Clostridium perfringens* (*cp*CBS-PPase), and *Fusobacterium nucleatum* (*fn*CBS-PPase) (Paper V) at pH 7.2, 0.1 mM Co^{2+}, 5 mM (*mt*CBS-PPase) or 20 mM Mg^{2+} (the rest) revealed that CBS-PPases are 100-1000-fold less active against PP_i (Paper I, V) than the previously characterized family II PPases (Parfenyev *et al*. 2001). Maximal activity was displayed at pH 7.2 for both enzymes (Paper I, V). *mt*CBS-PPase had a k_{cat} of 2.2 ± 0.2 s^{-1} with a K_m of 8 ± 3 μM (Paper II), while *cp*CBS-PPase had a k_{cat} of 50 ± 3 s^{-1} and K_m of 4.2 ± 0.3 μM (Paper V). Assaying *cp*CBS-PPase against a range of inorganic and organic polyP substrates (PP_i, P_3, $polyP_{25\text{-}1000}$, ATP, AP_4, AP_4A) for hydrolytic activity revealed that only PP_i or P_3 as substrate resulted in P_i production. In fact, the enzymes from *Clostridium perfringens* and *Fusobacterium nucleatum* (unpublished results) hydrolyze saturating concentrations of P_3 at rates similar to PP_i in contrast to any other known enzyme hydrolysing linear inorganic polyphosphates. *cp*CBS-PPase had a k_{cat} for P_3 of 55 ± 15 s^{-1} and a K_m of 2.3 ± 0.9 μM. In contrast, *mt*CBS-PPase, which lacks the DRTGG domain, possessed a weaker activity against P_3 (<0.1 s^{-1}). Exhaustive hydrolysis of P_3 resulted in 1.5 times more P_i compared to PP_i. This result confirmed that CBS-PPases are uniquely specialized short-chain inorganic polyPases (Paper V). Other family II PPases lacking the CBS-DRTGG insert possess an even weaker activity of 1/50,000 against P_3 in relation to PP_i hydrolytic activity in optimal catalytic conditions (Parfenyev *et al.* 2001, Zyryanov *et al.* 2002). These results lead to the conclusion that the DRTGG domain somehow enables P_3 hydrolysis by CBS-PPase. Perhaps this is due to structural or kinetic stabilization of protein-protein interactions or conformational dynamics, as the CBS and catalytic domains are highly flexible.

4.2 Ligand binding and regulation (Papers I-V)

4.2.1 Co^{2+} is the best transition metal cofactor of CBS-PPase

As metal ion-dependent phosphohydrolases require metal ions for catalysis, we evaluated the cation dependence of substrate hydrolysis by CBS-PPases. Cation dependencies were similar for both *mt*- and *cp*CBS-PPase, where maximal activities were observed in the presence of Co^{2+} and Mg^{2+}. Metals other than Co^{2+} are less effective at activation of PP_i and P_3 hydrolysis, in the order $Co^{2+}>Mn^{2+}>Ca^{2+}>Zn^{2+}>Mg^{2+}$. Also, metals other than Co^{2+} seem to leave rapidly the active site upon dilution or in the presence of an excess of other metals to yield a lower activity complex. The Co^{2+} complex is stable, however, even in the presence of 20 mM Mg^{2+} (Paper I, V). The family II PPase of *Methanococcus jannaschii* has similar cofactor specificity (Kuhn *et al.* 2000). In other known family II PPases, Mg^{2+} supports a moderate activity against PP_i (Parfenyev *et al.* 2001, Zyryanov *et al.* 2004b). Thus Co^{2+} is the probable physiological ligand of CBS-PPases but Mg^{2+} is required for maximal activity.

The PP_i and P_3 concentration dependencies of *cp*CBS-PPase were examined at fixed free Mg^{2+} concentrations. The dependencies obeyed the Michaelis-Menten equation. The dissociation constant for Mg^{2+} binding was 2–12 mM for free enzyme, and 1–4 mM for the ES complex. Substrate was deduced to bind independently to the Co^{2+}-bound enzyme as either $MgPP_i$ or MgP_3, and to bind one additional metal before hydrolysis occurs (Paper V).

4.2.2 CBS-PPases are dimers

Initially, it was observed that the CBS-PPases from *cp*CBS-PPase (Paper V) and *fn*CBS-PPase (Jämsen, unpublished results) were prone to aggregation, eluted at positions corresponding to tetramer and above during gel filtration and were soluble only at a concentration up to 5-15 mg/ml. However, *mt*CBS-PPase could be easily concentrated to above 100 mg/ml and displayed no higher order peaks in gel filtration, while migrating as a single band on native PAGE. Native PAGE analysis of *cp*CBS-PPase produced a smear of bands around the MW of 60 kDa with some of the sample

remaining in the well. Further, sedimentation and cross-linking experiments in the presence of Co^{2+} demonstrated that 10 μM *mt*CBS-PPase (Paper I), 9 μM *cp*CBS-PPase and 36 μM *cp*CBS insert, the latter including 2xCBS domains and one DRTGG domain (Paper III), are dimers in solution. The sedimentation coefficient ($s^0_{20,w}$) for *mt*CBS-PPase was 7.7 ± 0.1 S in the presence of Co^{2+} and 6.4 ± 0.1 S in the absence of metals (Paper I), while the same values for *cp*CBS-PPase were 8.4 ± 0.4 S and 4.9 ± 0.1 S (Paper III), respectively. Metal identity (Mg^{2+}, Mn^{2+} or Co^{2+}) or addition of nucleotides (AMP for *cp*CBS-PPase and AMP, ADP, ATP for *mt*CBS-PPase) had a negligible effect on oligomeric status (Papers I, III). However, the sedimentation coefficient for the insert remained approximately the same, indicating that it remains dimeric, while the full-length enzyme dissociated into monomers in the absence of metals (Paper III). The results parallel those for other family II PPases, where transition metals stimulated dimerization, whereas dimers dissociated into monomers upon metal removal (Parfenyev *et al.* 2001). The fact that *mt*CBS-PPase, which lacks the DRTGG domain, has no propensity for aggregation at concentrations up to 1 mM, and *cp*CBS-PPase displays severe aggregation at <100 μM suggests that the DRTGG domain has a role in multimerization of CBS-PPase. This is supported by the DRTGG-mediated oligomerization of phosphotransacetylase (Campos-Bermudez *et al.* 2010, Xu *et al.* 2005).

4.2.3 Regulation by adenosine-containing effectors

A wide range of ~80 adenosine- and other nucleoside-containing compounds were tested for effects on PP_i hydrolytic activity with *mt*- (Paper I) and *cp*CBS-PPase, in addition to effects on P_3 hydrolytic activity for the latter (Paper V). At saturating (200 μM) substrate concentration, most adenine nucleotides (100 μM) affected the PP_i hydrolytic activities of *mt*CBS-PPase, where other diphosphate containing nucleotides also had an effect. Diadenosine nucleotides had no effect on hydrolysis by *mt*CBS-PPase (Paper I). For *cp*CBS-PPase, (di)adenosine nucleotides had an effect on PP_i hydrolysis, except for cAMP, AP_4 and AP_2A, and activity against P_3 was strongly affected by guanine nucleotides (GMP, GDP, GTP) (Paper V). Regulation by adenine nucleotides is unique to CBS-PPases among inorganic pyrophosphatases and suggests a link to regulation of PPase activity *via* changes in overall metabolic energy levels.

The effects of AMP, ADP, and ATP on initial rate of *mt*- and *cp*CBS-PPase, including effect of AP_4A on the latter, were studied as a function of nucleotide and substrate concentrations (Papers I, V). Figure 6A shows representative concentration dependencies for *mt*CBS-PPase at saturating Mg_2PP_i concentration (200 μM). ATP was found to activate *mt*CBS-PPase 1.6-fold with a K_d of 0.20 ± 0.05 μM, while AMP and ADP inhibited 27 and 590-fold, respectively with K_d values of 0.22 ± 0.02 and 0.012 ± 0.001 μM (Paper I). For *cp*CBS-PPase, AMP and ADP inhibited 33-fold with K_d values of 1.4 ± 0.4 and 3.6 ± 1.3 μM, while ATP and AP_4A activated 2.7 and 3.6-fold, with K_d values of 24 ± 9 and 0.18 ± 0.05 μM, respectively. With P_3 as substrate, AMP inhibited to zero with a K_d of 0.7 ± 0.1 μM, ADP inhibited 50-fold with a K_d of 18 ± 4 μM, while ATP and AP_4A activated 1.9 and 2.8-fold with K_d's of 210 ± 20 and 0.14 ± 0.04 μM, respectively (Paper V).

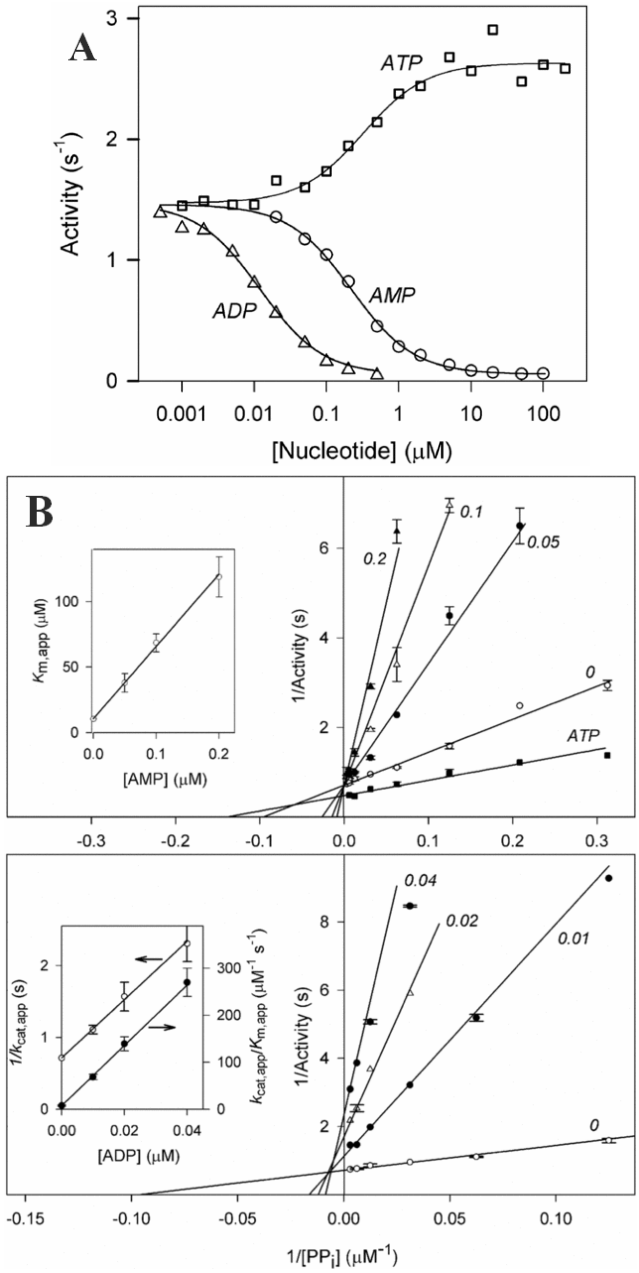


Figure 6. A) Concentration dependence of *mt*CBS-PPase activity at fixed [PP_i] (200 μM), B) Double-reciprocal plots of *mt*CBS-PPase activity as a function of substrate concentration at various fixed AMP (upper panel) and ADP (lower panel). The line marked ATP shows the effect of 10 μM ATP. The insets show secondary plots of derived parameters against [nucleotide] (Paper I).

Determination of inhibition and activation type, as shown in representative figure 6B, revealed that AMP inhibition was of competitive type for both *mt*- and *cp*CBS-PPase, whereas ADP inhibition and ATP activation were of mixed types for *mt*CBS-PPase, similarly to AP_4A activation type with *cp*CBS-PPase (Paper I,V).

