Specific Mutations Reverse Regulatory Effects of Adenosine Phosphates and Increase Their Binding Stoichiometry in CBS Domain-Containing Pyrophosphatase

Viktor A Anashkin *, Elena A. Kirillova , Victor N Orlov , Alexander A. Baykov *

Posted Date: 28 April 2024

doi: 10.20944/preprints202404.1753.v1

Keywords: allosteric regulation; cooperativity; cystathionine β-synthase domain; diadenosine tetraphosphate; enzyme regulation; isothermal calorimetry

 doi:10.20944/preprints202404.1753.v1

Article

Specific Mutations Reverse Regulatory Effects of Adenosine Phosphates and Increase Their Binding Stoichiometry in CBS Domain-Containing Pyrophosphatase

Viktor A. Anashkin *, Elena A. Kirillova, Victor N. Orlov and Alexander A. Baykov *

Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow 119899, Russia

* Correspondence: victor_anashkin@belozersky.msu.ru (V.A.A.); baykov@belozersky.msu.ru (A.A.B.)

Abstract: Regulatory cystathionine β-synthase (CBS) domains are widespread in proteins; however, difficulty in structure determination prevents a comprehensive understanding of the underlying regulation mechanism. Tetrameric microbial inorganic pyrophosphatase containing such domains (CBS-PPase) is allosterically inhibited by AMP and ADP and activated by ATP and cell alarmones diadenosine polyphosphates. Each CBS-PPase subunit contains a pair of CBS domains but binds cooperatively only one molecule of the mono-adenosine derivatives. We used site-directed mutagenesis of *Desulfitobacterium hafniense* CBS-PPase to identify the key elements determining the direction of the effect (activation or inhibition) and the "half-of-the-sites" ligand binding stoichiometry. Seven amino acid residues were selected in CBS1 domain based on the available X-ray structure of the regulatory domains and substituted by alanine and other residues. The interaction of 11 CBS-PPase variants with the regulating ligands was characterized by activity measurements and isothermal titration calorimetry. Lys100 replacement reversed the effect of ADP from inhibition to activation, whereas Lys95 and Gly118 replacements made ADP activator at low concentrations but inhibitor at high concentrations. Replacements of these residues for alanine increased the stoichiometry of mono-adenosine phosphate binding twofold. These findings identify several key protein residues and suggest "two non-interacting pairs of interacting regulatory sites" concept in CBS-PPase regulation.

Keywords: allosteric regulation; cooperativity; cystathionine β-synthase domain; diadenosine tetraphosphate; enzyme regulation; isothermal calorimetry

1. Introduction

The activities of many important enzymes, membrane transporters, and other proteins in all kingdoms of life are allosterically controlled via regulatory ligand binding to cystathionine β-synthase (CBS) domains. For instance, the human genome encodes 75 such CBS-proteins, some of which are associated with hereditary diseases [1,2]. Most commonly, the regulatory ligands are various adenosine or guanosine phosphates [3–6]. The CBS domains are generally found in pairs in the protein sequence and form tightly associated "Bateman modules" in the ternary structure. The Bateman modules of two subunits further form four-domain disk-like "CBS modules" in homo-oligomeric proteins. Interestingly, the same enzyme may be regulated via CBS domains in some species but lack them in other species. Furthermore, CBS domains can be deleted without loss of activity [7]. These findings suggest the utility of CBS domains as movable regulatory blocks for constructing regulated proteins. Because of the difficulty in obtaining structural data for full-size CBS-proteins, presumably associated with the unusual flexibility of CBS modules, only limited information on the structural mechanism of CBS domain-mediated regulation is available.

 doi:10.20944/preprints202404.1753.v1

Prokaryotic CBS domain-containing pyrophosphatase (CBS-PPase; EC 3.6.1.1), a typical CBS-protein, belongs to Family II PPases, better known as a nonregulated, CBS domain-lacking form [8]. The latter is a dimer of identical subunits, each formed by DHH and DHHA2 domains, with an active site in between. CBS-PPase contains an ~250-residue regulatory insert of two CBS and one DRTGG domain in the catalytic DHH domain (Figure 1A). AMP and ADP inhibit CBS-PPase, whereas ATP, Ap_4A, and other linear diadenosine polyphosphates (Ap_nA, with n > 4) activate it by binding to CBS domains [6,9,10]. The amino acid sequences of CBS domains reveal significant similarities between CBS-PPases and other CBS domain-containing proteins (CBS-proteins), despite a considerable degree of variability. The DRTGG domain is absent in other CBS-proteins and some CBS-PPases and hardly plays a significant role in their regulation.

Unlike the canonical Family II PPases, CBS-PPases form tetramers [11], which are stabilized by separate crosswise interactions of the catalytic and regulatory parts [12] (Figure 1B). The structure of tetrameric CBS-PPase from *Desulfitobacterium hafniense* (*dh*PPase) without the DHHA2 domain was obtained at 16-Å resolution by single-particle electron microscopy [12], whereas the structures of the dimeric regulatory part of homologous *Clostridium perfringens* CBS-PPase (*cp*PPase) in complexes with AMP (inhibitor) and Ap_4A (activator) were determined at 2.3 Å by X-ray crystallography [13]. Although each CBS domain can bind a mono-adenosine phosphate and there are two CBS domains per subunit, the structures indicated the binding stoichiometry of one molecule of AMP per subunit or one molecule of Ap4A per dimer [10,13]. The latter ligand bridges two subunits by placing its two adenine moieties into the same sites occupied by AMP in each neighboring subunit.

