



**HAL**  
open science

# Mutational analysis of residues in the regulatory CBS domains of *Moorella thermoacetica* pyrophosphatase corresponding to disease-related residues of human proteins

Joonas Jämsen, Heidi Tuominen, Alexander A Baykov, Reijo Lahti

## ► To cite this version:

Joonas Jämsen, Heidi Tuominen, Alexander A Baykov, Reijo Lahti. Mutational analysis of residues in the regulatory CBS domains of *Moorella thermoacetica* pyrophosphatase corresponding to disease-related residues of human proteins. *Biochemical Journal*, 2011, 433 (3), pp.497-504. 10.1042/BJ20101204 . hal-00558096

**HAL Id: hal-00558096**

**<https://hal.science/hal-00558096>**

Submitted on 21 Jan 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Mutational analysis of residues in the regulatory CBS domains of  
*Moorella thermoacetica* pyrophosphatase corresponding to  
disease-related residues of human proteins**

Joonas JÄMSEN\*, Heidi TUOMINEN\*, Alexander A. BAYKOV†<sup>1</sup> and Reijo LAHTI\*<sup>1</sup>

\*Department of Biochemistry and Food Chemistry, University of Turku, FIN-20014 Turku, Finland, and †A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia

Short title: CBS-pyrophosphatase of *M. thermoacetica*

Abbreviations used: CBS, cystathionine  $\beta$ -synthase; *mt*CBS-PPase, CBS domain-containing PPase from *Moorella thermoacetica*; PPase, inorganic pyrophosphatase.

<sup>1</sup> Correspondence may be addressed to either of these authors (email [reijo.lahti@utu.fi](mailto:reijo.lahti@utu.fi) or [baykov@genebee.msu.su](mailto:baykov@genebee.msu.su)).

Please address all prepublication correspondence to Professor Reijo Lahti, Tel.: 358-2-333-6845; Fax: 358-2-333-6860, email: [reijo.lahti@utu.fi](mailto:reijo.lahti@utu.fi).

*Moorella thermoacetica* pyrophosphatase (*mtCBS-PPase*) contains a pair of CBS domains that strongly bind adenine nucleotides, thereby regulating enzyme activity. Eight residues associated with the CBS domains of *mtCBS-PPase* were screened to explore for possible associations with regulation of enzyme activity. The majority of substitutions (V99A, R168A, Y169A, Y169F, Y188A, and H189A) enhanced the catalytic activity of *mtCBS-PPase*, two substitutions (R170A and R187G) decreased activity, and one (K100G) had no effect. AMP-binding affinity was markedly decreased in the V99A, R168A, Y169A, and elevated in the R187G and H189A mutant proteins. Remarkably, the R168A and Y169A substitutions changed the effect of AMP from inhibition to activation. The stoichiometry of AMP binding increased from one to two AMP molecules per CBS domain pair in the Y169F, R170A, R187G, and Y188A variants. The ADP-binding affinity decreased in three and increased in four mutant proteins. These findings identify residues determining the strength and selectivity of nucleotide binding, as well as the direction (inhibition or activation) of the subsequent effect. The data suggest that mutations in human CBS domain-containing proteins can be translated into a bacterial context. Furthermore, our data support the hypothesis that the CBS domains act as an “internal inhibitor” of *mtCBS-PPase*.

Key words: inorganic pyrophosphatase, CBS domain, *Moorella thermoacetica*, allosteric regulation, mutagenesis, adenine nucleotides.

Accepted Manuscript

## INTRODUCTION

In *Bacilli*, *Clostridia*, and other bacterial lineages, including several human pathogens, the intracellular  $PP_i$  level is controlled by family II pyrophosphatases (PPases) (EC 3.6.1.1) that convert  $PP_i$  to  $P_i$  [1,2]. Approximately one-quarter (67) of known family II PPase sequences, mostly from *Clostridia* and sulphate-reducing  $\delta$ -proteobacteria, contain a large insert (~100-125 residues) composed of two CBS domains that form a Bateman fold [3]. This structure, originally identified in cystathionine- $\beta$ -synthase [3], binds regulatory adenine nucleotides and S-adenosyl methionine and is present in a variety of proteins [4,5]. We recently expressed the CBS domain-containing PPase of *Moorella thermoacetica* (*mtCBS-PPase*) for the first time, and analysed enzyme regulation by kinetic and binding measurements [6,7]. The enzyme binds adenine nucleotides in the submicromolar range, is activated by ATP, and inhibited by AMP and ADP [6]. Stopped-flow kinetic measurements using a fluorescent AMP analogue revealed that nucleotide binding to *mtCBS-PPase* occurs in three steps with relaxation times from 0.01 to 100 s, implying conformational changes in the complex [7]. Nucleotide binding was modulated by the substrate ( $PP_i$ ), which additionally induced enzyme conversion to a more active form within a timescale of minutes [7]. Recently, we expressed the CBS domain containing PPase of *Clostridium perfringens* (*cpCBS-PPase*) and determined the structure of the regulatory region comprising two CBS domains [8].

According to the Pfam database [9], CBS domains have been identified in nearly 30,000 protein sequences from all kingdoms of life. In addition to the role played by CBS domains in cystathionine- $\beta$ -synthase [3], such domains are implicated in the control of intracellular trafficking by the chloride channel, CLC-5 [10]; gating of the osmoregulatory transporter, OpuA [11,12]; protein phosphorylation by AMP-dependent protein kinase [4];  $Mg^{2+}$  transfer by the transporter, MgtE [13]; plant vacuolar nitrate transport by the transporter, AtCLCa [14]; and adenylyate nucleotide synthesis (IMP dehydrogenase) [15]. Single mutations in the

CBS domains of these proteins are linked to a variety of human hereditary disorders, including homocystinuria, Wolff-Parkinson-White syndrome, retinitis pigmentosa, and Bartter syndrome [4,5]. However, little is known about the mechanism underlying the transduction of regulatory inhibition or activation signals from CBS domains to the substrate-binding regions of these proteins. Most crystal structures of CBS domains have been obtained for small proteins of unknown function containing only CBS domains, or separate regulatory regions of multi-domain proteins, which reveal no significant changes upon binding of nucleotides. Although the crystal structures of several complex proteins containing regulatory CBS domains are available [16–18], pairs of structures for activated/inhibited or nucleotide-bound/nucleotide-free forms have not been solved for any protein to date. Large structural differences in CBS domains have been observed in only two instances, dependent on the bound nucleotide: with the MJ0100 protein of unknown function from *Methanocaldococcus jannaschii* [19] and the regulatory part of *cp*CBS-PPase [8]. The *cp*CBS-PPase structures display marked conformational changes upon binding of activator (diadenosine tetraphosphate) compared to inhibitor (AMP) to the regulatory CBS domains, indicating that CBS-PPase is a promising model in which to explore the mechanism of regulation of CBS proteins.