The above results lead to the conclusion that CBS domains internally inhibit CBS-PPases compared to family II PPase lacking this regulatory region making the catalytic domain sensitive to structural changes caused by adenine nucleotide binding to the CBS domains (Paper I). The binding affinity and type of effect of ligands on CBS-PPases thus seems to be determined by the length of the phosphate chain and presence of the stabilizing DRTGG domain. Monophosphates, in fact, seem to work better as inhibitors of *cp*CBS-PPase (Paper I) compared to *mt*CBS-PPase, where ADP and other nucleotide diphosphates possess a more profound inhibitory potential than monophosphates (Paper V). Both types apparently stabilize inactive conformations of CBS-PPases. Three or more phosphate residues seem to be required for activation, where ATP activates both *mt*- and *cp*CBS-PPase, but with a much higher K_d than (di)adenosine polyphosphates with ≥3 phosphate residues. The activators seem to stabilize thermodynamically unfavorable conformations. ADP and ATP can bind to the enzyme-substrate complex and free enzyme, whilst AMP can only bind to the free enzyme (Paper I, V). The type of effect seems to be independent of the identity of metal activator, as determined for *mt*CBS-PPase in the presence of Mn^{2+}, but the magnitude of the effect was lower (Paper I). Additionally, presence of the DRTGG domain enables regulation by diadenosine nucleotides in addition to monoadenosine nucleotides (Paper III, V) perhaps due to greater flexibility in the regulatory insert and stabilization of the interaction between CBS domains, as these are notoriously flexible.

4.2.4 Substrate- and nucleotide-induced transitions

Addition of PP_i to a solution of *mt*CBS-PPase in the presence of 0.1 mM Co^{2+}, 5 mM Mg^{2+} at pH 7.2, and detection of resulting P_i accumulation, revealed a pre-steady state activity transition during the timescale (~30 min) of the assay (Figure 7A). The transition occurred from a lower activity to a higher activity with a halftime of 1 minute.

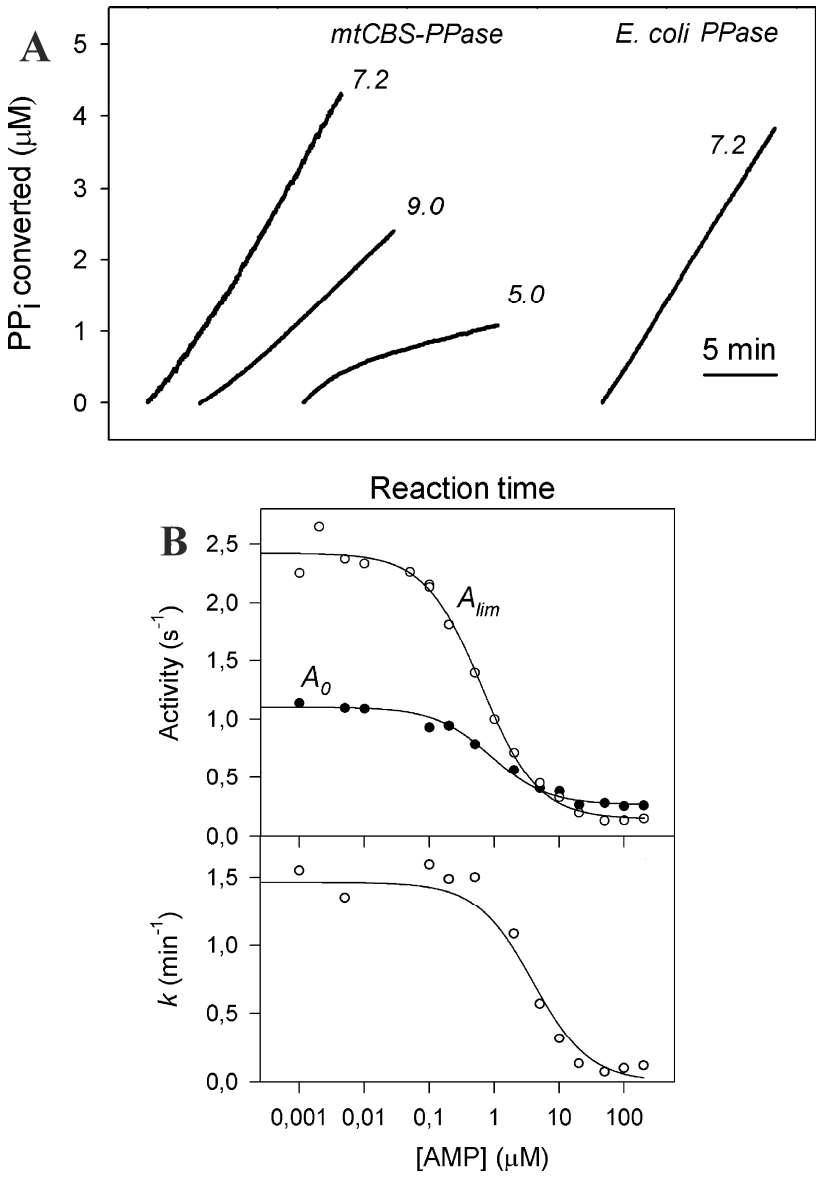


Figure 7. A) Timecourse of PP_i hydrolysis by *mt*CBS-PPase at pH 5, 7.2, and 9 in the presence of 100 μM Mg_2PP_i, 0.1 mM Co^{2+}, and 5 mM Mg^{2+} in comparison to the linear timecourse of *E. coli* PPase at pH 7.2 in the same conditions. B) dependence of k, A_0 and A_{lim} on increasing [AMP] determined in the same conditions as in A (Paper II).

This type of phenomenon is commonly presented, among other possibilities, when the enzyme exists in at least two different conformations (or populations of) at the beginning of the reaction, which interconvert to bind substrate (E→E*+S→E*S) or as

a consequence of substrate binding (E+S→ES→E*S) (Paper II). In contrast, *cp*CBS-PPase displayed a burst of product formation during the first ≤0.1 sec of PP_i hydrolysis, as determined with stopped-flow analysis of phosphate production *via* MDCC-labeled *E. coli* P_i-binding protein (unpublished results). This phenomenon is presented, for example, when there is a conformational change following formation of a reaction intermediate accompanied by product release.

Characterization of the transition observed for *mt*CBS-PPase in terms of a simple two-state model accounting for initial (A_0) and final (A_{lim}) activities, as well as, the rate of transition (k) between them, revealed that the substrate concentration dependence at various pH values could be accounted for by the Michaelis-Menten equation. Variation in pH generally slowed down the transition (basic pH) or inactivated the enzyme (acidic pH), while K_m values were similar. Omission of Co^{2+} or Mg^{2+} resulted in a linear P_i production curve. Importantly, incubation as a function of [E] or pH had no effect on the shape of the timecourse, indicating that the transition results from effects other than changes in oligomerization. The above data also indicates that the transition occurs in the ES complex and not in free E. Interestingly, increasing [AMP] led to an inactivating transition of the enzyme in contrast to activation in its absence, while the rate of transition approached zero at infinite AMP concentration (figure 7B). The dependencies revealed two nucleotide binding sites with apparent dissociation constants of ~0.5 and ~5 μM. The approximate fractions of each form of *mt*CBS-PPase could further be estimated from kinetic titration of ADP tight-binding sites (paper III). In the absence of substrate, the more active form comprised 15-33% of the stock enzyme, being more sensitive to effects of ADP (and AMP), while specific activities of the different forms differed 10-fold, yielding a 2.5-fold increase in activity (paper II). Similar differentially active enzyme forms have also been observed for the endosomal ClC-5 protein, a Cl^-/H^+ antiporter regulated by CBS domains (Zifarelli & Pusch 2009).

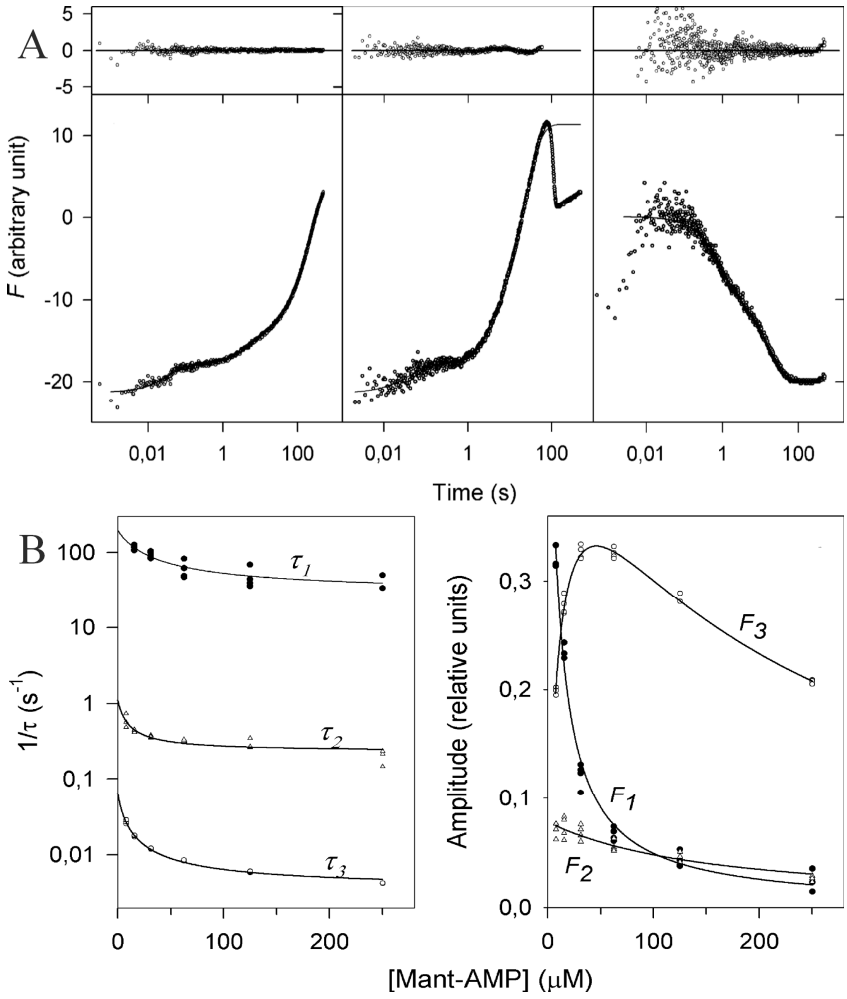


Figure 8. A) Stopped flow fluorescence traces of 500 μM Mant-AMP binding to 5 μM *mt*CBS-PPase at pH 7.2, 0.1 mM Co^{2+}, 5 mM Mg^{2+} (left), in the presence of 500 μM PP_i (middle) and binding of 30 μM Mant-AMP to a mixture of 250 μM AMP and 30 μM Mant-AMP. The upper panels show the distribution of residuals in the fit. Excitation of Trp fluorescence was at 295 nm, while Mant emission was monitored at >400 nm. Fitting was to two- or three exponential equation. B) Dependence of the inverse transit times (left) and amplitudes (right) derived from A) on [Mant-AMP] (right). Fitting was to hyperbolic equations.

Stopped-flow measurements using a fluorescent AMP analogue, 2'-(3-')-O-(N-methylanthraniloyl)-AMP (Mant-AMP) were used to directly monitor nucleotide binding to CBS domains. These measurements utilized an increase in fluorescence resonance energy transfer (FRET) between W128 and Mant-AMP, located in the boundary between the regulatory and catalytic domains. In other words, protein fluorescence was excited at 295 nm and Mant-AMP fluorescence emission was

monitored at 455 nm upon mixing enzyme with Mant-AMP. As shown in figure 8A, nucleotide binding revealed a 3-step increase in Mant-AMP fluorescence emission upon binding to *mt*CBS-PPase with relaxation times of 25 ms, 12 s, and 156 s implying large conformational changes in the complex (Paper II). When mixing Mant-AMP with *cp*CBS-PPase, there were only two transitions, with relaxation times of 10 and 4 ms (Paper V). In both cases, the binding step was too fast to be observed with our apparatus. In the case of *mt*CBS-PPase, the effect was reversed by addition of unlabeled AMP in two steps with relaxation times of 0.85 and 20 s (Paper II). The relaxation times and amplitudes of the fluorescence signals of *mt*CBS-PPase depended on [Mant-AMP] in a hyperbolic manner, which suggests the presence of two binding sites on the protein with K_d values of <1μM and 10 μM (figure 8B, Paper II). The dependence of the relaxation times of *cp*CBS-PPase displayed hyperbola with a K_d value of 20 μM (Paper V). Mixing *mt*CBS-PPase with Mant-AMP in the presence and absence of different metals indicated that addition of Mg^{2+} seemed to generally decrease the rates of all 3 steps of the nucleotide-induced transition, while metal-depleted rates of steps 2-3 by 3-fold, and Co^{2+}-bound enzyme increased rates of steps 1-2 by 12-fold. Mg^{2+} and Co^{2+} increased the amplitude of step 3, while metal-depleted enzyme displayed lower amplitude for this step. Addition of PP_i to a mixture of Mant-AMP and enzyme at catalytic concentrations was shown to produce a distinct conformation separate from the conformations generated by nucleotide binding. This is also supported by AMP incubation with enzyme prior to substrate addition in P_i accumulation measurements (Paper II).