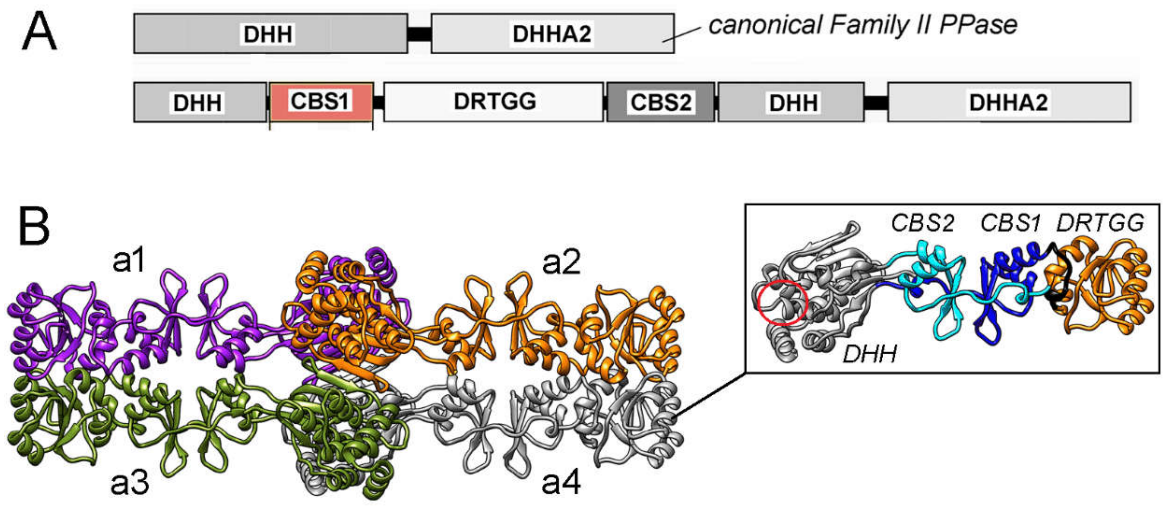


Figure 1. (A) Domain composition of canonical PPase and CBS-PPase of Family II. **(B)** 3D structure of tetrameric *dh*PPase without the DHHA2 domain suggested by low-resolution cryo-EM and molecular modeling [12]. The four subunits a1–a4 are shown in different colors. The side panel shows separate subunit a4 with differently colored and labeled domains. The region of the active site retained in the truncated subunit is indicated by a red circle and is the location where the missing DHHA2 domain is attached. Panel B was reproduced from Zamakhov et al. [12] under the Creative Commons CC BY license.

CBS-PPase provides a good model for studying CBS domain-mediated regulation because it is easily accessible and stable, is differentially regulated by adenosine phosphates, including cell alarmones diadenosine polyphosphates, and its activity can be conveniently and precisely measured. In recent site-directed mutagenesis studies of *dh*PPase, we identified Arg276 and Arg295 in the CBS2 and Asn312 and Arg334 in DHH domains as having crucial roles in kinetic cooperativity (active site interaction) [14,15]. Another important finding was that replacement of Arg295 or Asn312 by alanine reversed the effect of ATP from activation to inhibition [15].

In this study, we continued this line of research by seeking other residues that control the direction of the ligand effect (activation or inhibition). Our analysis of the available structural data

 doi:10.20944/preprints202404.1753.v1

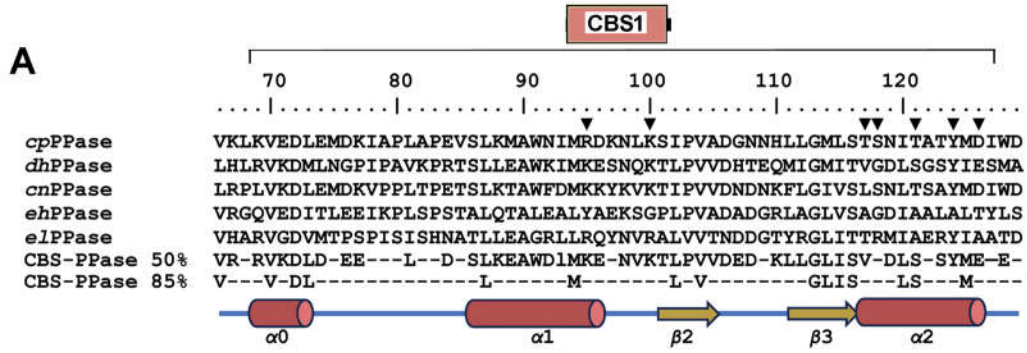
predicted that such residues may belong to the CBS1 domain. Another goal of our study was to identify the structural determinants of the "half-of-the-sites" ligand binding stoichiometry and increase it in CBS-PPase.