Here, we report the effects of point mutations in the CBS domains of *mt*CBS-PPase on interactions with adenine nucleotides, with a view to identifying residues that determine the strength and selectivity of nucleotide binding as well as the direction of the effect (inhibition or activation). Additionally, we characterize the effects of substitutions on the substrate- and nucleotide-induced activity transitions observed in previous studies [7]. Residues were selected for substitution based on approximate correspondence with physiologically disruptive mutations in other CBS domains and the X-ray structure of homologous *cp*CBS-PPase.

## EXPERIMENTAL PROCEDURES

### Reagents

Restriction enzymes and T4 DNA ligase were obtained from Fermentas (Lithuania) and *Pfu* Turbo DNA polymerase from Stratagene (USA). Oligonucleotides were purchased from MWG Biotech (Germany). AMP, ADP and ATP were from Fluka (Switzerland). [<sup>14</sup>C]AMP was acquired from Moravек Biochemicals and Radiochemicals (USA). All other reagents and kits were obtained from Sigma (USA).

### Site-directed mutagenesis, protein expression, and purification

Mutagenesis of the *mtCBS*-PPase gene was performed in pBluescript SK (Stratagene, USA) with a QuikChange kit (Stratagene). Primer sequences used are shown in Supplementary Table S1. Variant genes were cloned for expression into the pET36b+ plasmid (Novagen, USA) using *Nde*I and *Xho*I restriction sites. Gene integrity and presence of the desired point mutations were confirmed via DNA sequencing. Expression and purification of mutant proteins was performed as described for wild-type *mtCBS*-PPase [6]. Protein concentrations were determined from absorbance at 280 nm ( $\epsilon = 22,920 \text{ M}^{-1} \text{ cm}^{-1}$ , calculated using ProtParam [20]). Protein expression level and purity of the final preparation were analysed on 8–25% SDS-PAGE gels using the Phast system (GE Healthcare, USA).

### Activity measurements

Unless otherwise stated, activity measurements were performed at 25°C in 25–40 ml of 100 mM Mops/KOH buffer (pH 7.2), including 0.1 mM CoCl<sub>2</sub>, 5 mM MgCl<sub>2</sub> and 0.16 mM PP<sub>i</sub>. PP<sub>i</sub> hydrolysis was initiated by adding an aliquot (5–200  $\mu$ l) of diluted enzyme solution to obtain 0.005–5  $\mu$ g/ml final enzyme concentration. P<sub>i</sub> formation was monitored for 20–25 min

using a semi-automatic  $P_i$  analyser [21]. Stock nucleotide solutions used in kinetic measurements were prepared immediately before use, and kept on ice.

### Membrane filtration

[ $^{14}\text{C}$ ]AMP binding to *mt*CBS-PPase was determined with a filtration method [7] using Multiscreen 96-well opaque filtration plates with an IP membrane (Millipore Corp., USA). The membrane was washed with 20% ethanol using buffer supplemented with 1 mM unlabelled AMP, and twice with buffer without AMP. Next, 50  $\mu\text{l}$  of 10  $\mu\text{M}$  *mt*CBS-PPase solution was added and incubated for 10 min. Finally,  $^{14}\text{C}$ -labelled AMP (200–5,000 cpm, 100  $\mu\text{l}$ ) was added, the complete mixture incubated for 10 min, and the wells emptied using a vacuum pump (time to empty about 1 s). Wells were washed twice with 0.1 mL of buffer, 0.2 mL of Ultima Gold XR scintillant added, and the radioactivity adsorbed on the membrane was determined with a Chameleon V multimode microplate reader (Hidex Ltd, Turku, Finland). Bovine serum albumin was added instead of enzyme to estimate non-specific [ $^{14}\text{C}$ ]AMP binding.

### Data analyses

All non-linear least-squares fittings were performed with the program Scientist (MicroMath, USA). Dependence of activity ( $A$ ) on nucleotide concentration ( $[L]$ ) at a fixed substrate concentration was fitted to eqn 1, where  $A_{+L}$  and  $A_{-L}$  represent activities with and without bound nucleotide, respectively, and  $K_{i,\text{app}}$  is the apparent nucleotide binding constant:

$$A = A_{+L} + (A_{-L} - A_{+L}) / (1 + [L] / K_{i,\text{app}}) \quad (1)$$

Time-courses of  $\text{PP}_i$  hydrolysis were fit to eqns 2 and 3 [7], where  $A_0$ ,  $A$  and  $A_{\text{lim}}$

represent velocities at time zero,  $t$ , and infinity, respectively,  $[E]_0$  is the total enzyme concentration, and  $K_m$  is the Michaelis constant. Equation 3 assumes that activity increases with a first-order rate constant,  $k$ , to a constant level due to enzyme transition to a more active form during substrate conversion, and decreases upon substrate depletion.  $A_0$ ,  $A_{lim}$  and  $k$  values were treated as adjustable parameters. The  $K_m$  value, initially set at 10  $\mu\text{M}$  at the initial round of calculation, was subsequently estimated from the calculated dependences of  $A_{lim}$  on  $[S]_0$ , and the calculation cycle repeated with the refined  $K_m$  value until the results converged.

$$d[P]/dt = A[E]_0 \quad (2)$$

$$A = [A_0 + (A_{lim} - A_0)(1 - e^{-kt})](1 + K_m/[S]_0) / \{1 + K_m/([S]_0 - [P])\} \quad (3)$$

## RESULTS

### Selection and production of enzyme variants

Eight residues were selected for substitution in the two CBS domains of *mt*CBS-PPase, based on (a) the positions of disease-causing mutations in CBS proteins (Figure 1A and Supplementary Table S2), (b) the identity of the ligands binding to nucleotides in *cp*CBS-PPase (Figure 1B), and, (c) the residue conservation profiles of CBS-PPases and other CBS proteins (Figure 1A). Residues corresponding to Lys100, Tyr169, and Arg170 are nucleotide ligands in the *cp*CBS-PPase structure (Lys, Tyr, and Ser, respectively) (Figure 1) and are well conserved (61%, 69%, and 87%, respectively) in 67 putative CBS-PPase sequences.

Furthermore, residues at positions 168–170 form part of the RYRN sequence corresponding to the RYSN loop of *cp*CBS-PPase, which undergoes the most significant conformational change when the activator- and inhibitor-bound forms of *cp*CBS-PPase are compared [8].

Residues 187–189 are part of a conserved CBS domain motif. Finally, Val99 is located at the

same relative position in the second CBS domain of *mt*CBS-PPase compared to Y169 of the first domain (Figure 1).

Keeping in mind that replacement of disease-causing residues yields varying effects in different proteins (Figure 1A and Supplementary Table S2), Ala is commonly used as a substituting residue, to eliminate any effect of polar side-chain groups. In addition, Tyr169 was conservatively replaced with Phe. Two Arg residues were mutated to Gly, because such substitutions are known to affect the function of pig AMP-dependent protein kinase  $\gamma$ 3 linked to CBS domains [22]. For substitution of R168, the choice between Ala and Gly was made on the basis of protein solubility, as the R168G variant showed a tendency to aggregate during purification. Expression levels for all *mt*CBS-PPase variants (~ 10% of total cell protein) were similar to that for wild-type enzyme, and variants were easily separated from other proteins, including *E. coli* PPase. The final preparations were at least 90 % pure as estimated by SDS-PAGE analysis (Supplementary Figure S1).