Equilibrium fluorescence titration experiments with *mt*CBS-PPase and increasing concentrations of Mant-AMP or unlabeled AMP also supported the presence of two binding sites with dissociation constants of 1 and 10 μM, whereas equilibrium dialysis and membrane filtration measurements with ^{14}C-AMP revealed binding of one AMP molecule per enzyme monomer with a K_d of 10 μM (Paper II). Interestingly, AMP caused a decrease, while AP_3A– AP_6A caused an increase and AP_2A failed to induce a change in *cp*CBS-PPase intrinsic protein fluorescence, indicating that different conformations are generated by these compounds. Mant-AMP and AP_4A had K_d values of 60 ± 20 and 0.4 ± 0.1 μM, respectively (Paper V). Collectively, these observations suggest negative cooperativity in nucleotide binding. The site with the lower K_d value is probably the inhibitory site, while the other site is a modulatory site. Nucleotides are unable to induce interconversion of the populations without addition of substrate and

generate a conformation which could be distinguished in our studies. Overall, the data suggest a tight coupling between the regulatory and catalytic domains.

4.2.5 The structural basis of regulation

In paper III, we presented the structure of the regulatory region of *cp*CBS-PPase. The insert consists of two CBS domains separated in sequence by 118 residues corresponding to the DRTGG domain, in complex with two molecules of AMP in one structure per PPase dimer, and one molecule of AP_4A in the other. The topology of the CBS domains is similar to that of several previously solved structures of CBS domains, while the structure of DRTGG domain was determined for the first time. Figure 9A shows a model of the full-length structure of *cp*CBS-PPase, which includes the determined structure of the regulatory insert. The two CBS domain pairs are adjacent to each other, while the DRTGG domains are next to them, proximal to the catalytic domains.

Upon superimposition of the CBS domain dimers, the activator and inhibitor-bound structures display a difference in conformation. The adenine in both the AMP and AP_4A structures interacts with M114 from CBS1, and Y278, N280 and V258 from CBS2. The ribose sugar hydrogen bonds to N119 and T253 in CBS2, while the phosphate of AMP interacts with S116, S118 and K100 from the other subunit. In the activated structure bound to AP_4A, the ribose interacts with S116 and S118, while the phosphates are bound by S279 and K100 from the other subunit. Thus, activator binding shifts the $RY_{278}SN$ loop connecting α-helix1 and β-strand2 of CBS2 by 4.2 Å, leading to an opened CBS2 domain dimer interface. This in turn may allow the DHHA2 domain to move more freely relative to the DHH domain and so fine-tune the structure for catalysis. Figure 9B shows the two segments between the catalytic and regulatory domains (residues 66–70 and 305–308), which may be responsible for signal transduction between these domains. The regulatory region is also close to D311, which coordinates metals, and through which changes in conformation may affect catalysis. Dimerization is head-to-head in the CBS tetramer, with the dimer contacts formed by the two α-helices of CBS1 and helix 5 of the DRTGG domain. These contacts change little in inhibited vs. activated structures, indicating that oligomerization of *cp*CBS-PPase is not affected by nucleotide binding.

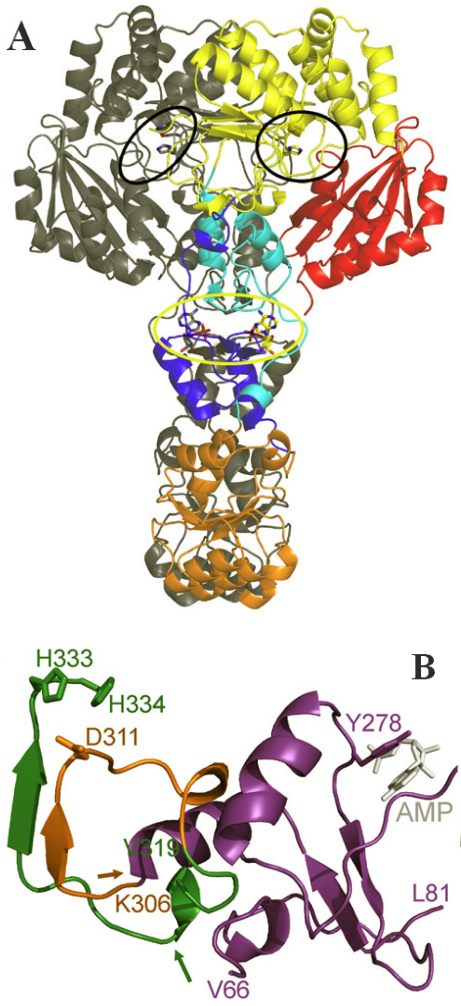


Figure 9. A) Homology model of the structure of dimeric full-length *cp*CBS-PPase with active sites highlighted in black and CBS domain binding site in yellow. The DHH domain is in yellow, DHHA2 in red, DRTGG in orange, and CBS domains are in blue and cyan (Paper III). B) Elements of the interdomain signal transduction system, which transmits the signal from the regulatory to the catalytic domain. The arrows show possible movement in the depicted segments.

4.2.6 Possible physiological functions of CBS-PPases

Biosynthetic reactions produce PP_i, the intracellular concentration of which increases during the logarithmic phase of microbial growth. The PP_i must be degraded so as to prevent inhibition of critical biosynthetic reactions (Heinonen *et al.* 2001). High ATP concentrations enable microbial growth, but this also stimulates the CBS-

PPase mediated hydrolysis of the PP_i produced by these ATP-dependent reactions, as CBS-PPases are activated by ATP and inhibited by AMP and ADP. In contrast, when growth slows down, the decrease in the [ATP]/([AMP]+[ADP]) ratio would inhibit PPase activity. During the latter phase, membrane PPase is able to maintain metabolism by creating an H^+ potential gradient across various membranes together with H^+-ATPase (Paper I). E. g., a basal level of PP_i is maintained in *E. coli* even when [PPase] increases, thus the control of PPase activity is crucial to the maintenance of metabolism under cellular stress (Heinonen *et al.* 2001).

The P_3 hydrolytic function of CBS-PPases, which possess a DRTGG domain in addition to two CBS domains, seems to be the predominant activity of this enzyme, as the k_{cat}/K_m ratio of *cp*CBS-PPase for P_3 at 1 mM Mg^{2+} (approximate physiological concentration) is about 11 times higher than for PP_i. There is an *E. coli*-type PPX in *Clostridium perfringens*, which degrades longer chain polyphosphate and is inhibited by P_3, which CBS-PPase is able to degrade into phosphate, thus partly completing the phosphate metabolic cycle. *cp*CBS-PPase is the only PPase in *Clostridium perfringens*, and so the PP_i hydrolytic activity is also significant (Paper V).

4.3 CBS Domain mutational analysis (Paper IV)

We generated enzyme with point mutations in the CBS domains of *mt*CBS-PPase and characterized their effects on catalytic efficiency, interactions with adenine nucleotides, and effects on the previously observed substrate-induced transition of *mt*CBS-PPase (Paper II).

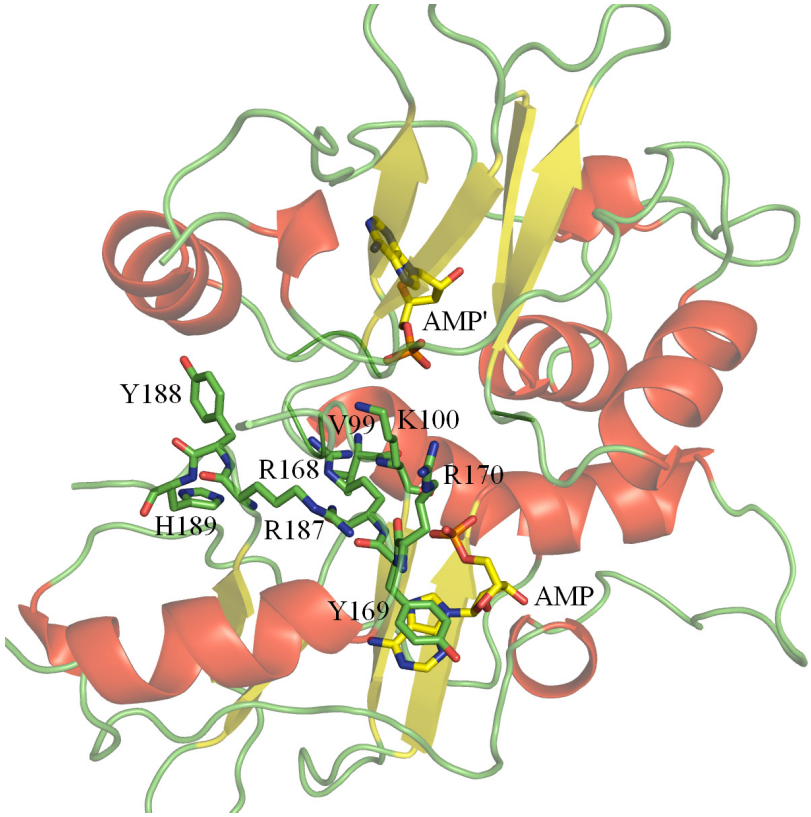


Figure 10. Model of the CBS domains (CBS1 above, CBS2 below) of *mt*CBS-PPase with bound AMP based on the structure of *cp*CBS (PDB ID 3LB2, Paper III). The residues chosen for substitution in this study are labeled. Residue numbering is for *mt*CBS-PPase. The primed AMP is located in the other subunit.

As shown in figure 10, the residues corresponding to Lys100, Tyr169 and Arg170 in the CBS domains of *mt*CBS-PPase bind the AMP. V99 is in the same position as Y169 in CBS2, while residues at positions 168−170 are part of a conserved RYRN loop. Based on the crystal structure, conformational changes in this loop between activated and inhibited enzyme are thought to control coupling between the CBS and PPase domains. We mutated most residues into alanine, conservatively replacing Y169 with phenylalanine, while K100 and R187 were substituted with glycine, because these substitutions are known to affect the function of pig AMP-dependent protein kinase.

4.3.1 CBS and PPase domains are tightly coupled in native CBS-PPase

The effects of mutations on the activity at the end of the reaction (A_{lim}) was used to quantify the effects of CBS domain mutations on catalysis. The effects on k_{cat} could be divided into three groups (table 4): (i) similar to wild type (K100G), (ii) more active (V99A, R168A, Y169A, Y169F, Y188A, H189A) by 1,3-3,3-fold, and (iii) less active (R170A, R187G) by 2,5-fold. The K_m in each case was approximately 10 μM for all variants within experimental error. The fact that substitutions in a regulatory domain affect catalysis to such a large extent, signifies that the two domains are tightly coupled to each other and that even small changes in the structure of one domain are translated to changes in the other. Since the majority of substitutions caused an increase in catalytic efficiency (k_{cat}/K_m), overall the results support the hypothesis that the CBS insert functions as an autoinhibitor of CBS-PPase, where mutations either increase or decrease the inhibitory coupling between domains.