2. Results

2.1. Selection of Residues for Substitution

CBS1 domains participate in three types of interactions: (a) with the CBS2 domain of the same subunit to form Bateman module; (b) with the CBS1′ and CBS2′ domains of the neighboring subunit to form the four-domain CBS module; (c) with the adenine, ribose, and phosphate groups of regulatory adenosine phosphates (Figure 2). Seven residues potentially important for these interactions were selected in the *dh*PPase CBS1 domain (Figure 2B), based on the structure of the regulatory part of homologous *cp*PPase [13] and the in-depth structural analysis of CBS-proteins performed by Ereño-Orbea et al. [16]. Two of them (Lys100 and Tyr124) are identical in *dh*PPase and *cp*PPase, and the other are conservatively replaced in the pairs Lys/Arg95, Val/Thr117, Gly/Ser118, Ser/Thr121, and Glu/Asp126, where nonidentical residues found in the same position in *dh*PPase and *cp*PPase are separated by slashes. Six of these residues (except Gly/Ser118) form contacts with the neighboring subunit, and six residues (except Lys100) belong to helical regions. Lys/Arg95 and Glu/Asp126′, belonging to different subunits, form two symmetrical ionic pairs at the subunit interface in the AMP complex. Lys100 and Ser121 form similar symmetrical intersubunit H-bonds with CBS2 Arg276′ and CBS1 Ser121′residues, respectively. In addition, the Lys100 side chain neutralizes the charge of the α-phosphate group of the bound regulatory ligand. Val1/Thr117 and Tyr124 form a hydrophobic and stacking contact, respectively, with identical residues (Val/Thr117′ and Tyr124′) of the neighboring subunit. Residue 118 (Gly/Ser) is not involved in subunit contact, but, like Val/Thr117, belongs to the ribose-phosphate binding motif G-h-h-S/T-x-x-D/N (residues 113–119; h is a hydrophobic residue, and x is any residue) [17]. This motif determines the permissible chemical structure of the substituent at the ribose O5′ atom of the regulating ligand [16].

All seven selected residues were replaced with alanine. In addition, Val117 was replaced by threonine, which is found in this position in >50% CBS-PPases, including cpPPase (Figure 2A). Gly118 was additionally replaced by Ser, found in *cp*PPase (Figure 2A), and hydrophobic Ile and Met, found in S-adenosyl methionine-binding CBS domains [16]. All *dh*PPase variants were produced in a tag-less form in *Escherichia coli* and purified to apparent homogeneity using ion exchange and size-exclusion chromatography.



 doi:10.20944/preprints202404.1753.v1

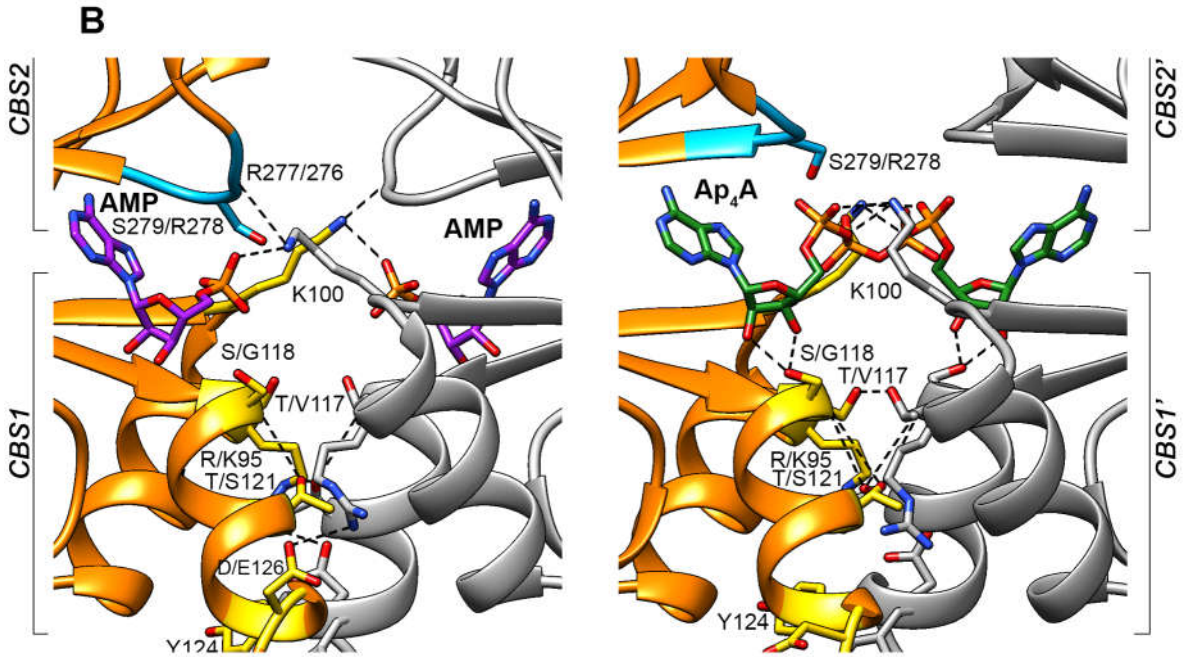


Figure 2. Selection of residues for substitution. (A) Partial sequence alignment of five CBS-PPases in the region of the CBS1 domain (residues 69–127). The residue numbers shown are identical for *cp*PPase and *dh*PPase. The residues substituted in this study are marked with inverted triangles. The consensus sequences based on 50 % and 85 % identities in 457 CBS-PPases (KEGG GENES database of March 28, 2024) are shown below. *cp*PPase, *C. perfringens* CBS-PPase (KEGG ID: CPF_2312); *dh*PPase, *D. hafniense* CBS-PPase (Dhaf_1515); *cn*PPase, *Clostridium novyi* CBS-PPase (NT01CX_1928); *eh*PPase, *Ethanoligenens harbinense* CBS-PPase (Ethha_0050); *el*PPase, *Eggerthella lenta* CBS-PPase (Elen_0953). (**B**) AMP and Ap_4A coordination in the regulatory S1 sites of *cp*PPase as determined by X-ray crystallography (PDB IDs 3L31 and 3L2B, respectively) [13]. The ribbon models of the two subunits are shown in different colors. The bound AMP and Ap_4A molecules and the *cp*PPase amino acid residues corresponding to those substituted in *dh*PPase are shown as stick models; the *dh*PPase residues are indicated after the slash in the residue labels. RYRN loop is blue. Atoms forming an H-bond or ionic pair are connected by dashed lines.