### Catalytic activity

Earlier studies on *mt*CBS-PPase revealed ~2-fold activation by substrate, with a half-time of approximately 40 s, during PP<sub>i</sub> hydrolysis [7]. This effect was preserved in all variants studied, except for the Y169A and R170A proteins (Figure 2). Accordingly, steady-state rate was attained after several minutes of reaction. The corresponding activity ( $A_{lim}$ ) was physiologically more relevant than was initial activity in estimation of the catalytic constants ( $k_{cat,lim}$  values) and Michaelis constants ( $K_m$  values) for variant enzymes. Based on the effect of mutations on  $k_{cat,lim}$ , the variant enzymes were divided into three groups (Table 1): (1) those displaying  $k_{cat,lim}$  values similar to that of the wild-type enzyme (K100G), (2) more active enzymes (V99A, R168A, Y169A, Y169F, Y188A, and H189A), and (3) less active enzymes (R170A and R187G). The effects of mutation on  $K_m$  were smaller, and in most instances, did

not exceed the error of determination.

### Nucleotide inhibition

The effects of substitutions on activity regulation by AMP and ADP were significant (Figure 3). Substrate (PP<sub>i</sub>) concentration was fixed at 160 μM for these measurements, ensuring at least 85% saturation of enzyme active sites, based on the  $K_m$  values listed in Table 1. Thus the observed effects of nucleotides on activity mainly reflect an influence on  $k_{cat,lim}$ .

Remarkably, AMP inhibited wild-type and the majority of variant enzymes, but activated the R168A and Y169A mutants (Figure 3, Table 1). Activation was particularly marked (2.8-fold) for the R168A protein. In keeping with these observations, the maximal degree of inhibition, characterized by the residual activity of an enzyme-AMP complex, was increased in the Y188A and H189A variants compared with wild-type enzyme (Table 1). The AMP-binding affinity, estimated using  $K_{i,app}$ , was not significantly affected in the Y169F and R170A mutants; somewhat decreased in the K100G and Y188A substituents; markedly less in R168A, Y169A, and, particularly, V99A; and significantly elevated in the R187G and H189A mutant proteins.

The effects of ADP were more uniform in that all variants were inhibited by more than 5-fold (Table 1). However, the  $K_{i,app}$  values showed more variation. Specifically, increases were seen for the V99A, Y169A, and R170A variants, and falls were noted for other variants, except for R168A and Y188A. In most instances, the effects of ADP on  $K_{i,app}$  paralleled those of AMP.

### Equilibrium AMP binding

Direct measurements of [<sup>14</sup>C]AMP binding to *mt*CBS-PPase and variants thereof in the absence of substrate were performed using the membrane filtration technique (Figure 4). In an

earlier study, we presented evidence that, in wild-type *mt*CBS-PPase, AMP binding to a dimeric enzyme is characterized by negative cooperativity [7]. Unfortunately, the binding curves obtained from filtration assays showed some scatter, precluding accurate analysis of curve form, which was subsequently approximated as a simple, two-parameter hyperbola. Furthermore, the membrane filtration assay may underestimate the binding affinity because of partial dissociation of the complex during the removal of unbound AMP, which occurs in about 1 s. This may explain why the  $K_d$  values obtained were markedly larger than those of  $K_{i,app}$  derived from kinetic assays. However, regardless of the actual significance of filtration-derived  $K_d$  values, a marked increase upon residue substitution, as observed with V99A, K100G, Y169A, Y188A, and H189A proteins (Table 1), is clearly indicative of altered (most likely, weakened) AMP binding. Another important finding is that the stoichiometry of AMP binding increased from 1 to 2 mol/mol monomer in the Y169F, R170A, R187G, and Y188A variants, implying occupancy of both tandem CBS domains.

### Activity transition

In the absence of substrate, wild-type *mt*CBS-PPase exists as a mixture of two conformations with different activities, and is converted to the more active conformation on the time-scale of the enzyme assay upon addition of substrate [7]. Nucleotides bound to wild-type *mt*CBS-PPase slowed this transition. Similar analyses were performed for the variant enzymes. Fitting of eqns 2 and 3 to the time-courses of  $PP_i$  hydrolysis (Figure 2) yielded both a ratio of initial to final activities ( $k_{cat,0}/k_{cat,lim}$ ) (Table 2) and a rate constant for the substrate-induced transition ( $k_{tr}$ ). For most variants, the  $k_{cat,0}/k_{cat,lim}$  ratio was not significantly different from that of wild-type. The rate constant for transition in the absence of nucleotides was 1.5-3-fold less, compared with wild-type enzyme, and nucleotide addition further decelerated the transition (Figure 5), analogous to what was seen with wild-type PPase. Conversely, the Y169A and

R170A variants exhibited no activity transition in the absence of nucleotides ( $k_{cat,0}/k_{cat,lim} = 1$ ). However, this transition was clearly accelerated in the presence of AMP or ADP. Another interesting finding was that ADP accelerated transition in the Y188A variant. The lack of effect of ADP on the  $k_{tr,lim}$  value of the V99A variant appears to result from weak nucleotide binding (Table 1).

## DISCUSSION

### CBS Domain pair as an internal inhibitor in CBS-PPase

The specific activities of two CBS-PPases characterized to date, those of *M. thermoacetica* [6] and *C. perfringens* [8], are 2-3 orders of magnitude lower than that reported for common family II PPases lacking CBS domains [1,2,23]. Previously, we hypothesized that CBS domains act as “intrinsic inhibitors” in CBS-PPases [6], supported by the observation that the activity of CBS-PPase and other CBS proteins may either decrease or rise, depending on the nature of the bound nucleotide. These allosteric effects are clearly mediated by conformational changes in CBS domains, and it is possible that other factors causing similar structural changes in CBS domains also affect PPase activity. The activation effect is particularly interesting, as it may provide relief against the “inhibition” induced by CBS domains.

The present results provide further support for the above theory. First, five of nine substitutions in CBS domains resulted in enhanced activity (as measured by  $k_{cat,lim}$  values; Table 1). Although functional improvements in proteins after artificial mutation have been documented in the literature, function-disrupting mutations are far more common. This observation reflects the general dogma that functions tailored by nature are difficult to improve upon, but easy to damage. Our findings are in line with this belief if the apparent

“improvement” in function (activation) actually results in a disruption of the auto-inhibition mechanism underlying the regulatory activity of the CBS domains.