4.3.2 Effects on nucleotide binding and inhibition

As shown in table 4, AMP and ADP generally inhibited most variants similarly to wild type. A remarkable activation of 1.2- and 2.8-fold by AMP was observed with the substitutions Y169A and R168A, respectively. The degree of AMP inhibition was increased by the substitutions Y188A and H189A. The kinetically determined AMP-binding affinity ($K_{i,app}$) was significantly lower in V99A, R168A and Y169A, while an increase of 10-100 fold was observed with the substitutions R187G and H189A. The stoichiometry of ^{14}C-AMP binding was increased from 1 to 2 molecules per CBS pair in the variants Y169A, R170A, R187G and Y188A. This can be explained by conformational changes in CBS2, which increase affinity for nucleotides, decrease the negative cooperativity between domains, or change the interaction between CBS domain dimers. ADP inhibited all variants by around 5-fold, where the substitutions V99A, Y169A and R170A significantly decreased, while the substitutions Y169F, R187G and H189A increased the affinity of ADP to enzyme by 100-1000 fold. Elimination of the tyrosine ring of Y169 eliminated ADP binding and increased the K_d for AMP binding.

K100 is thought to bind the first phosphate after the adenine ring, thus the substitution K100G decreased K_i and K_d for AMP while K_i for ADP was increased. Similarly, R170 binds the second phosphate of ADP, thus the substitution of this residue had the opposite effect on constants for AMP and ADP binding. In general, the substitution of residues K100, Y169 and R170, which were thought to change contacts with the nucleotide ligand based on the structure of *cp*CBS-PPase, had the expected effects on constants. The other substituted residues reside in a mobile loop (R168), which adopts different conformations in activated and inhibited form of the regulatory region, or are nearby in the structure (R187, Y188). R187 apparently has different conformations in activated and inhibited structures, while Y188 is a critical residue in the CBS dimer interface. The latter three residues were thus hypothesized to control the tight coupling between the CBS and PPase domains.

Analysis of the effects of substitutions on the substrate-induced transition of *mt*CBS-PPase revealed that the ratio of initial to final activity ($k_{cat,0}/k_{cat,lim}$) was not much different from wild-type, except for the substitutions Y169A and R170A, which displayed no transition. This can be explained by the substitutions decoupling the catalytic and regulatory domains, supported by the fact that these residues are part of the conserved RYRN loop, which undergoes the largest change in conformation when comparing AMP and AP_4A bound structure of the regulatory region of *cp*CBS-PPase. The rate constant characterizing the transition was 1.5-3-fold lower in mutants in comparison to wild type, and nucleotide binding was generally found to slow the transitions.

Table 4. Effects of substitutions on catalytic activity, nucleotide inhibition and AMP binding. Values with large deviation from wild type are shown in bold.

	Catalytic activity		AMP inhibition/activation*		ADP inhibition*		AMP binding†	
Enzyme	$k_{cat,lim}$ (s^{-1})	K_m (μM)	$K_{i,app}$ (μM)	Residual activity‡ (%)	$K_{i,app}$ (μM)	Residual activity‡ (%)	K_d (μM)	Binding stoichiometry (mol/mol)
Wild-type	2.2 ± 0.2	8 ± 3	0.6 ± 0.1	6 ± 2	1.0 ± 0.4	14 ± 5	13 ± 1	0.94 ± 0.01
V99A	**7.3 ± 0.5**	3.7 ± 1.6	**>1000**	n.d.	**>1000**	n.d.	**>2000**	n.d.
K100G	1.9 ± 0.1	13 ± 4	1.7 ± 0.2	4 ± 2	0.13 ± 0.03	8 ± 3	340 ± 90	1.1 ± 0.02
KR168A	2.8 ± 0.3	11 ± 5	**11 ± 5**	**280 ± 30**	0.8 ± 0.2	12 ± 4	35 ± 8	1.06 ± 0.07
Y169A	**5.0 ± 0.7**	9.4 ± 2.8	**8 ± 3**	**122 ± 2**	**>1000**	n.d.	**1000 ± 200**	0.7 ± 0.3
Y169F	**7.1 ± 0.8**	12 ± 6	0.39 ± 0.06	<2	**0.017 ± 0.001**	<2	50 ± 10	**1.8 ± 0.1**
R170A	0.88 ± 0.06	7.9 ± 2.6	0.9 ± 0.1	2 ± 3	**57 ± 36**	<20	17 ± 1	**2.0 ± 0.1**
R187G	0.98 ± 0.09	5.7 ± 2.4	**0.051 ± 0.007**	<4	**0.009 ± 0.001**	<2	27 ± 7	**1.6 ± 0.1**
Y188A	**5.8 ± 0.9**	15 ± 9	2.0 ± 0.7	**37 ± 4**	1.8 ± 0.8	5 ± 1	150 ± 50	**2.1 ± 0.1**
H189A	**3.8 ± 0.2**	17 ± 4	**0.0015 ± 0.0006**	**44 ± 3**	**0.011 ± 0.004**	5 ± 3	1400 ± 100	1.3 ± 0.1

*Measured at 160 μM PP_i.
†Measured using the filtration assay.
‡Activity at infinite nucleotide concentration compared with activity at zero nucleotide concentration.

5. ACKNOWLEDGEMENTS

This study was carried out at the University of Turku, Department of Biochemistry during the years 2006-2011. Major support was received from the National Doctoral School in Informational and Structural Biology (ISB), the Alfred Kordelin Foundation and the Turku University Foundation.

I would especially like to thank my supervisor Prof. Reijo Lahti for giving me the opportunity to work in his group and for his flexibility during my studies. I also especially want to thank my supervisor from Moscow, Prof. Alexander Baykov, for lending his expertise to help me. Prof. Jyrki Heino is acknowledged for providing equipment and facilities. I would like thank the reviewers of this thesis, Prof. Ari Hinkkanen and Prof. Simo Laakso for their valuable comments, and Prof. Markku Kulomaa for acting as the opponent in my defence.

I would like to thank my co-authors for their help in preparing the papers presented in this thesis. I want to thank Heidi Tuominen for co-authorship in papers I, III and IV, and for help in general. I also want to thank Drs. Anu Salminen and Georgiy Belogurov for assistance, and the students Juho Vuononvirta and Minna Peippo for contributing to paper III. I also would like to thank PPase lab members Anssi Malinen, Marko Tammenkoski, Heidi Tuominen and Heidi Luoto for great times in the lab, and Anssi for guiding me through my M.Sc thesis.

I would like to thank Prof. Mark Johnson for his advice, for taking me into his lab at a young age during high school and for all his support during these years. Thanks also to Prof. Henri Xhaard, for being my mentor during my first experience in a scientific laboratory and for the fun times later on in Helsinki. I also want to thank Yuezhou, Michal, Gloria, Aino-Leena and Lasse in the Xhaard Lab at the Department of Pharmacy, University of Helsinki.

During these five years, the graduate school has organized outings, as well as, excellent meetings and offered travel grants to many places which I was glad to be able to visit. I want to thank Mark Johnson and Fredrik Karlsson for providing support from the graduate school.

I would like to thank my friends for their support during the years, even if I am not in touch with some of them. I would like to thank various people for interesting conversations and outings (in alphabetical order): Anssi Malinen, Anu Salminen, Esko Oksanen, Georgiy Belogurov, Joeri Mertens, Juho Kellosalo, Kalle Rytkönen, Matti Ilonen, Matti Lahti, Pauli Kallio, Pasi Kankaanpää, Sami Lamppu and Teijo Pellinen.

Finally, I am grateful to my parents Pekka and Katriina. You helped me when times were rough, especially when I was working on my thesis. I thank you for that. You both have tried to teach me what is important in life and that's everything a son could ask from his parents.

6. REFERENCES

Abu Kwaik Y. (1998) Induced expression of the *Legionella pneumophila* gene encoding a 20-kilodalton protein during intracellular infection. *Infect Immun* 66, 203-12.

Accardi A & Miller C. (2004) Secondary active transport mediated by a prokaryotic homologue of ClC Cl^- channels. *Nature* 427, 803-7.

Aguado-Llera D, Oyenarte I, Martínez-Cruz LA and Neira JL. (2010) The CBS domain protein MJ0729 of *Methanocaldococcus jannaschii* binds DNA. *FEBS Lett* 584, 4485-9.

Ahn S, Milner AJ, Fütterer K, Konopka M, Ilias M, Young TW and White SA. (2001) The "open" and "closed" structures of the type-C inorganic pyrophosphatases from *Bacillus subtilis* and *Streptococcus gordonii*. *J Mol Biol* 313, 797-811.

Aherne A, Kennan A, Kenna PF, McNally N, Lloyd DG, Alberts IL, Kiang AS, Humphries MM, Ayuso C, Engel PC, Gu JJ, Mitchell BS, Farrar GJ and Humphries P. (2004) On the molecular pathology of neurodegeneration in IMPDH1-based retinitis pigmentosa. *Hum Mol Genet* 13, 641–650.

Aitken SM & Kirsch JF. (2003a) Kinetics of the yeast cystathionine β-synthase forward and reverse reactions: continuous assays and the equilibrium constant for the reaction. *Biochemistry* 42, 571–578.

Aitken SM & Kirsch JF. (2003b) *Escherichia coli* cystathionine β-synthase does not obey ping-pong kinetics. Novel continuous assays for the elimination and substitution reactions. *Biochemistry* 42, 11297–11306.

Aitken SM & Kirsch JF. (2005) The enzymology of cystathionine biosynthesis: strategies for the control of substrate and reaction specificity. *Arch Biochem Biophys* 433, 166-75.

Akiyama M, Crooke E and Kornberg A. (1993). An exopolyphosphatase of *Escherichia coli*. The enzyme and its *ppx* gene in a polyphosphate operon. *J Biol Chem* 268, 633–639.

Amodeo GA, Rudolph MJ and Tong L. (2007) Crystal structure of the heterotrimer core of *Saccharomyces cerevisiae* AMPK homologue SNF1. *Nature* 449, 492-495.

Anderson J & Sartorelli A. (1968) Inosinic acid dehydrogenase of sarcoma 180 cells. *J Biol Chem* 243, 4762-4768.

Apic G, Gough J and Teichmann SA. (2001) Domain combinations in archaeal, eubacterial and eukaryotic proteomes. *J Mol Biol* 310, 311–325

Aravind L & Koonin EV. (1998) A novel family of predicted phosphoesterases includes Drosophila prune protein and bacterial RecJ exonuclease. *Trends Biochem Sci* 23, 17-19.

Baltrop JA, Grobb PW and B Heap B. (1963) Mechanisms for oxidative phosphorylation at the pyridine nucleotide flavoprotein level. *Nature (London)* 199, 759-761.

Baltscheffsky H, Von Stedingk LV, Heldt HW, Klingenberg M. (1966) Inorganic pyrophosphate: formation in bacterial photophosphorylation. *Science* 153, 1120-2.

Banerjee R, Evande R, Kabil O, Ojha S and Taoka S. (2003) Reaction mechanism and regulation of cystathionine β-synthase. *Biochim Biophys Acta* 1647, 30-5.

Banerjee R & Zou CG. (2005) Redox regulation and reaction mechanism of human cystathionine-β-synthase: a PLP-dependent hemesensor protein. *Arch Biochem Biophys* 433, 144-56.

Bashton M & Chothia C. (2002) The geometry of domain combination in proteins. *J Mol Biol* 315, 927-39.

Basu MK, Carmel L, Rogozin IB and Koonin EV. (2008) Evolution of protein domain promiscuity in eukaryotes. *Genome Res* 18, 449–61.

Bateman A. (1997) The structure of a domain common to archaebacteria and the homocystinuria disease protein. *Trends Biochem Sci* 22, 12-13.

Baykov AA & Avaeva SM. (1981) A simple and sensitive apparatus for continuous monitoring of orthophosphate in the presence of acid-labile compounds. *Anal Biochem* 116, 1-4.

Baykov AA, Shestakov AS, Kasho VN, Vener AV, Ivanov AH. (1990) Kinetics and thermodynamics of catalysis by the inorganic

pyrophosphatase of *Escherichia coli* in both directions. *Eur J Biochem* 194, 879-87.

Baykov AA & Shestakov AS. (1992) Two pathways of pyrophosphate hydrolysis and synthesis by yeast inorganic pyrophosphatase. *Eur J Biochem* 206, 463-70.

Baykov AA, Bakuleva NP and Rea PA. (1993). Steady-state kinetics of substrate hydrolysis by vacuolar H^+-pyrophosphatase. A simple three-state model. *Eur J Biochem* 217, 755-762.

Baykov AA, Cooperman BS, Goldman A and Lahti R. (1999) Cytoplasmic inorganic pyrophosphatase. *Prog Mol Subcell Biol* 23, 127-50.