2.2. Effects of Substitutions on Tetramer Activity and Stability

The tendency to dissociate into inactive dimers in diluted solutions complicates activity assay of *dh*PPase. Similar to the wild-type enzyme [11], dissociation was a relatively slow process and did not occur significantly during the time of the activity assay (2–3 min), but was evident in preincubated diluted stock solutions. Therefore, the time-courses of enzyme activity upon dilution were measured to estimate tetramer stability and activity (Figure 3A). All variants were fairly active in PP_i hydrolysis and demonstrated a first-order transition to a new equilibrium with lower tetramer content and, consequently, lower activity (Figure 3B). The time courses were analyzed in terms of the equilibrium in Scheme 1, as described previously [11], to derive the values of k_a, k_d, and activities at zero and infinite time.

 doi:10.20944/preprints202404.1753.v1

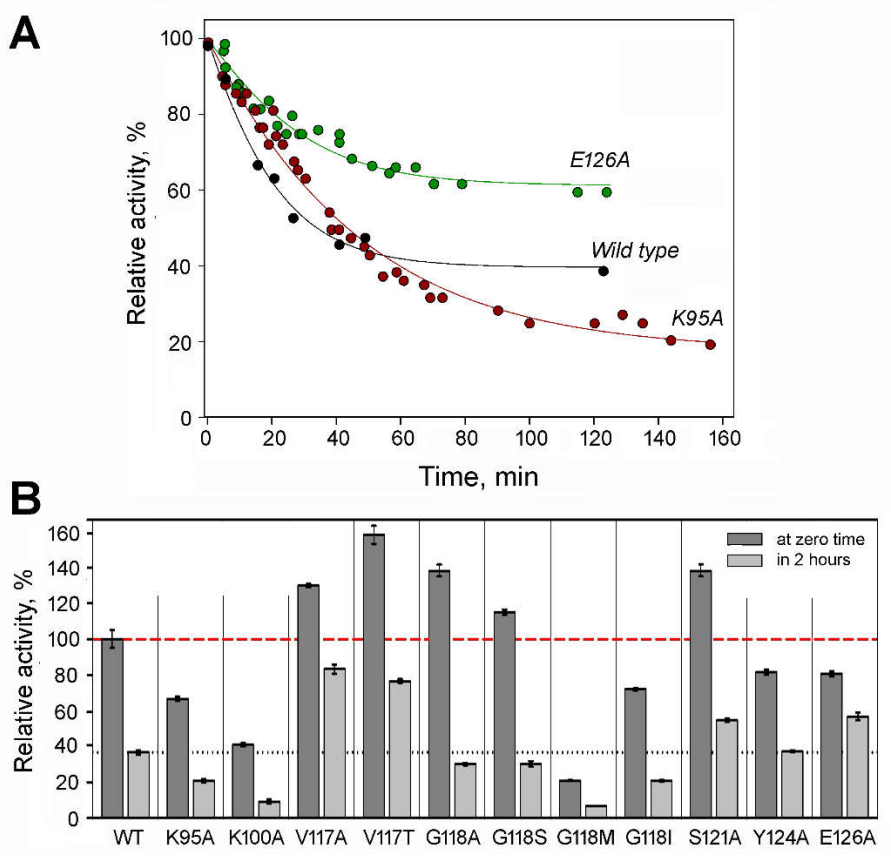


Figure 3. Specific activity and tetramer stability of *dh*PPase variants. (**A**) Activity time-courses for wild-type *dh*PPase and selected variants upon dilution at 25 °C. Stock enzyme solution containing 160-200 μM enzyme, 0.1 M Mops/KOH, pH 7.2, 2 mM $MgCl_2$, and 0.1 mM $CoCl_2$ was diluted into the same buffer to 0.2 μM enzyme. Aliquots were withdrawn in time, and activity was measured using the standard procedure. The lines were created using the best-fit parameter values found in Table 1. Activity extrapolated to zero time by the fitting procedure (A_T) was taken as 100 % for each curve (actual values are found in Table 1). (**B**) Specific activities before dilution and after equilibration in the diluted state. Wild-type *dh*PPase activity before dilution (590 ± 30 IU/mg) was considered 100 % for all variants.

$$T \underset{k_d}{\overset{k_a}{\leftrightarrows}} 2D$$

Scheme 1. Reversible dissociation of tetrameric enzyme into dimers.