Second, the “internal inhibitor” mechanism is supported by reversal of the AMP effect from inhibition to activation in the R168A and Y169A substituents. To the best of our knowledge, no change in the direction of a nucleotide effect after amino acid substitution has ever been previously reported for any CBS protein. The possibility of such reversal implies high plasticity of the regulatory mechanism. Importantly, in the Y169A variant, the basal  $k_{cat,lim}$  value in the absence of nucleotides ( $5.0 \text{ s}^{-1}$ ) significantly exceeded that of wild-type *mtCBS-PPase* ( $2.2 \text{ s}^{-1}$ ). The combined effect of mutation and AMP binding led to an increase in  $k_{cat,lim}$  values by 2.8-3.5-fold in the R168A and Y169A variants, compared with the wild-type basal value. Most constructed mutations thus widen the range of activities attainable in the absence and presence of nucleotides, fully consistent with the “internal inhibitor” hypothesis.

### Nucleotide-binding stoichiometry

Each CBS domain of the Bateman fold contains a potential binding site for nucleotides. However, *cpCBS-PPase* structures [8], most other reported structures [24–26] and those found only in PDB (3LFR, 3FWR, 3FWS, 3FNA, 2YZQ, and 3DDJ) contain only one nucleotide molecule per two CBS domains. However, both sites are occupied by nucleotides in several reported protein structures [19,27–29].

Wild-type *mtCBS-PPase* is similar to *cpCBS-PPase* in that the enzyme binds only one nucleotide per two CBS domains (Table 1). Remarkably, four substitutions in CBS2 elevated the AMP binding stoichiometry of *mtCBS-PPase* to two nucleotides per CBS domain pair (Table 1). Two of the replaced residues (Y169 and R170) correspond to *cpCBS-PPase* AMP ligands, and two (R187 and Y188) belong to a conservative CBS motif. One possible

explanation is that mutations in CBS2 induce conformational changes in this domain, inhibiting interactions with the CBS1-bound nucleotide and facilitating intrinsic nucleotide binding. Interestingly, the  $K_{i,app}$  value was markedly decreased in the R187G variant, indicating tighter binding. This may indicate that, at least in this variant, the  $K_{i,app}$  value is referable to the CBS2-bound nucleotide. Gómez-García *et al.* [29] recently reported unusual binding of AMP to the linker sequence connecting the two CBS domains in a CBS protein from *Methanocaldococcus jannaschii*. However, this type of binding is unlikely in *mtCBS*-PPase, because the linker sequence is relatively short.

### Possible roles of mutated residues in regulation

The effects of substitutions on nucleotide binding affinity are easily explainable in terms of changes in nucleotide contact (group 1 residues), whereas effects on activity may be attributable to closure/opening of the CBS domain interface (group 2 residues), as seen in *cpCBS*-PPase structures [8]. Group 1 residues include K100, Y169 and R170. The finding that AMP binding is weakened and ADP binding strengthened upon K100G substitution is consistent with the theory that K100 interacts with P1 of the nucleotides [8]. Y169 stacks onto the adenine ring in *cpCBS*-PPase structures [8], and alteration to Ala is thus expected to weaken binding, leading to malpositioning of the bound nucleotide, in turn facilitating opening of the CBS domain interface with concomitant enzyme activation. An opposite effect was predicted and indeed confirmed for the Y169F mutant. Based on the *cpCBS*-PPase structure [8], R170 is expected to play a role in ADP P2 binding. Indeed, the  $K_{i,app}$  values obtained for the R170A substitution indicate that this residue is clearly important for ADP, but not AMP, binding.

Group 2 residues, which do not interact with bound nucleotide in the *cpCBS*-PPase structure [8], include R168, R187, and Y188. R168 is located at the CBS domain interface

within the RYRN sequence. The loop adopts markedly different conformations in activated and inhibited *cp*CBS-PPase, and may therefore play a central role in transmitting the regulatory signal to catalytic domains. This hypothesis is confirmed by the finding that mutation of both residues of this loop region reversed the effect of AMP from inhibition to activation. The importance of R187, a highly conserved (90%) residue in CBS-PPase, is explained by the interactions of this residue with the RYRN loop. Interestingly, the conformations adopted by R187 in inhibited and activated *cp*CBS-PPase structures [8] are markedly different. Similar data were obtained for Y188, which is displaced from the domain interface to a hydrophobic cavity when a complex is formed with activating nucleotide. After Y188A substitution, the domain interface is expected to open, resulting in increased activity of the variant enzyme.

In summary, six CBS-domain residues, including four corresponding to disease-related residues of human proteins, are important in adenine nucleotide regulation of CBS-PPase. Together with the previously determined structures of the inhibitor- and activator-bound forms of the regulatory domain of *cp*CBS-PPase, our results significantly clarify the mechanism underlying the regulatory effect of the CBS domain. Elucidation of the detailed mechanism of the cooperation between regulatory and catalytic domains is the next challenging step, and the current results will serve as a basis for this work.

## ACKNOWLEDGEMENT

The authors wish to acknowledge Dr. Georgiy Belogurov for mutant design and the initial concept as well as Juho Vuononvirta and Minna Peippo for expression and purification of the mutant proteins.

This work was supported by Academy of Finland Grant 114706, Russian Foundation for

Basic Research Grant 09-04-00869, and a grant from the Ministry of Education and the Academy of Finland (for the National Graduate School in Informational and Structural Biology).

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101204

## REFERENCES

- 1 Shintani, T., Uchiumi, T., Yonezawa, T., Salminen, A., Baykov, A. A., Lahti, R. and Hachimori, A. (1998) Cloning and expression of a unique inorganic pyrophosphatase from *Bacillus subtilis*: evidence for a new family of enzymes. *FEBS Lett.* **439**, 263–266
- 2 Young, T. W., Kuhn, N. J., Wadson, A., Ward, S., Burges, D. and Cooke, G. D. (1998) *Bacillus subtilis* ORF yybQ encodes a manganese-dependent inorganic pyrophosphatase with distinctive properties: the first of a new class of soluble pyrophosphatase? *Microbiology* **144**, 2563–2571
- 3 Bateman, A. (1997) The structure of a domain common to archaeobacteria and the homocystinuria disease protein. *Trends Biochem. Sci.* **22**, 12–13
- 4 Scott, J. W., Hawley, S. A., Green, K. A., Anis, M., Stewart, G., Scullion, G. A., Norman, D. G. and Hardie, D. G. (2004) CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. *J. Clin. Invest.* **113**, 274–284
- 5 Ignoul, S. and Eggermont, J. (2005) CBS domains: structure, function, and pathology in human proteins. *Am. J. Physiol. Cell Physiol.* **289**, C1369–C1378
- 6 Jämsen, J., Tuominen, H., Salminen, A., Belogurov, G. A., Magretova, N. N., Baykov, A. A. and Lahti, R. (2007) A CBS domain-containing pyrophosphatase of *Moorella thermoacetica* is regulated by adenine nucleotides. *Biochem. J.* **408**, 327–333
- 7 Jämsen, J., Baykov, A. A. and Lahti, R. (2009) Nucleotide- and substrate-induced conformational transitions in the CBS domain-containing pyrophosphatase of *Moorella thermoacetica*. *Biochemistry* **49**, 1005–1013
- 8 Tuominen H., Salminen A., Oksanen E., Jämsen J., Heikkilä O., Lehtiö L., Magretova N.N., Goldman A., Baykov A.A. and 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