Baykov AA, Fabrichniy IP, Pohjanjoki P, Zyryanov AB and Lahti R. (2000) Fluoride effects along the reaction pathway of pyrophosphatase: evidence for a second enzyme.pyrophosphate intermediate. *Biochemistry* 39, 11939-47.

Belogurov GA, Fabrichniy IP, Pohjanjoki P, Kasho VN, Lehtihuhta E, Turkina MV, Cooperman BS, Goldman A, Baykov AA and Lahti R. (2000) Catalytically important ionizations along the reaction pathway of yeast pyrophosphatase. *Biochemistry* 39, 13931-8.

Belogurov GA, Turkina MV, Penttinen A, Huopalahti S, Baykov AA and Lahti R. (2002) H+-pyrophosphatase of *Rhodospirillum rubrum*. High yield expression in Escherichia coli and identification of the Cys residues responsible for inactivation by mersalyl. *J Biol Chem* 277, 22209-14.

Bennetts B, Rychkov GY, Ng HL, Morton CJ, Stapleton D, Parker MW and Cromer BA. (2005) Cytoplasmic ATP-sensing domains regulate gating of skeletal muscle ClC-1 chloride channels. *J Biol Chem* 280, 32452-8.

Bennetts B, Parker MW and Cromer BA. (2007) Inhibition of skeletal muscle ClC-1 chloride channels by low intracellular pH and ATP. *J Biol Chem* 282, 32780-91.

Bielen AA, Willquist K, Engman J, van der Oost J, van Niel EW and Kengen SW. (2010) Pyrophosphate as a central energy carrier in the hydrogen-producing extremely thermophilic *Caldicellulosiruptor saccharolyticus*. *FEMS Microbiol Lett* 307, 48-54.

Biemans-Oldehinkel E, Mahmood NA and Poolman B. (2006) A sensor for intracellular ionic strength. *Proc Natl Acad Sci U S A* 103, 10624-9.

Blair E, Redwood C, Ashrafian H, Oliveira M, Broxholme J, Kerr B, Salmon A, Ostman-Smith I and Watkins H. (2001) Mutations in the γ2 subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum Mol Genet* 10, 1215-20.

Borcsok E & Abeles RH. (1982) Mechanism of action of cystathionine synthase. *Arch Biochem Biophys* 213, 695-707.

Bowne SJ, Sullivan LS, Blanton SH, Cepko CL, Blackshaw S, Birch DG, Hughbanks-Wheaton D, Heckenlively JR and Daiger SP (2002) Mutations in the inosine monophosphate dehydrogenase 1 gene (IMPDH1) cause the RP10 form of autosomal dominant retinitis pigmentosa. *Hum Mol Genet* 11, 559–568

Bowne SJ, Sullivan LS, Mortimer SE, Hedstrom L, Zhu J, Spellicy CJ, Gire AI, Hughbanks-Wheaton D, Birch DG, Lewis RA, Heckenlively JR and Daiger SP. (2006) Why do mutations in the ubiquitously expressed housekeeping gene IMPDH1 cause retina-specific photoreceptor degeneration? *Investig Ophthalmol Vis Sci* 47, 34–42

Brown MR & Kornberg A. (2004) Inorganic polyphosphate in the origin and survival of species. *Proc Natl Acad Sci USA* 101, 16085-16087.

Brown MR & Kornberg A. (2008) The long and short of it—polyphosphate, PPK and bacterial survival. *Trends Biochem Sci* 33, 284-290.

Brox LW & Hampton A. (1968) Inosine 5'-phosphate dehydrogenase. Kinetic mechanism and evidence for selective reaction of the 6-chloro analog of inosine 5'-phosphate with a cysteine residue at the inosine 5'-phosphate site. *Biochemistry* 7, 2589-96.

Brune M, Hunter JL, Corrie JE and Webb MR. (1994) Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase. *Biochemistry* 33, 8262-71.

Bukovska G, Kery V and Kraus JP. (1994) Expression of human cystathionine β-synthase in *Escherichia coli*: purification and characterization. *Protein Expr Purif* 5, 442–448.

Campos-Bermudez VA, Bologna FP, Andreo CS and Drincovich MF. (2010) Functional dissection of *Escherichia coli* phosphotransacetylase structural domains and analysis of key compounds involved in activity regulation. *FEBS J* 277, 1957-66.

Carr SF, Papp E, Wu JC and Natsumeda Y. (1993) Characterization of human type I and type II IMP dehydrogenases. *J Biol Chem* 268, 27286–27290.

Carr G, Simmons N and Sayer J. (2003) A role for CBS domain 2 in trafficking of chloride channel CLC-5. *Biochem Biophys Res Commun* 310, 600-5.

Chen J, Brevet A, Fromant M, Lévêque F, Schmitter JM, Blanquet S and Plateau P. (1990) Pyrophosphatase is essential for growth of *Escherichia coli*. *J Bacteriol* 172, 5686-9.

Chen G, Gharib TG, Huang CC, Thomas DG, Shedden KA, Taylor JM, Kardia SL, Misek DE, Giordano TJ, Iannettoni MD, Orringer MB, Hanash SM and Beer DG. (2002) Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. *Clin Cancer Res* 8, 2298-2305.

Chothia C & Lesk AM. (1986) The relation between the divergence of sequence and structure in proteins. *EMBO J* 5, 823-6.

Chothia C, Gough J, Vogel C and Teichmann SA. (2003) Evolution of the protein repertoire. *Science* 300, 1701–3.

Cleiren E, Bénichou O, Van Hul E, Gram J, Bollerslev J, Singer FR, Beaverson K, Aledo A, Whyte MP, Yoneyama T, deVernejoul MC and Van Hul W. (2001) Albers-Schönberg disease (autosomal dominant osteopetrosis, type II) results from mutations in the ClCN7 chloride channel gene. *Hum Mol Genet* 10, 2861-7.

Colby TD, Vanderveen K, Strickler MD, Markham GD and Goldstein BM. (1999) Crystal structure of human type II inosine monophosphate dehydrogenase: implications for ligand binding and drug design. *Proc Natl Acad Sci U S A*. 96, 3531-6.

Collart FR & Huberman E (1988) Cloning and sequence analysis of the human and Chinese hamster inosine-50-monophosphate dehydrogenase cDNAs. *J Biol Chem* 263, 15769–15772.

Cooperman BS. (1982) The mechanism of action of yeast inorganic pyrophosphatase. *Methods Enzymol* 87, 526-48.

Cooperman BS, Baykov AA and Lahti R. (1992) Evolutionary conservation of the active site of soluble inorganic pyrophosphatase. *Trends Biochem Sci* 17, 262-6.

Cornuel JF, Moraillon A and Gueron M. (2002) Participation of yeast inosine 5'-monophosphate dehydrogenase in an in vitro complex with a fragment of the C-rich telomeric strand. *Biochimie* (*Paris*) 84, 279–289.

Day P, Sharff A, Parra L, Cleasby A, Williams M, Hörer S, Nar H, Redemann N, Tickle I and Yon J. (2007) Structure of a CBS-domain pair from the regulatory γ1 subunit of human AMPK in complex with AMP and ZMP. *Acta Crystallogr D Biol Crystallogr* 63, 587-96.

De Clerq E. (1993) Antiviral agents: characteristic activity spectrum depending on the molecular target with which they interact. *Ad*V. *Virus Res. 42*, 1-55.

de Diego C, Gámez J, Plassart-Schiess E, Lasa A, Del Río E, Cervera C, Baiget M, Gallano P and Fontaine B. (1999) Novel mutations in the muscle chloride channel CLCN1 gene causing myotonia congenita in Spanish families. *J Neurol* 246, 825-9.

de Graaf BH, Rudd JJ, Wheeler MJ, Perry RM, Bell EM, Osman K, Franklin FC and Franklin-Tong VE. (2006) Self-incompatibility in Papaver targets soluble inorganic pyrophosphatases in pollen. *Nature* 444, 490-493.

Digits JA & Hedstrom L. (1999) Kinetic mechanism of *Tritrichomonas foetus* inosine 5'-monophosphate dehydrogenase. *Biochemistry* 38, 2295-306.

Docampo R, de Souza W, Miranda K, Rohloff P and Moreno SN. (2005) Acidocalcisomes—conserved from bacteria to man. *Nat. Rev. Microbiol.* **3**: 251-261.

Dutzler R, Campbell EB, Cadene M, Chait BT and MacKinnon R. (2002) X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. *Nature* 415, 287-94.

Estévez R, Pusch M, Ferrer-Costa C, Orozco M and Jentsch TJ. (2004) Functional and structural conservation of CBS domains from CLC chloride channels. *J Physiol* 557, 363-78.

Evande R, Ojha S and Banerjee R. (2004) Visualization of PLP-bound intermediates in hemeless variants of human cystathionine-β-synthase: evidence that lysine 119 is a general base. *Arch Biochem Biophys* 427, 188–196.

Evande R, Blom H, Boers GH and Banerjee R. (2002) Alleviation of intrasteric inhibition by the pathogenic activation domain mutation, D444N, in human cystathionine β-synthase. *Biochemistry* 41, 11832-7.

Fabrichniy IP, Lehtiö L, Salminen A, Zyryanov AB, Baykov AA, Lahti R and Goldman A. (2004) Structural studies of metal ions in family II pyrophosphatases: the requirement for a Janus ion. *Biochemistry* 43, 14403-11.

Fabrichniy IP, Lehtiö L, Tammenkoski M, Zyryanov AB, Oksanen E, Baykov AA, Lahti R and Goldman A. (2007) A trimetal site and substrate distortion in a family II inorganic pyrophosphatase. *J Biol Chem* 282, 1422-31.

Frank N, Kery V, Maclean KN and Kraus JP. (2006) Solvent-accessible cysteines in human cystathionine β-synthase: Crucial role of cysteine 431 in S-adenosyl L-methionine binding. *Biochemistry* **45**, 11021–11029.

Franklin T & Cook J. (1969) The inhibition of nucleic acid synthesis by mycophenolic acid. *Biochem. J. 113*, 515-524.

Feng L, Yan H, Wu Z, Yan N, Wang Z, Jeffrey PD and Shi Y. (2007) Structure of a site-2 protease family intramembrane metalloprotease. *Science* 318, 1608-12.

Ferrer A, Caelles C, Massot N and Hegardt FG. (1985) Activation of rat liver cytosolic 3-hydroxy-3-methylglutaryl coenzyme A reductase kinase by adenosine 5'-monophosphate. *Biochem Biophys Res Commun* 132, 497-504.

Finkelstein JD, Kyle WE, Martin JL and Pick AM. (1975) Activation of cystathionine synthase by adenosylmethionine and adenosylethionine. *Biochem Biophys Res Commun* 66, 81–87.

Finkelstein JD & Martin JJ. (1984a) Methionine metabolism in mammals. Distribution of homocysteine between competing pathways. *J Biol Chem* 259, 9508–9513.

Finkelstein JD & Martin JJ. (1984b) Inactivation of betaine-homocysteine methyltransferase by adenosylmethionine and adenosylethionine. *Biochem Biophys Res Commun* 118, 14-19

Gan L, Petsko G A and Hedstrom L. (2002) Crystal structure of a ternary complex of *Tritrichomonas foetus* inosine 5'-monophosphate dehydrogenase: NAD+ orients the active site loop for catalysis. *Biochemistry* 41, 13309–13317

Gan L, Seyedsayamdost MR, Shuto, S, Matsuda A, Petsko GA and Hedstrom L. (2003) The immunosuppressive agent mizoribine monophosphate forms a transition state analogue complex with inosine monophosphate dehydrogenase. *Biochemistry* 42, 857.

Gollob MH, Green MS, Tang AS, Gollob T, Karibe A, Ali Hassan AS, Ahmad F, Lozado R, Shah G, Fananapazir L, Bachinski LL and Roberts R. (2001) Identification of a gene responsible for familial Wolff-Parkinson-White syndrome. *N Engl J Med* 344, 1823-31.

Gomez-Garcia MR, Losada M and Serrano A. (2003). Concurrent transcriptional activation of ppa and ppx genes by phosphate deprivation in the cyanobacterium *Synechocystis sp.* strain PCC 6803. *Biochem Biophys Res Commun* 302, 601-609.