As Table 1 highlights, the substitutions generally had a moderate effect on tetramer activity. The effect varied from a 4.8-fold decrease in the G118M variant to a 1.4-fold increase in the V117T variant. The effects of the substitutions on the rate constants for the reversible dissociation were more pronounced and decreased both k_d and k_a, except for the E126A variant, for which k_a slightly increased. The effects on k_a generally prevailed, resulting in tetramer destabilization in terms of the equilibrium dissociation constant $K_d = k_d/k_a$, except for the V117T and E126A variants, which were more stable than the wild-type enzyme. Based on the K_d values obtained, the percentage of dissociated tetramer in the stock enzyme solutions before dilution in the experiments for Figure 3A and similar did not exceed 10 % for the least stable variants (with Lys100 and Gly118 substituted).

 doi:10.20944/preprints202404.1753.v1

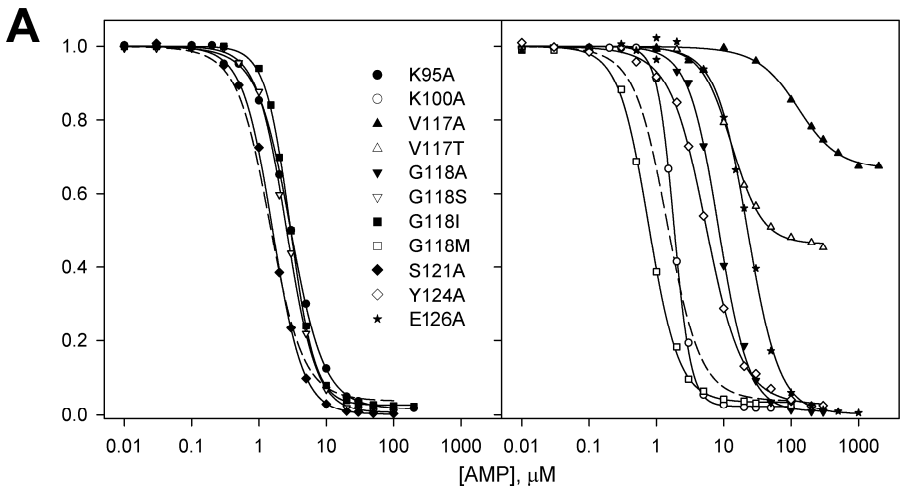
Table 1. Parameters for tetramer ⇆ dimer equilibrium in *dh*PPase variants. Values of k_a, k_d, and their ratio, $K_d = k_d/k_a$, were estimated from inactivation time-courses upon dilution (Figure 3) [11]. Zero-time wild-type *dh*PPase activity (A_T = 590 ± 30 U/mg) was taken as 100 %. Effects greater than 10-fold are indicated in boldface.

Enzyme variant	A_T, %	k_d, h^{-1}	k_a, $\mu M^{-1}\cdot h^{-1}$	K_d, μM
WT	100 ± 5	1.6 ± 0.1 (1.8 ± 0.1) [a]	2.1 ± 0.5 (2.4 ± 0.4) [a]	0.8 ± 0.2 (0.7 ± 0.2) [a]
K95A	66 ± 2	1.00 ± 0.04	0.3 ± 0.1	3.2 ± 0.8
K100A	42 ± 1	1.1 ± 0.1	**0.19 ± 0.04**	5.9 ± 1.6
V117A	113 ± 1	0.44 ± 0.02	1.5 ± 0.2	0.3 ± 0.1
V117T	141 ± 5	0.6 ± 0.1	0.6 ± 0.3	1.0 ± 0.6
G118A	120 ± 3	1.06 ± 0.04	**0.15 ± 0.04**	6.8 ± 0.6
G118S	117 ± 2	0.85 ± 0.04	**0.14 ± 0.06**	6.0 ± 2.7
G118M	21 ± 1	0.70 ± 0.04	**0.18 ± 0.10**	4 ± 2
G118I	73 ± 1	0.75 ± 0.03	**0.12 ± 0.05**	6.2 ± 2.9
S121A	120 ± 3	1.5 ± 0.1	0.5 ± 0.1	3.3 ± 0.9
Y124A	83 ± 2	1.0 ± 0.1	0.9 ± 0.1	1.1 ± 0.2
E126A	82 ± 2	0.7 ± 0.1	3.4 ± 0.5	0.21 ± 0.05

[a] Values in parentheses are from Anashkin et al. [11].

2.3. Regulation of dhPPase Variants by Adenosine Phosphates

The effects of four physiological modulators, including AMP, ADP, ATP, and Ap_4A, were tested for *dh*PPase variants (Figure 4A–D). The first two inhibit the wild-type enzyme, whereas the last two activate it [9,10]. Furthermore, other linear diadenosine polyphosphates, Ap_nA, with n = 3, 5, and 6 also activate *dh*PPase [10]. Thus, there is a correlation between the phosphate chain length in the CBS-PPase regulator and the direction of the effect—inhibition or activation. As the border lies between two and three phosphate groups, changes in the residues that sense polyphosphate length might shift this border and confer activation by ADP or inhibition by ATP to CBS-PPase.