- 9 Finn, R. D., Mistry, J., Tate, J., Coghill, P., Heger, A., Pollington, J. E., Gavin, O. L., Guneseckaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res.* **38**, D211–222
- 10 Carr, G., Simmons, N. and Sayer, J. (2003) A role for CBS domain 2 in trafficking of chloride channel CLC-5. *Biochem. Biophys. Res. Comm.* **310**, 600–605
- 11 Bieman-Oldehinkel, E., Mahmood, N. B. N. and Poolman, B. (2006) A sensor for intracellular ionic strength. *Proc. Natl. Acad. Sci. USA* **103**, 10624–10629
- 12 Mahmood, N. A., Biemans-Oldehinkel, E. and Poolman, B. (2009) *J. Biol. Chem.* **284**, 14368–14376
- 13 Ishitani, R., Sugita, Y., Dohmae, N., Furuya, N., Hattori, M. and Nureki, O. (2008) Mg<sup>2+</sup>-sensing mechanism of Mg<sup>2+</sup> transporter MgtE probed by molecular dynamics study. *Proc. Natl. Acad. Sci. USA* **105**, 15393–15398
- 14 De Angeli, A., Moran, O., Wege, S., Filleur, S., Ephritikhine, G., Thomine, S., Barbier-Brygoo, H. and Gambale, F. (2009) ATP binding to the C terminus of the *Arabidopsis thaliana* nitrate/proton antiporter, AtCLCa, regulates nitrate transport into plant vacuoles. *J. Biol. Chem.* **284**, 26526–26532
- 15 Pimkin, M., Pimkina, J., and Markham, G. D. (2009) A regulatory role of the Bateman domain of IMP dehydrogenase in adenylate nucleotide biosynthesis. *J. Biol. Chem.* **284**, 7960–7969
- 16 Zhang, R., Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., Joachimiak, A. and Collart, F. R. (1999) Characteristics and crystal structure of bacterial inosine-5'-monophosphate dehydrogenase. *Biochemistry* **38**, 4691–4700
- 17 Hattori, M., Iwase, N., Furuya, N., Tanaka, Y., Tsukazaki, T., Ishitani, R., Maguire, M.E., Ito, K., Maturana, A. and Nureki, O. (2009) Mg<sup>2+</sup>-dependent gating of bacterial MgtE channel underlies Mg<sup>2+</sup> homeostasis. *EMBO J.* **28**, 3602–3612
- 18 Proudfoot, M., Sanders, S. A., Singer, A., Zhang, R., Brown, G., Binkowski, A., Xu, L., Lukin, J. A., Murzin, A. G., Joachimiak, A., Arrowsmith, C. H., Edwards, A. M., Savchenko,

- A. V. and Yakunin, A. F. (2008) Biochemical and structural characterization of a novel family of cystathionine beta-synthase domain proteins fused to a Zn ribbon-like domain. *J. Mol. Biol.* **375**, 301–315
- 19 Lucas, M., Encinar, J. A., Arribas, E. A., Oyenarte, I., García, I. G., Kortazar, D., Fernández, J. A., Mato, J. M., Martínez-Chantar, M. L. and Martínez-Cruz, L. A. (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**, 800–820
- 20 Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2005) in *The Proteomics Protocols Handbook* (Walker, J. M., Ed.), pp. 571-607, Hmana Press Inc., Totowa
- 21 Baykov, A. A. and Avaeva, S. M. (1981) A simple and sensitive apparatus for continuous monitoring of orthophosphate in the presence of acid-labile compounds. *Anal. Biochem.* **116**, 1–4
- 22 Barnes, B., Marklund, S., Steiler, T., Walter, M., Hjälml, G., Amarger, V., Mahlapuu, M., Leng, Y., Johansson, C., Galuska, D., Lindgren, K., Abrink, M., Stapleton, D., Zierath, J. and Andersson, L. (2004) The 5'-AMP-activated protein kinase gamma3 isoform has a key role in carbohydrate and lipid metabolism in glycolytic skeletal muscle. *J. Biol. Chem.* **279**, 38441–38447
- 23 Parfenyev, A. N., Salminen, A., Halonen, P., Hachimori, A., Baykov, A. 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**, 24511-24518
- 24 Townley, R. and Shapiro, L. (2007) Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science* **315**, 1726–1729
- 25 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 gamma1

subunit of human AMPK in complex with AMP and ZMP. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 587–596

- 26 Meyer, S., Savaresi, S., Forster, I. and Dutzler, R. (2007) Nucleotide recognition by the cytoplasmic domain of the human chloride transporter CIC-5. *Nat. Struct. Mol. Biol.* **14**, 60–67
- 27 King, N., Lee, T., Sawaya, M., Cascio, D. and Yeates, T. (2008) Structures and functional implications of an AMP-binding cystathionine beta-synthase domain protein from a hyperthermophilic archaeon. *J. Mol. Biol.* **380**, 181–192
- 28 Xiao, B., Heath, R., Saiu, P., Leiper, F., Leone, P., Jing, C., Walker, P., Haire, L., Eccleston, J., Davis, C., Martin, S., Carling, D. and Gamblin, S. (2007) Structural basis for AMP binding to mammalian AMP-activated protein kinase. *Nature*, **449**, 496–500
- 29 Gómez-García, I., Oyenarte, I. and Martínez-Cruz, L. A. (2010) The crystal structure of protein MJ1225 from *Methanocaldococcus jannaschii* shows strong conservation of key structural features seen in the eukaryal gamma-AMPK. *J. Mol. Biol.* **399**, 53-70

**Table 1 Catalytic and nucleotide-binding parameters for wild-type and variant *mtCBS-PPases***

Values showing large deviations from wild-type enzyme are in bold. Parameter values for catalytic activity and nucleotide inhibition refer to substrate-activated enzyme.