Gómez-García I, Oyenarte I and Martínez-Cruz LA. (2010) The crystal structure of protein MJ1225 from *Methanocaldococcus jannaschii* shows strong conservation of key structural features seen in the eukaryal γ-AMPK. *J Mol Biol* 399, 53-70.

Gómez-García I, Oyenarte I and Martínez-Cruz LA. (2011) Purification, crystallization and preliminary crystallographic analysis of the CBS pair of the human metal transporter CNNM4. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 67, 349-53.

Gonzalez MA, Webb MR, Welsh KM, Cooperman BS. (1984) Evidence that catalysis by yeast inorganic pyrophosphatase proceeds by direct phosphoryl transfer to water and not via phosphoryl enzyme intermediate. *Biochemistry* 35, 4655-4661.

Gough J, Karplus K, Hughey R and Chothia C. (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J Mol Biol* 313, 903-19.

Gough J. (2005) Convergent evolution of domain architectures (is rare). *Bioinformatics* 21, 1464-71.

Gouet P, Robert X and Courcelle E. (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 31, 3320-3.

Gu JJ, Spychala J and Mitchell BS (1997) Regulation of the human inosine monophosphate dehydrogenase type I gene. *J Biol Chem* 272, 4458–4466.Gu JJ, Stegmann S, Gathy K, Murray R, Laliberte J, Ayscue L and Mitchell BS. (2000) Inhibition of T lymphocyte activation in mice heterozygous for loss of the IMPDH II gene. *J Clin Inves* 106, 599–606.

Gu JJ, Tolin AK, Jain J, Huang H, Santiago L and Mitchell BS. (2003) Targeted disruption of the inosine 50-monophosphate dehydrogenase type I gene in mice. *Mol Cell Biol* 23, 6702–6712.

Gulick PJ, Drouin S, Yu Z, Danyluk J, Poisson G, Monroy AF and Sarhan F. (2005). Transcriptome comparison of winter and spring wheat responding to low temperature. *Genome* 48, 913-923.

Gunter JH, Thomas EC, Lengefeld N, Kruger SJ, Worton L, Gardiner EM, Jones A, Barnett NL and Whitehead JP. (2008) Characterisation of inosine monophosphate dehydrogenase expression during retinal development: differences between variants and isoforms. *Int J Biochem Cell Biol* 40, 1716–1728

Halonen P, Baykov AA, Goldman A, Lahti R and Cooperman BS. (2002) Single-turnover kinetics of Saccharomyces cerevisiae inorganic pyrophosphatase. *Biochemistry* 41, 12025-31.

Halonen P, Tammenkoski M, Niiranen L, Huopalahti S, Parfenyev AN, Goldman A, Baykov A and Lahti R. (2005) Effects of active site mutations on the metal binding affinity, catalytic competence, and stability of the family II pyrophosphatase from Bacillus subtilis. *Biochemistry* 44, 4004-10.

Hara-Chikuma M, Yang B, Sonawane ND, Sasaki S, Uchida S, Verkman AS. (2005) ClC-3 chloride channels facilitate endosomal acidification and chloride accumulation. *J Biol Chem* 280, 1241-7.

Hardie DG & Hawley SA. (2001) AMP-activated protein kinase: the energy charge hypothesis revisited. *Bioessays* 23, 1112-1119.

Hardie DG. (2007) AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. *Nat Rev Mol Cell Biol* 8, 774-85.

Harding MM. (2001) Geometry of metal-ligand interactions in proteins. *Acta Crystallogr D Biol Crystallogr* 57, 401-11.

Hattori M, Tanaka Y, Fukai S, Ishitani R and Nureki O. (2007) Crystal structure of the MgtE Mg^{2+} transporter. *Nature* 448: 1072–1075.

Hattori M, Iwase N, Furuya N, Tanaka Y, Tsukazaki T, Ishitani R, Maguire ME, Ito K, Maturana A and Nureki O. (2009) Mg^{2+}-dependent gating of bacterial MgtE channel underlies Mg^{2+} homeostasis. *EMBO J* 28, 3602-12.

Hedstrom L & Wang CC. (1990) Mycophenolic acid and thiazole adenine dinucleotide inhibition of *Trichomonas foetus* inosine 5'-monophosphate dehydrogenase: Implications on enzyme mechanism. *Biochemistry* 8, 2303.

Hedstrom L & Gan L. (2006) IMP dehydrogenase: structural schizophrenia and an unusual base. *Curr Opin Chem Biol* 10, 520–525

Hedstrom L. (2008) IMP dehydrogenase-linked retinitis pigmentosa. *Nucleosides Nucleotides Nucleic Acids* 27, 839–849.

Heikinheimo P, Lehtonen J, Baykov A, Lahti R, Cooperman BS and Goldman A. (1996) The structural basis for pyrophosphatase catalysis. *Structure* 4, 1491-508.

Heikinheimo P, Tuominen V, Ahonen AK, Teplyakov A, Cooperman BS, Baykov AA, Lahti R and Goldman A. (2001) Toward a quantum-mechanical description of metal-assisted phosphoryl transfer in pyrophosphatase. *Proc Natl Acad Sci U S A* 98, 3121-6.

Heinonen JK. (2001) Biological role of inorganic pyrophosphate. Kluwer Academic Publishers, London.

Heyde E, Nagabhushanam A, Vonarx M and Morrison J. (1976) Studies on inosine monophosphate dehydrogenase. Steady state kinetics. *Biochim Biophys Acta* 429, 645-660.

Holmes E, Pehlke D and Kelley W. (1974) Human IMP dehydrogenase. Kinetics and regulatory properties. *Biochim Biophys Acta* 364, 209-217.

Hooley P, Whitehead MP and Brown MR. (2008) Eukaryote polyphosphate kinases: is the 'Kornberg' complex ubiquitous? *Trends Biochem Sci* 33, 577-582.

Huang S, Colmer TD and Millar AH. (2008) Does anoxia tolerance involve altering the energy

currency towards PP_i? *Trends Plant Sci* 13, 221-227.

Huete-Perez JA, Wu JC, Witby FG and Wang CC. (1995) Identification of the IMP binding site in the IMP dehydrogenase from *Tritrichomonas foetus*. *Biochemistry* 34, 13889-13894.

Hupe D, Azzolina B and Behrens N. (1986) IMP dehydrogenase from the intracellular parasitic protozoan *Eimeria tenella* and its inhibition by mycophenolic acid. *J Biol Chem* 261, 8363-8369.

Höhne WE & Heitmann P. (1974) Tripolyphosphate as a substrate of the inorganic pyrophosphatase from baker's yeast; the role of divalent metal ions. *Acta Biol Med Ger* 33, 1-14.

Ignoul S & Eggermont J. (2005) CBS domains: structure, function, and pathology in human proteins. *Am J Physiol* 289, C1369–C1378

Janosík M, Kery V, Gaustadnes M, Maclean KN and Kraus JP. (2001) Regulation of human cystathionine β-synthase by S-adenosyl-L-methionine: evidence for two catalytically active conformations involving an autoinhibitory domain in the C-terminal region. *Biochemistry* 40, 10625-33.

Jentsch TJ & Günther W. (1997) Chloride channels: an emerging molecular picture. *Bioessays* 19, 117-26.

Jentsch TJ. (2008) CLC chloride channels and transporters: from genes to protein structure, pathology and physiology. *Crit Rev Biochem Mol Biol* 43, 3-36.

Jhee KH, McPhie P and Miles EW. (2000a) Domain architecture of the heme-independent yeast cystathionine-β-synthase provides insights into mechanisms of catalysis and regulation. *Biochemistry* 39, 10548–10556.

Jhee KH McPhie P and Miles EW. (2000b) Yeast cystathionine β-synthase is a pyridoxal phosphate enzyme but, unlike the human enzyme, is not a heme protein. *J Biol Chem 275*, 11541-4.

Jhee KH, Niks D, McPhie P, Dunn MF and Miles EW. (2001) The reaction of yeast cystathionine β-synthase is rate-limited by the conversion of aminoacrylate to cystathionine. *Biochemistry* 40, 10873–10880.

Jin X, Townley R and Shapiro L. (2007) Structural insight into AMPK regulation: ADP comes into play. *Structure* 15, 1285-95.

Josse J. (1966) Constitutive inorganic pyrophosphatase of *Escherichia coli*. 1. Purification and catalytic properties. *J Biol Chem* 241, 1938-47.

Kabil O & Banerjee R. (1999) Deletion of the regulatory domain in the pyridoxal phosphate-dependent heme protein cystathionine β-synthase alleviates the defect observed in a catalytic site mutant. *J Biol Chem* 274, 31256-60.

Kabil O, Taoka S, LoBrutto R, Shoemaker R and Banerjee R. (2001) The pyridoxal phosphate binding sites are similar in human hemedependent and yeast heme-independent cystathionine β-synthases. Evidence from 31P NMR and pulsed EPR spectroscopy that the heme and the PLP cofactors are not proximal in the human enzyme. *J Biol Chem* 276, 19350–19355.

Kabil O, Zhou Y and Banerjee R. (2006). Human cystathionine β-synthase is a target for sumoylation. *Biochemistry* 45, 13528–36.

Kay HD. (1928) The phosphatases of mammalian tissues: Pyrophosphatase. *Biochem J* 22, 1446-8.

Kennan A, Aherne A, Palfi A, Humphries M, McKee A, Stitt A, Simpson DA, Demtroder K, Orntoft T, Ayuso C, Kenna PF, Farrar GJ and Humphries P. (2002) Identification of an IMPDH1 mutation in autosomal dominant retinitis pigmentosa (RP10) revealed following comparative microarray analysis of transcripts derived from retinas of wild-type and Rho(-/-) mice. *Hum Mol Genet* 11, 547–557.

Kennan A, Aherne A and Humphries P. (2005) Light in retinitis pigmentosa. *Trends Genet* 21, 103–110

Kery V, Poneleit L and Kraus J. (1998) Trypsin cleavage of human cystathionine β-synthase into an evolutionarily conserved active core: structural and functional consequences *Arch Biochem Biophys* 355, 222–232.

Klemme JH. (1976) Regulation of intracellular pyrophosphatase-activity and conservation of the phosphoanhydride-energy of inorganic pyrophosphate in microbial metabolism. *Z Naturforsch C* 31, 544-50.

Kluijtmans LA, Boers GH, Stevens EM, Renier WO, Kraus JP, Trijbels FJ, van den Heuvel LP and Blom HJ. (1996) Defective cystathionine β-synthase regulation by S-adenosylmethionine in

a partially pyridoxine responsive homocystinuria patient. *J Clin Invest* 98, 285-9.

Konopka MA, White SA and Young TW. (2002) Bacillus subtilis inorganic pyrophosphatase: the C-terminal signature sequence is essential for enzyme activity and conformational integrity. *Biochem Biophys Res Commun* 290, 806-12.

Koutmos M, Kabil O, Smith JL and Banerjee R. (2010) Structural basis for substrate activation and regulation by cystathionine β-synthase (CBS) domains in cystathionine β-synthase. *Proc Natl Acad Sci U S A* 107, 20958-63.

Kornberg A. (1962) On the metabolic significance of phosphorylic and pyrophosphorylic reactions. *Horizons in biochemistry* p. 251-64. Eds. Kasha, M. and Pullman, B. Academic Press, New York.

Kornberg A (1995) Inorganic polyphosphate: toward making a forgotten polymer unforgettable. *J Bacteriol* **177**: 491–496.

Kornberg A, Rao NN and Ault-Riché D. (1999) Inorganic polyphosphate: a molecule of many functions. *Annu Rev Biochem* 68, 89-125.

Kuhla B, Kuhla S, Rudolph PE, Albrecht D and Metges CC. (2007). Proteomics analysis of hypothalamic response to energy restriction in dairy cows. *Proteomics* 7, 3602-3617.

Kuhn NJ & Ward S. (1998) Purification, properties, and multiple forms of a manganese-activated inorganic pyrophosphatase from *Bacillus subtilis*. *Arch Biochem Biophys* 354, 47-56.

Kuhn NJ, Wadeson A, Ward S, Young TW. (2000) *Methanococcus jannaschii* ORF mj0608 codes for a class C inorganic pyrophosphatase protected by Co^{2+} or Mn^{2+} ions against fluoride inhibition. *Arch Biochem Biophys* 379, 292-298.