 doi:10.20944/preprints202404.1753.v1

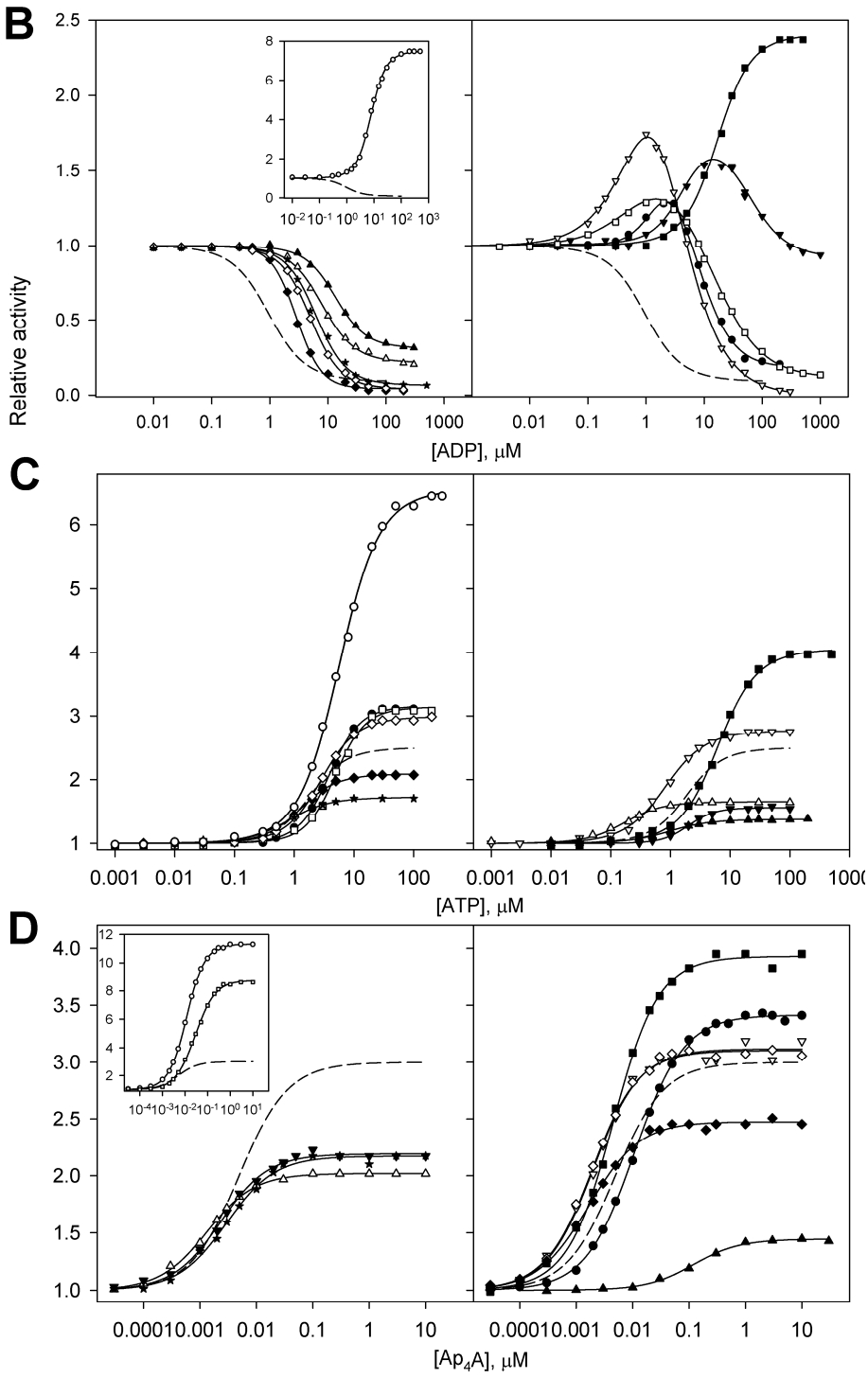


Figure 4. Effects of AMP **(A)**, ADP **(B)**, ATP **(C)**, and Ap_4A **(D)** on the activities of 11 *dh*PPase variants. Activity measured without adenosine phosphates (A_0 in Eqns 1 and 2) was taken as unity for each variant. The symbols used are detailed in the top panel. The insets show the data for variants with the largest effects. The reaction of PP_i hydrolysis was initiated by adding enzyme pre-incubated in stock 0.8–3.2 μM (0.05–0.2 mg/ml) solution to the assay mixture containing 0.1 M Mops/KOH buffer, pH 7.2, 5 mM $MgCl_2$, and 140 μM PP_i (yielding 50 μM $MgPP_i$ complex). The solid lines show the best fits of Eqn 1 or 2. The wild-type data are from previous publications [9,10].

In accordance with these expectations, only ADP inhibition was reversed in six variants obtained by substituting three residues: Lys95, Lys100, and Gly118 (Figure 4B). The effects were, however, not uniform. While the activities of the K100A and G118I variants monotonically increased to a constant

 doi:10.20944/preprints202404.1753.v1

level up to 500 μM ADP concentration, four other variants demonstrated bell-shaped dependence with activity finally dropping to ~1 % (G118S), ~10 % (G118M), 20 % (K95A), or 90 % (G118A) at the highest ADP concentrations.

In contrast, the effects of AMP, ATP, and Ap_4A on the *dh*PPase variants were qualitatively similar to their effects on the wild-type enzyme—AMP inhibited the enzymes, whereas ATP and Ap_4A activated them. However, there were important quantitative differences. Thus, the V117A and V117T variants demonstrated increased residual activity at saturating AMP levels (Figure 4A). Interestingly, the same effect was observed with these two variants in ADP inhibition (Figure 4B), and they additionally demonstrated decreased activation by ATP and Ap_4A compared with the wild-type enzyme (Figure 4C,D). For three variants (V117A, V117T, and E126A), AMP profiles were markedly shifted to higher AMP concentrations, indicating weaker ligand binding. For the K100A variant, the AMP profile was much steeper than those for all other enzyme forms, indicating increased binding cooperativity. The K100A variant surpassed all other enzyme forms in terms of the degree of activation by ATP, and the V117A variant demonstrated the highest sensitivity to low ATP concentrations. Two substitutions (K100A and G118M) increased activation by Ap_4A to 9–11-fold, and one (V117A) markedly suppressed the activation by this ligand in terms of both binding affinity and the degree of activation.