Enzyme variant	Catalytic activity			AMP inhibition/activation <sup>a</sup>		ADP inhibition <sup>a</sup>		AMP binding <sup>b</sup>	
	$k_{cat,lim}$ ( $s^{-1}$ )	$K_m$ ( $\mu M$ )	$K_{i,app}$ ( $\mu M$ )	Residual activity <sup>c</sup> (%)	$K_{i,app}$ ( $\mu M$ )	Residual activity <sup>c</sup> (%)	$K_d$ ( $\mu M$ )	Binding stoichiometry (mol/mol)	
WT	$2.2 \pm 0.2$	$8 \pm 3$	$0.6 \pm 0.1$	$6 \pm 2$	$1.0 \pm 0.4$	$14 \pm 5$	$13 \pm 1$	$0.94 \pm 0.01$	
V99A	<b><math>7.3 \pm 0.5</math></b>	$3.7 \pm 1.6$	<b>&gt;1000</b>	n.d. <sup>d</sup>	<b>&gt;1000</b>	n.d. <sup>d</sup>	<b>&gt;2000</b>	n.d. <sup>d</sup>	
K100G	$1.9 \pm 0.1$	$13 \pm 4$	$1.7 \pm 0.2$	$4 \pm 2$	$0.13 \pm 0.03$	$8 \pm 3$	<b><math>340 \pm 90</math></b>	$1.1 \pm 0.2$	
R168A	$2.8 \pm 0.3$	$11 \pm 5$	<b><math>11 \pm 5</math></b>	<b><math>280 \pm 30</math></b>	$0.8 \pm 0.2$	$12 \pm 4$	$35 \pm 8$	$1.06 \pm 0.07$	
Y169A	<b><math>5.0 \pm 0.7</math></b>	$9.4 \pm 2.8$	<b><math>8 \pm 3</math></b>	<b><math>122 \pm 2</math></b>	<b>&gt;1000</b>	n.d. <sup>d</sup>	<b><math>1000 \pm 200</math></b>	$0.7 \pm 0.3$	
Y169F	<b><math>7.1 \pm 0.8</math></b>	$12 \pm 6$	$0.39 \pm 0.06$	<2	<b><math>0.017 \pm 0.001</math></b>	<2	$50 \pm 10$	<b><math>1.8 \pm 0.1</math></b>	
R170A	$0.88 \pm 0.06$	$7.9 \pm 2.6$	$0.9 \pm 0.1$	$2 \pm 3$	<b><math>57 \pm 36</math></b>	<20	$17 \pm 1$	<b><math>2.0 \pm 0.1</math></b>	
R187G	$0.98 \pm 0.09$	$5.7 \pm 2.4$	<b><math>0.051 \pm 0.007</math></b>	<4	<b><math>0.009 \pm 0.001</math></b>	<2	$27 \pm 7$	<b><math>1.6 \pm 0.1</math></b>	
Y188A	<b><math>5.8 \pm 0.9</math></b>	$15 \pm 9$	$2.0 \pm 0.7$	<b><math>37 \pm 4</math></b>	$1.8 \pm 0.8$	$5 \pm 1$	<b><math>150 \pm 50</math></b>	<b><math>2.1 \pm 0.2</math></b>	

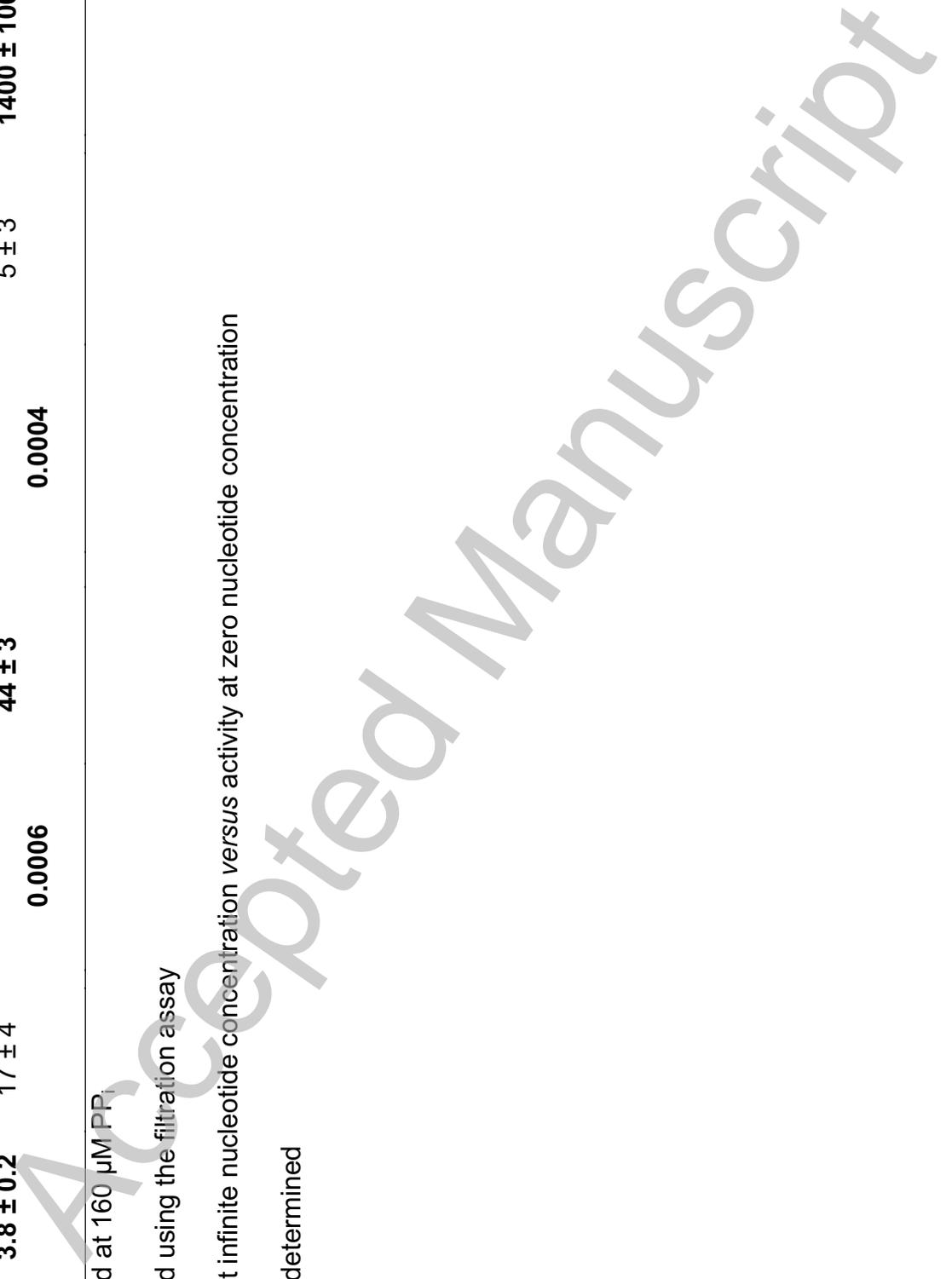
H189A	<b>3.8 ± 0.2</b>	<b>17 ± 4</b>	<b>0.0015 ±</b>	<b>44 ± 3</b>	<b>0.0011 ±</b>	<b>5 ± 3</b>	<b>1400 ± 100</b>	<b>1.3 ± 0.1</b>
			<b>0.0006</b>		<b>0.0004</b>			

<sup>a</sup> Measured at 160 μM PP<sub>i</sub>

<sup>b</sup> Measured using the filtration assay

<sup>c</sup> Activity at infinite nucleotide concentration versus activity at zero nucleotide concentration

<sup>d</sup> n.d., not determined



**Table 2 Parameters describing activity transition in wild-type and variant *mtCBS-PPases***

Values showing large deviations from wild-type enzyme are presented in bold.