Kulaev IS, Vagabov VM and Kulakovskaya TV. (2004) *The Biochemistry of Inorganic Polyphosphates*, John Wiley & Sons, Ltd., West Sussex.

Kummerfeld SK & Teichmann SA. (2009) Protein domain organisation: adding order. *BMC Bioinformatics* 10, 39.

Lahti R. (1983) Microbial inorganic pyrophosphatases. *Microbiol Rev* 47, 169-78.

Letunic I, Doerks T and Bork P. (2009) SMART 6: recent updates and new developments. *Nucleic Acids Res* 37(Database issue), D229-32.

Lexander H, Palmberg C, Auer G, Hellstrom M, Franzen B, Jornvall H and Egevad L. (2005). Proteomic analysis of protein expression in prostate cancer. *Anal Quant Cytol Histol* 27, 263-272.

Lindner SN, Knebel S, Wesseling H, Schoberth SM and Wendisch VF. (2009) Exopolyphosphatases PPX1 and PPX2 from *Corynebacterium glutamicum*. *Appl Envir Microbiol* 75, 3161–3170.

Link JO & Straub K. (1996) Trapping of an IMP dehydrogenase-substrate covalent intermediate by mycophenolic acid. *J Am Chem Soc* 118, 2091-2092.

Lucas M, Encinar JA, Arribas EA, Oyenarte I, García IG, Kortazar D, Fernández JA, Mato JM, Martínez-Chantar ML and Martínez-Cruz LA. (2010) Binding of S-methyl-5'-thioadenosine and S-adenosyl-L-methionine to protein MJ0100 triggers an open-to-closed conformational change in its CBS motif pair. *J Mol Biol* 396, 800-20.

Maclean KN, Gaustadnes M, Oliveriusová J, Janosík M, Kraus E, Kozich V, Kery V, Skovby F, Rüdiger N, Ingerslev J, Stabler SP, Allen RH and Kraus JP. (2002) High homocysteine and thrombosis without connective tissue disorders are associated with a novel class of cystathionine β-synthase (CBS) mutations. *Hum Mutat* 19, 641-55.

Maclean KN, Kraus E, Kraus JP. (2004) The dominant role of Sp1 in regulating the cystathionine beta-synthase -1a and -1b promoters facilitates potential tissue-specific regulation by Kruppel-like factors. *J Biol Chem* 279, 8558-66.

Mansurova SE, Ermakova SA and Kulaev IS. (1976) Extramitochondrial energy-dependent synthesis of inorganic pyrophosphate in yeast. *Biokhimiya* 41, 1716-1719.

Markovic S & Dutzler R. (2007) The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface. *Structure* 15, 715-25.

Matsuno K, Miyamoto T, Yamaguchi K, Abu Sayed M, Kajiwara T and Hatano S. (1995) Identification of DNA-binding proteins changed after induction of sporulation in *Bacillus cereus*. *Biosci Biotechnol Biochem* 59, 231–235

McLean JE, Hamaguchi N, Belenky P, Mortimer SE, Stanton M and Hedstrom L. (2004) Inosine

5'-monophosphate dehydrogenase binds nucleic acids in vitro and in vivo. *Biochem J* 379, 243–251

Meier M, Janosik M, Kery V, Kraus JP and Burkhard P. (2001) Structure of human cystathionine β-synthase: a unique pyridoxal 5-phosphate-dependent heme protein. *EMBO J* 20, 3910–3916.

Merckel MC, Fabrichniy IP, Salminen A, Kalkkinen N, Baykov AA, Lahti R and Goldman A. (2001) Crystal structure of *Streptococcus mutans* pyrophosphatase: a new fold for an old mechanism. *Structure* 9, 289-97.

Meyer S & Dutzler R. (2006) Crystal structure of the cytoplasmic domain of the chloride channel ClC-0. *Structure* 14, 299-307.

Meyer S, Savaresi S, Forster IC and Dutzler R. (2007) Nucleotide recognition by the cytoplasmic domain of the human chloride transporter ClC-5. *Nat Struct Mol Biol* 14, 60-7.

Miles EW & Kraus JP. (2004) Cystathionine β-synthase: structure, function, regulation, and location of homocystinuria-causing mutations. *J Biol Chem* 279, 29871–29874.

Miller MD, Schwarzenbacher R, von Delft F, Abdubek P, Ambing E, Biorac T, Brinen LS, Canaves JM, Cambell J, Chiu HJ, Dai X, Deacon AM, DiDonato M, Elsliger MA, Eshagi S, Floyd R, Godzik A, Grittini C, Grzechnik SK, Hampton E, Jaroszewski L, Karlak C, Klock HE, Koesema E, Kovarik JS, Kreusch A, Kuhn P, Lesley SA, Levin I, McMullan D, McPhillips TM, Morse A, Moy K, Ouyang J, Page R, Quijano K, RobbA, Spraggon G, Stevens RC, van den Bedem H, Velasquez J, Vincent J, Wang X, West B, Wolf G, Xu Q, Hodgson KO, Wooley J and Wilson IA. (2004) Crystal structure of a tandem cystathionine-β-synthase (CBS) domain protein (TM0935) from Thermotoga maritima at 1.87 A resolution. *Proteins* 57, 213-7.

Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, Best AA, Cande WZ, Chen F, Cipriano MJ, Davids BJ, Dawson SC, Elmendorf HG, Hehl AB, Holder ME, Huse SM, Kim UU, Lasek-Nesselquist E, Manning G, Nigam A, Nixon JE, Palm D, Passamaneck NE, Prabhu A, Reich CI, Reiner DS, Samuelson J, Svard SG and Sogin ML. (2007) Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia. Science* 317, 1921-6.

Mortimer SE & Hedstrom L. (2005) Autosomal dominant retinitis pigmentosa mutations in inosine 5'-monophosphate dehydrogenase type I disrupt nucleic acid binding. *Biochem. J.* 390, 41–47.

Mortimer SE, Xu D, McGrew D, Hamaguchi N, Lim HC, Bowne SJ, Daiger SP and Hedstrom L. (2008) IMP dehydrogenase type 1 associates with polyribosomes translating rhodopsin mRNA. *J Biol Chem* 283, 36354–36360.

Mudd SH, Finkelstein JD, Irreverre F and Laster L. (1964) Homocystinuria: an enzymatic defect. *Science* 27, 1443-1445.

Mudd SH, Finkelstein JD, Irreverre F and Laster L. (1965) Transsulfuration in mammals. Microassays and tissue distributions of three enzymes of the pathway. *J Biol Chem* 240, 4382–4392.

Murzin AG. (1998) How far divergent evolution goes in proteins. *Curr Opin Struct Biol* 8, 380-7.

Natsumeda Y, Ikegami T, Murayama K and Weber G. (1988) De novo guanylate synthesis in the commitment to replication in hepatoma 3924A cells. *Cancer Res* 48, 507–511.

Natsumeda Y, Ohno S, Kawasaki H, Konno Y, Weber G and Suzuki K. (1990) Two distinct cDNAs for human IMP dehydrogenase. *J Biol Chem* 265, 5292-5.

Nelson DL & Cox MM. (2000) "Lehninger Principles of biochemistry", 3rd edition. Worth Publishers, 41 Madison avenue, New York, NY 10010

Nimmesgern E, Fox T, Fleming MA and Thomson JA. (1996) Conformational changes and stabilization of inosine 5'-monophosphate dehydrogenase associated with ligand binding and inhibition by mycophenolic acid. *J Biol Chem* 271(32):19421-7.

Nimmesgern E, Black J, Futer O, Fulghum JR, Chambers SP, Brummel CL, Raybuck SA and Sintchak MD. (1999) Biochemical analysis of the modular enzyme inosine 5'-monophosphate dehydrogenase. *Protein Expr Purif* 17, 282-9.

Ogasawara N. (2000) Systematic function analysis of Bacillus subtilis genes. *Res Microbiol* 151, 129-34.

Oksanen E, Ahonen AK, Tuominen H, Tuominen V, Lahti R, Goldman A and Heikinheimo P. (2007) A complete structural description of the

catalytic cycle of yeast pyrophosphatase. *Biochemistry* 46, 1228-39.

Oliveira SM, Ehtisham J, Redwood CS, Ostman-Smith I, Blair EM and Watkins H. (2003) Mutation analysis of AMP-activated protein kinase subunits in inherited cardiomyopathies: implications for kinase function and disease pathogenesis. *J Mol Cell Cardiol* 35, 1251-5.

Orengo CA. (1999) CORA--topological fingerprints for protein structural families. *Protein Sci* 8, 699–715.

Orengo CA & Thornton JM. (2005) Protein families and their evolution – A structural perspective. *Annu Rev Biochem* 74, 867-900.

Pankiewicz KW & Goldstein BM. (2003) *Inosine Monophosphate Dehydrogenase: A Major Therapeutic Target*, Oxford University Press, Washington, D. C.

Pankiewicz KW, Patterson SE, Black PL, Jayaram HN, Risal D, Goldstein BM, Stuyver LJ and Schinazi RF. (2004) Cofactor mimics as selective inhibitors of NAD-dependent inosine monophosphate dehydrogenase (IMPDH) - the major therapeutic target. *Curr Med Chem* 11, 887-900.

Parfenyev AN, Salminen A, Halonen P, Hachimori A, Baykov AA and Lahti R. (2001) Quaternary structure and metal ion requirement of family II pyrophosphatases from *Bacillus subtilis*, *Streptococcus gordonii*, and *Streptococcus mutans*. *J Biol Chem* 276, 24511-8.

Peller L. (1976) On the free-energy changes in the synthesis and degradation of nucleic acids. *Biochemistry* 15, 141-6.

Peña-Münzenmayer G, Catalán M, Cornejo I, Figueroa CD, Melvin JE, Niemeyer MI, Cid LP and Sepúlveda FV. (2005) Basolateral localization of native ClC-2 chloride channels in absorptive intestinal epithelial cells and basolateral sorting encoded by a CBS-2 domain di-leucine motif. *J Cell Sci* 118, 4243-52.

Picollo A & Pusch M. (2005) Chloride/proton antiporter activity of mammalian CLC proteins ClC-4 and ClC-5. *Nature* 436 ,420-3.

Pimkin M & Markham GD. (2008) The CBS subdomain of inosine 5'-monophosphate dehydrogenase regulates purine nucleotide turnover. *Mol Microbiol* 68, 342–359

Pimkin M & Markham GD. (2009) Inosine 5'-monophosphate dehydrogenase. *Adv Enzymol Relat Areas Mol Biol* 76, 1–53

Pohjanjoki P, Fabrichniy IP, Kasho VN, Cooperman BS, Goldman A, Baykov AA and Lahti R. (2001) Probing essential water in yeast pyrophosphatase by directed mutagenesis and fluoride inhibition measurements. *J Biol Chem* 276, 434-41.

Purich DL. (2010) *Enzyme Kinetics: Catalysis and control*, Elsevier Inc, London, UK, p. 685-728.

Pusch M, Ludewig U, Rehfeldt A and Jentsch TJ. (1995) Gating of the voltage-dependent chloride channel CIC-0 by the permeant anion. *Nature* 373, 527-31.

Rajagopal L, Clancy A and Rubens CE. (2003). A eukaryotic type serine/threonine kinase and phosphatase in *Streptococcus agalactiae* reversibly phosphorylate an inorganic pyrophosphatase and affect growth, cell segregation, and virulence. *J Biol Chem* 278, 14429-14441.

Ranea JA, Buchan DW, Thornton JM and Orengo CA. (2004) Evolution of protein superfamilies and bacterial genome size. *J Mol Biol*.336, 871–87.

Rantanen MK, Lehtiö L, Rajagopal L, Rubens CE and Goldman A. (2007) Structure of the *Streptococcus agalactiae* family II inorganic pyrophosphatase at 2.80 A resolution. *Acta Crystallogr D Biol Crystallogr* 63, 738-43.

Rao NN, Gómez-García MR and Kornberg A. (2009) Inorganic polyphosphate: essential for growth and survival. *Annu Rev Biochem* 78, 605–647.