For quantitative comparison, dose dependences were analyzed using the Hill equation. For monotonous dependencies, this equation was used in its simple form:

$$v = \frac{A_1 + A_0(\frac{K}{[N]})^h}{1 + (\frac{K}{[N]})^h}, \quad (1)$$

where A_0 and A_1 are activities at zero and infinite ligand (N) concentrations, respectively, K is an apparent binding constant [$v = (A_0 + A_1)/2$ at $[N] = K$], and h is the Hill coefficient. The parameter values obtained by fitting Eqn 1 with the program Scientist (MicroMath) and listed in Table 2 support the conclusions made above based on the visual inspection of Figure 4. The K100A variant demonstrated an extraordinarily high Hill coefficient of 3.5 in AMP inhibition, indicating the presence of at least four strongly interacting binding sites for this ligand. In all other cases, h value did not significantly exceed 2.0. The V117A substitution had the greatest effect on the binding affinity (decreased it) for all adenosine phosphates, except ATP.

The bell-shaped dependence describing the ADP effects on the four variants were treated with a modified Hill equation:

$$v = \frac{A_1 + A_0(\frac{K_1}{[N]})^{h1} + A_2(\frac{[N]}{K_2})^{h2}}{1 + (\frac{K_1}{[N]})^{h1} + (\frac{[N]}{K_2})^{h2}}, \quad (2)$$

This equation contains two binding constants and two Hill coefficients, implying binding of ≥ h_1 activating and ≥ h_2 inhibiting ligand molecules; the activity A_2 refers to the saturating concentration of the ligand; and A_1 is the activity of the enzyme with only the activating ligand bound. The parameter values obtained using Eqn 2 for the four variants are listed in Table 3. Interestingly, of the four Gly118 replacements, Gly/Ala had the largest effect on the binding affinity (K_1 and K_2), clearly indicating that the side chain size is not the key parameter for this residue position. The h value of 0.9–1.1 for the ascending and descending parts of the ADP profiles in the G118S and G118M variants suggests that they result from independent binding of two ADP molecules. In other words, these substitutions cancel ADP-binding cooperativity. In contrast, the cooperativity was retained in two other variants with the bell-shaped profiles (Table 3), which is only possible if the number of interacting sites increases above two. Thus, either the network of interactions involves all four binding sites present in the wild-type enzyme, or new site(s) appear in the K95A and G118A variants.

 doi:10.20944/preprints202404.1753.v1

Table 2. Parameters of Eqn 1 derived from the effects of the adenosine phosphates on the activities of 11 *dh*PPase variants (Figure 4).

Enzyme variant	Adenosine phosphate and parameter value [a]											
	AMP			ADP			ATP			Ap_4A		
	A_1, %	K, μM	h	A_1 %	K, μM	h	A_1 %	K, μM	h	A_1 %	K, nM	h
WT [b]	3.7 ± 0.1	1.4 ± 0.1	1.7 ± 0.1	9 ± 1	0.95 ± 0.05	1.34 ± 0.03	250 ± 20	1.9 ± 0.2	1.45 ± 0.13	300 ± 10	4.9 ± 0.2	1.00 ± 0.02
K95A	1.4 ± 0.2	2.90 ± 0.04	1.7 ± 0.1	N.A. [c]	N.A.	N.A.	317 ± 5	3.9 ± 0.1	1.63 ± 0.06	341 ± 6	10.9 ± 0.3	1.00 ± 0.02
K100A	2.3 ± 0.4	1.82 ± 0.02	**3.5 ± 0.1**	**750 ± 30**	**7.8 ± 0.2**	1.55 ± 0.05	**650 ± 20**	5.6 ± 0.1	1.26 ± 0.04	**1130 ± 20**	11.5 ± 0.1	1.07 ± 0.01
V117A	**66 ± 1**	**124 ± 5**	1.4 ± 0.1	**31 ± 1**	**13.4 ± 0.5**	1.60 ± 0.08	138 ± 1	0.95 ± 0.05	1.22 ± 0.07	145 ± 2	**120 ± 10**	1.06 ± 0.09
V117T	**46 ± 1**	**13.2 ± 0.4**	1.9 ± 0.1	**21 ± 1**	**7.9 ± 0.3**	1.48 ± 0.07	165 ± 4	**0.17 ± 0.02**	1.34 ± 0.12	202 ± 2	1.30 ± 0.05	1.02 ± 0.04
G118A	0.8 ± 0.7	8.4 ± 0.2	1.9 ± 0.1	N.A.	N.A.	N.A.	156 ± 2	2.00 ± 0.07	1.92 ± 0.12	219 ± 4	2.3 ± 0.2	1.02 ± 0.06
G118S	0.5 ± 0.6	2.55 ± 0.04	2.0 ± 0.1	N.A.	N.A.	N.A.	277 ± 5	0.82 ± 0.03	1.25 ± 0.05	312 ± 11	1.9 ± 0.2	1.03 ± 0.08
G118M	3.4 ± 0.6	0.76 ± 0.02	1.9 ± 0.1	N.A.	N.A.	N.A.	314 ± 8	4.9 ± 0.2	1.70 ± 0.12	**870 ± 30**	28 ± 1	0.99 ± 0.02
G118I	1.6 ± 1.0	2.90 ± 0.03	2.2 ± 0.1	240 ± 3	**17 ± 1**	1.5 ± 0.1	410 ± 12	6.2 ± 0.2	1.30 ± 0.06	393 ± 9	4.3 ± 0.2	1.05 ± 0.04
S121A	1 ± 4	1.60 ± 0.02	1.9 ± 0.1	4.2 ± 0.9	2.8 ± 0.1	2.0 ± 0.10	209 ± 2	1.27 ± 0.03	1.41 ± 0.05	247 ± 4	1.6 ± 0.1	0.97 ± 0.05
Y124A	2.6 ± 0.6	5.4 ± 0.1	1.6 ± 0.1	3.8 ± 0.6	5.0 ± 0.1	1.63 ± 0.04	299 ± 4	2.81 ± 0.06	1.40 ± 0.04	311 ± 7	1.8 ± 0.1	1.04 ± 0.05
E126A	± 0.9	**22.5 ± 0.5**	1.8 ± 0.1	6.5 ± 1.6	**6.2 ± 0.3**	1.6 ± 0.1	172 ± 2	0.80 ± 0.06	1.25 ± 0.10	217 ± 4	2.9 ± 0.2	1.02 ± 0.07