Enzyme variant	$k_{cat,0}/k_{cat,lim}^c$	$k_{tr,lim}^b$		
		No nucleotide	AMP	ADP
		$min^{-1}$		
WT	0.60	1.8 ± 0.2	<0.04	<0.04
V99A	0.55	0.81 ± 0.13	<0.04	<b>0.81 ± 0.13</b>
K100G	0.70	1.1 ± 0.1	0.36 ± 0.11	0.43 ± 0.07
R168A	0.45	1.2 ± 0.2	0.29 ± 0.04	<0.04
Y169A	<b>1.0</b>	<b>&lt;0.04</b>	<0.1	<b>0.92 ± 0.12</b>
Y169F	0.65	0.62 ± 0.03	<0.04	0.19 ± 0.02
R170A	<b>1.0</b>	<b>&lt;0.04</b>	<b>0.40 ± 0.02</b>	<b>0.31 ± 0.05</b>
R187G	0.35	0.66 ± 0.05	0.11 ± 0.02	0.094 ± 0.043
Y188A	0.35	0.76 ± 0.12	0.47 ± 0.02	<b>6.5 ± 0.2</b>
H189A	0.40	0.71 ± 0.03	<0.04	<0.04

<sup>a</sup>  $k_{tr,lim}$  values refer to >85% substrate saturation.

<sup>b</sup> Ratio of catalytic activities estimated with eqns 2 and 3 from the initial and final rates of substrate conversion in the absence of nucleotide.

## FIGURE LEGENDS

### Figure 1 Selection of residues for mutation

(A) Sequence alignment of CBS domains 1 and 2 of two characterized CBS-PPases and CBS proteins with disease-causing mutations. Also shown are consensus sequences of CBS-PPases (based on 67 proteins) with 30% and 85% identity. Residue numbering is for *mt*CBS-PPase. Some sequences contain inserts compared to *mt*CBS-PPase; the number of residues in the insert is indicated in parentheses. Residues mutated in *mt*CBS-PPase are marked with asterisks, AMP-binding residues (side-chain contacts only) in *cp*CBS-PPase [8] are shown in italics, and disease-associated residues are shaded in grey. The box indicates a conserved CBS domain motif. Non-PPase CBS protein designation: hCBS, human cystathionine  $\beta$ -synthase; hCIC-7, human chloride channel CIC-7; hIMPDPH1, human IMP dehydrogenase 1; hAMPK $\gamma$ 2, human AMP-activated protein kinase  $\gamma$ 2; hAMPK $\gamma$ 3, human AMP-activated protein kinase  $\gamma$ 3; pigAMPK $\gamma$ 3, pig AMP-activated protein kinase  $\gamma$ 3. For AMPKs containing two pairs of CBS domains per subunit, CBS domain numbers, beginning from the N-termini, are presented after protein names.

(B) Stereoview of the nucleotide-binding site of *cp*CBS-PPase with bound AMP [8] showing the residues that were selected for mutation in the present study. Residue numbering is for *cp*CBS-PPase/*mt*CBS-PPase. The primed K100 belongs to the other subunit and forms an ion pair with P $\alpha$  of AMP. S279/R170 is expected to interact with P $\beta$  of ADP, and Y278/Y169 forms a stacking interaction with the adenine moiety [8].

### Figure 2 Time-courses of PP<sub>i</sub> hydrolysis by wild-type *mt*CBS-PPase (WT) and

**selected variants**

The assay medium contained 160  $\mu\text{M}$   $\text{PP}_i$ . The reaction was initiated by adding an aliquot of enzyme solution pre-equilibrated at 25°C. Curve labels refer to variant identity and final enzyme concentration in nanograms per millilitre.

**Figure 3 Effects of AMP (A) and ADP (B) on the equilibrium activity of wild-type *mtCBS*-PPase (WT) and selected variants**

The assay medium contained 160  $\mu\text{M}$   $\text{PP}_i$ . The lines show the best fit to eqn 1. The dashed line, representing the effect of AMP on wild-type enzyme (WT), is obtained from Jämsen et al. [7].

**Figure 4 [ $^{14}\text{C}$ ]AMP binding to wild-type *mtCBS*-PPase (WT) and selected variants, determined with membrane filtration**

The abscissa depicts the concentration of free unbound AMP, calculated as the difference between total and bound AMP concentrations. The lines represent the best fit to eqn 1 with  $A_L$  set to zero.

**Figure 5 Effects of AMP (A) and ADP (B) on the rate constant for substrate-induced activity transition in wild-type *mtCBS*-PPase (WT) and selected variants**

The assay medium contained 160  $\mu\text{M}$   $\text{PP}_i$ . The lines represent the best fit to eqn 1.