Rudolph MJ, Amodeo GA, Iram SH, Hong SP, Pirino G, Carlson M and Tong L. (2007) Structure of the Bateman2 domain of yeast Snf4: dimeric association and relevance for AMP binding. *Structure* 15, 65–74.

Rychkov GY, Pusch M, Astill DS, Roberts ML, Jentsch TJ and Bretag AH. (1996) Concentration and pH dependence of skeletal muscle chloride channel ClC-1. *J Physiol* 497, 423-35.

Schröder HC & Müller WEG. (1999) Inorganic polyphosphates: biochemistry, biology, biotechnology. Progress in molecular and subcellular biology, vol. 23. Springer, Berlin, Germany.

Scott JW, Hawley SA, Green KA, Anis M, Stewart G, Scullion GA, Norman DG and Hardie DG. (2004) CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. *J Clin Invest* 113, 274–284.

Sen S & Banerjee R. (2007) A pathogenic linked mutation in the catalytic core of human cystathionine β-synthase disrupts allosteric regulation and allows kinetic characterization of a full-length dimer. *Biochemistry* 46, 4110–4116.

Serrano A, Perez-Castineira JR, Baltscheffsky M and Baltscheffsky H. (2007) H^+-PPases: yesterday, today and tomorrow. *IUBMB Life* 59, 76-83.

Shakhnovich B, Dokholyan N, DeLisi C and Shakhnovich E. (2003) Functional fingerprints of folds: evidence for correlated structure-function evolution. *J Mol Biol* 326, 1–9.

Shan X & Kruger WD. (1998) Correction of disease causing CBS mutations in yeast, *Nat. Genet.* 19, 91–93.

Shintani T, Uchiumi T, Yonezawa T, Salminen A, Baykov AA, 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, 263-6.

Shizawa N, Uchiumi T, Taguchi J, Kisseleva NA, Baykov AA, Lahti R and Hachimori A. (2001) Directed mutagenesis studies of the C-terminal fingerprint region of *Bacillus subtilis* pyrophosphatase. *Eur J Biochem* 268, 5771-5.

Shu Q & Nair V. (2008) Inosine monophosphate dehydrogenase (IMPDH) as a target in drug discovery. *Med Res Rev* 28, 219–232.

Sintchak MD, Fleming MA, Futer O, Raybuck SA, Chambers SP, Caron PR, Murcko M and Wilson KP. (1996) Structure and mechanism of inosine monophosphate dehydrogenase in complex with the immunosuppressant mycophenolic acid. *Cell* 85, 921-930.

Sivula T, Salminen A, Parfenyev AN, Pohjanjoki P, Goldman A, Cooperman BS, Baykov AA and Lahti R. (1999) Evolutionary aspects of inorganic pyrophosphatase. *FEBS Lett* 454, 75-80.

Skovby F, Kraus JP and Rosenberg LE. (1984) Homocystinuria: biogenesis of cystathionine β-synthase subunits in cultured fibroblasts and in an in vitro translation system programmed with fibroblast messenger RNA. *Am J Hum Genet* 36, 452–459.

Springs B, Welsh KM and Cooperman BS. (1981) Thermodynamics, kinetics, and mechanism in yeast inorganic pyrophosphatase catalysis of inorganic pyrophosphate: inorganic phosphate equilibration. *Biochemistry* 20, 6384-91.

Stapleton D, Mitchelhill KI, Gao G, Widmer J, Michell BJ, Teh T, House CM, Fernandez CS, Cox T, Witters LA and Kemp BE. (1996) Mammalian AMP-activated protein kinase subfamily. J Biol Chem 271, 611–4.

Suda M, Nakagawa H and Kimura H. (1973) Cystathionine synthase. in *Methods Enzymol* 17, 454–458.

Tammenkoski M, Moiseev VM, Lahti M, Ugochukwu E, Brondijk TH, White SA, Lahti R and Baykov AA. (2007) Kinetic and mutational analyses of the major cytosolic exopolyphosphatase from *Saccharomyces cerevisiae*. *J Biol Chem* 282, 9302-11.

Tammenkoski M, Koivula K, Cusanelli E, Zollo M, Steegborn C, Baykov AA and Lahti, R. (2008) Human metastasis regulator protein h-prune is a short-chain exopolyphosphatase. *Biochemistry* 47, 9707-9713.

Tanaka Y, Hattori M, Fukai S, Ishitania R, Nureki O. (2007) Crystallization and preliminary X-ray diffraction analysis of the cytosolic domain of the Mg^{2+} transporter MgtE. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 63, 678-81.

Taoka S, Widjaja L and Banerjee R. (1999) Assignment of enzymatic functions to specific regions of the PLP-dependent heme protein cystathionine β-synthase. *Biochemistry* 38, 13155–13161.

Taoka S & Banerjee R. (2001) Characterization of NO binding to human cystathionine β-synthase: possible implications of the effects of CO and NO binding to the human enzyme. *J Inorg Biochem* 87, 245–251.

Taoka S, Lepore BW, Kabil O, Ojha S, Ringe D and Banerjee R. (2002) Human cystathionine β-synthase is a heme sensor protein. Evidence that the redox sensor is heme and not the vicinal cysteines in the CXXC motif seen in the crystal structure of the truncated enzyme. *Biochemistry* 41, 10454–10461.

Tono H & Kornberg A. (1967) Biochemical studies of bacterial sporulation. 3. Inorganic pyrophosphatase of vegetative cells and spores of *Bacillus subtilis*. *J Biol Chem* 242, 2375-82.

Townley R & Shapiro L. (2007) Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science* 315, 1726–1729.

Uchida S. (2000) In vivo role of CLC chloride channels in the kidney. *Am J Physiol Renal Physiol* 279, F802-8.

Ueda A, Kathiresan A, Bennett J and Takabe T. (2006). Comparative transcriptome analyses of barley and rice under salt stress. *Theor Appl Genet* 112, 1286-1294.

Ugochukwu E, Lovering AL, Mather OC, Young TW, White SA. (2007) The crystal structure of the cytosolic exopolyphosphatase from Saccharomyces cerevisiae reveals the basis for substrate specificity. *J Mol Biol* 371, 1007-21.

Urreizti R, Balcells S, Rodés M, Vilarinho L, Baldellou A, Couce ML, Muñoz C, Campistol J, Pintó X, Vilaseca MA and Grinberg D. (2003) Spectrum of CBS mutations in 16 homocystinuric patients from the Iberian Peninsula: high prevalence of T191M and absence of I278T or G307S. *Hum Mutat* 22, 103.

Veech RL, Cook GA and King MT. (1980) Relationship of free cytoplasmic pyrophosphate to liver glucose content and total pyrophosphate to cytoplasmic phosphorylation potential. *FEBS Lett* 117, Suppl:K65-72.

Verham R, Meek TD, Hedstrom L and Wang CC. (1987) Purification, characterization, and kinetic analysis of inosine 5'-monophosphate dehydrogenase of Tritrichomonas foetus. Mol Biochem Parasitol. *Mol Biochem Parasitol* 24, 1-12.

Wang W & Hedstrom L. (1997) Kinetic mechanism of human inosine 5'-monophosphate dehydrogenase type II: random addition of substrates and ordered release of products. *Biochemistry* 36, 8479-83.

Wang CY, Shi JD, Yang P, Kumar PG, Li QZ, Run QG, Su YC, Scott HS, Kao KJ and She JX. (2003a) Molecular cloning and characterization of a novel gene family of four ancient conserved domain proteins (ACDP). *Gene* 306, 37-44.

Wang L, Fraley CD, Faridi J, Kornberg A and Roth RA. (2003b) Inorganic polyphosphate stimulates mammalian TOR, a kinase involved in the proliferation of mammary cancer cells. *Proc Natl Acad Sci USA* 100, 11249–11254.

Wang XQ, Deriy LV, Foss S, Huang P, Lamb FS, Kaetzel MA, Bindokas V, Marks JD, Nelson DJ. (2006) CLC-3 channels modulate excitatory synaptic transmission in hippocampal neurons. *Neuron* 52, 321-33.

Wang J, Zeevi A, Webber S, Girnita DM, Addonizio L, Selby R, Hutchinson IV and Burckart GJ. (2007) A novel variant L263F in human inosine 50-monophosphate dehydrogenase 2 is associated with diminished enzyme activity. *Pharmacogenet Genomics* 17, 283–290.

Weber G. (1983) Biochemical strategy of cancer cells and the design of chemotherapy: G. H. A. Clowes Memorial Lecture. *Cancer Res* 43, 3466-3492.

Weiner JBBE 3rd & Moore AD. (2008) Just how versatile are domains? *BMC Evol Biol* 8, 285.

Woehl EU, Tai CH, Dunn MF and Cook PF. (1996) Formation of the alpha-aminoacrylate immediate limits the overall reaction catalyzed by O-acetylserine sulfhydrylase. *Biochemistry* 35, 4776–4783.

Woese C. (2002) On the evolution of cells. *Proc. Natl. Acad. Sci.USA* 99, 8742–47.

Wurst H & Kornberg A. (1994) A soluble exopolyphosphatase of *Saccharomyces cerevisiae*. Purification and characterization. *J Biol Chem* 269, 10996–11001.

Xiang B, Taylor JC and Markham GD. (1996) Monovalent cation activation and kineticmechanism of inosine 5'-monophosphate dehydrogenase. *J Biol Chem 271*, 1435-1440.

Xiang B & Markham GD. (1997) Probing the mechanism of inosine monophosphate dehydrogenase with kinetic isotope effects and NMR determination of the hydride transfer stereospecificity. *Arch Biochem Biophys* 348, 378-82.

Xu QS, Jancarik J, Lou Y, Kuznetsova K, Yakunin AF, Yokota H, Adams P, Kim R and Kim SH. (2005) Crystal structures of a phosphotransacetylase from Bacillus subtilis and its complex with acetyl phosphate. *J Struct Funct Genomics* 6, 269-79.

Xu D, Cobb G, Spellicy CJ, Bowne SJ, Daiger SP and Hedstrom L. (2008) Retinal isoforms of inosine 5'-monophosphate dehydrogenase type 1 are poor nucleic acid binding proteins. *Arch Biochem Biophys* 472, 100–104

Young TW, Kuhn NJ, Wadeson A, Ward S, Burges D and Cooke GD. (1998) Bacillus subtilis ORF yybQ encodes a manganese-dependent inorganic pyrophosphatase with distinctive properties: the first of a new class of soluble pyrophosphatase? *Microbiology* 144, 2563-71.

Zhang R, Evans G, Rotella FJ, Westbrook EM, Beno D, Huberman E, Joachimiak A and Collart FR. (1999) Characteristics and crystal structure of bacterial inosine-5'-monophosphate dehydrogenase. *Biochemistry* 38, 4691–4700.

Zifarelli G & Pusch M. (2007) CLC chloride channels and transporters: a biophysical and physiological perspective. *Rev Physiol Biochem Pharmacol* 158, 23-76.

Zifarelli G, Murgia AR, Soliani P and Pusch M. (2008) Intracellular proton regulation of ClC-0. *J Gen Physiol* 132, 185-98.

Zifarelli G & Pusch M. (2009) Conversion of the 2 Cl(-)/1 H+ antiporter ClC-5 in a NO3(-)/H+ antiporter by a single point mutation. *EMBO J* 28, 175-82.

Zyryanov AB, Shestakov AS, Lahti R and Baykov AA. (2002) Mechanism by which metal cofactors control substrate specificity in pyrophosphatase. *Biochem J* 367, 901-6.

Zyryanov AB, Tammenkoski M, Salminen A, Kolomiytseva GY, Fabrichniy IP, Goldman A, Lahti R and Baykov AA. (2004a) Site-specific effects of zinc on the activity of family II pyrophosphatase. *Biochemistry* 43, 14395-402.

Zyryanov AB, Vener AV, Salminen A, Goldman A, Lahti R and Baykov AA. (2004b) Rates of elementary catalytic steps for different metal forms of the family II pyrophosphatase from *Streptococcus gordonii*. *Biochemistry* 2004 43, 1065-74.

Zyryanov AB, Lahti R and Baykov AA. (2005) Inhibition of family II pyrophosphatases by analogs of pyrophosphate and phosphate. *Biochemistry (Mosc)* 70, 908-12.