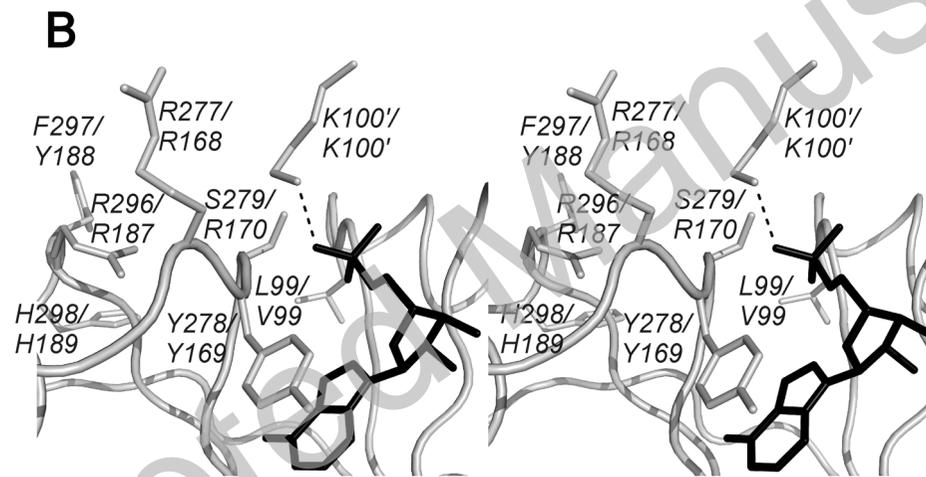


Figure 1B

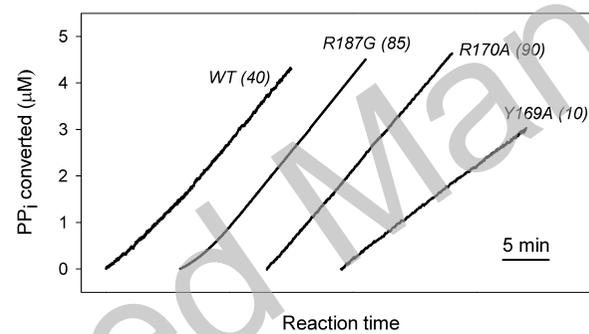


Figure 2

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101204

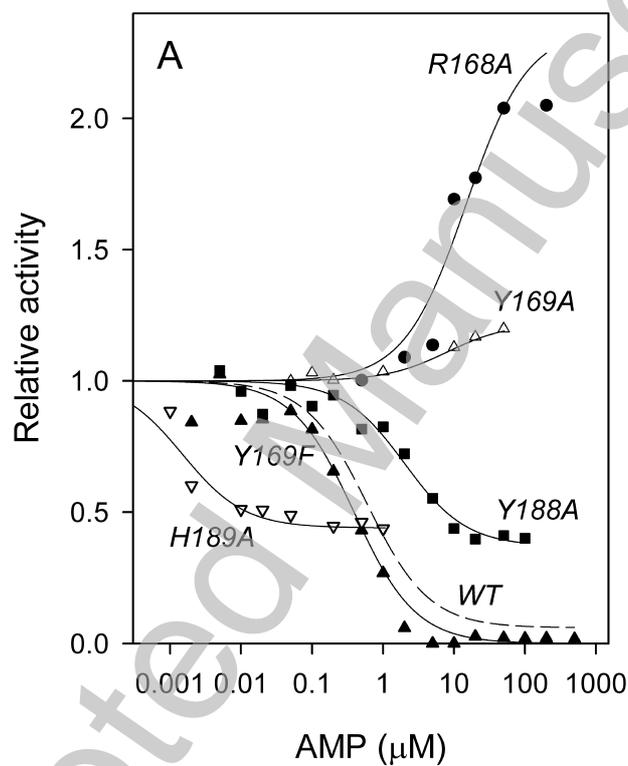


Figure 3A

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101204

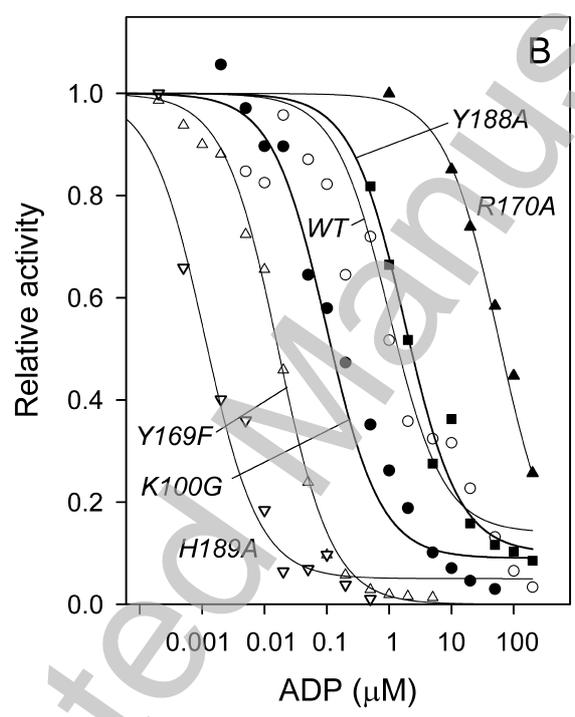


Figure 3B

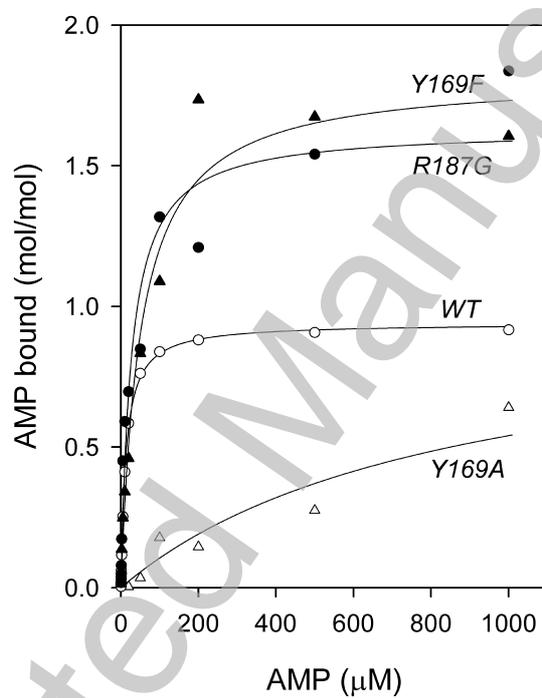


Figure 4

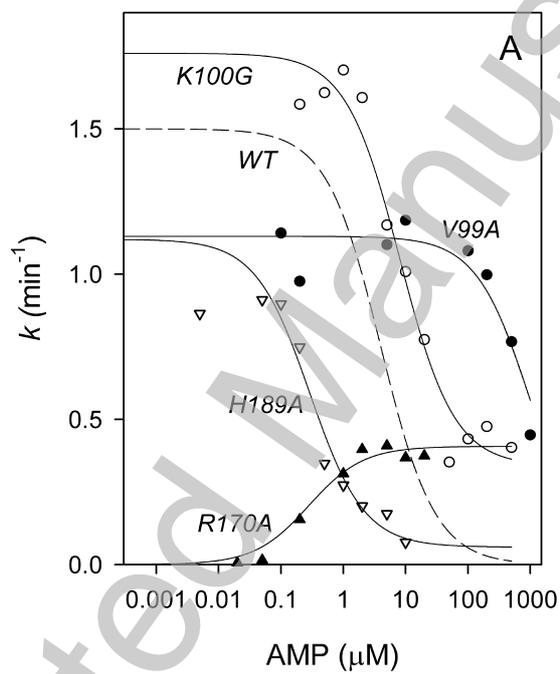


Figure 5A

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101204

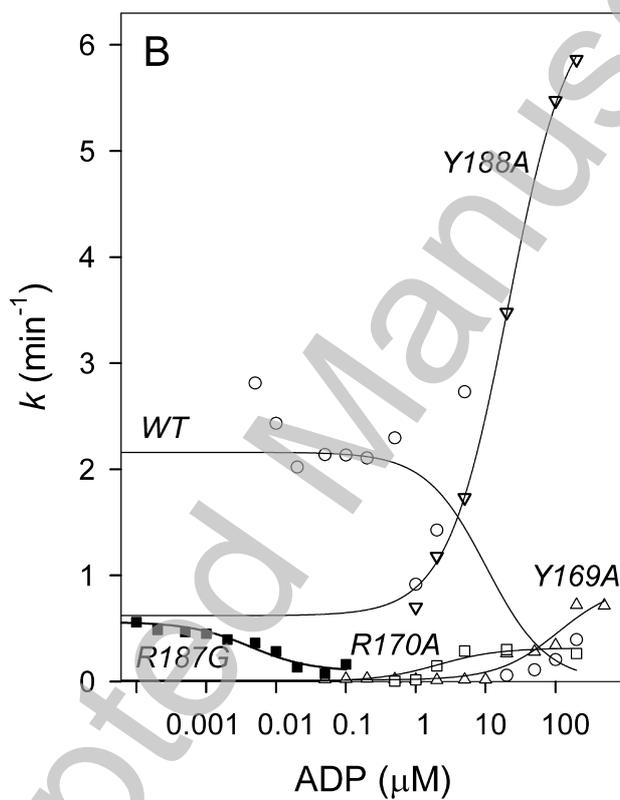


Figure 5B