[a] Activity measured without adenosine phosphates (A_0 in Eqn 1) was taken as 100 % for each variant. Large effects are marked by boldface. [b] From Anashkin et al. [10] (Ap_4A data) or from Anashkin et al. [9] (other adenosine phosphates). [c] N.A., not attendant.

 doi:10.20944/preprints202404.1753.v1

Table 3. Parameters of Eqn 2 for the four variants with the bell-shaped activity versus [ADP] profiles in Figure 4 [a].

Enzyme variant	A_1, %	A_2, %	K_1, μM	K_2, μM	h_1	h_2
K95A	170 ± 30	18 ± 2	1.7 ± 0.6	7 ± 2	1.8 ± 0.3	1.4 ± 0.2
G118A	180 ± 20	91 ± 5	4.8 ± 1.5	50 ± 20	1.5 ± 0.2	1.3 ± 0.2
G118S	270 ± 30	0.8 ± 1.1	0.8 ± 0.2	3.0 ± 0.6	1.1 ± 0.1	1.1 ± 0.1
G118M	174 ± 19	10 ± 1	1.0 ± 0.5	10 ± 2	0.9 ± 0.1	0.9 ± 0.1

[a] The activity measured without ADP (A_0 in Eqn 2) was taken as 100 % for each variant.

S-Adenosyl methionine regulates cystathionine β-synthase by binding to its CBS domains [16], which contain Met in the position corresponding to *dh*PPase Gly118. However, this adenosine derivative had no effect on the activities of wild-type *dh*PPase and its G118M and G118I variants Thus, it appears that these substitutions alone were not enough to confer the ability to accommodate S-adenosyl methionine to *dh*PPase CBS domains.

2.4. AMP/ADP and ADP/Ap4A Competition in Variants with Unusual ADP Effects

The AMP and, presumably, ADP binding stoichiometry is one molecule per Bateman module [10,13] and, hence, four molecules per tetramer, all of which appear to be inhibitory in the wild-type enzyme. Two replacements (K100A and G118I) apparently converted all ADP-binding sites into activating. In the four variants demonstrating bell-shaped ADP profiles (K95A, G118A, G118S, and G118M), only part of the binding sites became activating ones. Alternatively, the latter substitutions might result in the appearance of additional ADP-binding sites with an activating effect on the enzyme.

Two types of experiments were conducted to select between these alternatives. First, the binding stoichiometry was indirectly estimated by measuring the competition between AMP, ADP, and Ap_4A in the activity assay (Figure 5). In these experiments, activity was measured at an increasing concentration of one ligand (AMP or ADP) and a fixed concentration of the other ligand (ADP or Ap_4A). If the effects of the two ligands were independent of each other, the activity profiles would shift upward due to activation by ADP and Ap_4A, without changing their shapes. This is clearly not the case. Moreover, when the dependences in Figure 4 for the "two-ligand" systems were analyzed with Eqn 1, the Hill coefficient and the binding constants differed significantly from the values found in Tables 2 and 3 for the corresponding "one-ligand" dependence. For the K100A variant, Ap_4A and ADP decreased the Hill coefficient for AMP inhibition from 3.5 to 1.8 and 2.8, respectively. This is an expected behavior because the competing ligands (Ap_4A and ADP) bind with no or lower cooperativity than AMP. This also explains the only moderate effects of ADP and Ap_4A on the value of K for AMP (Figure 5 and Table 2).

In the G118A and G118S variants, Ap_4A eliminated the activating effects of ADP and shifted the inhibition part of the bell-shaped profiles to larger ADP concentrations (Figure 5). These observations are consistent with Ap_4A competition for both ADP binding sites responsible for the bell-shaped appearance of the profiles in its absence.

 doi:10.20944/preprints202404.1753.v1

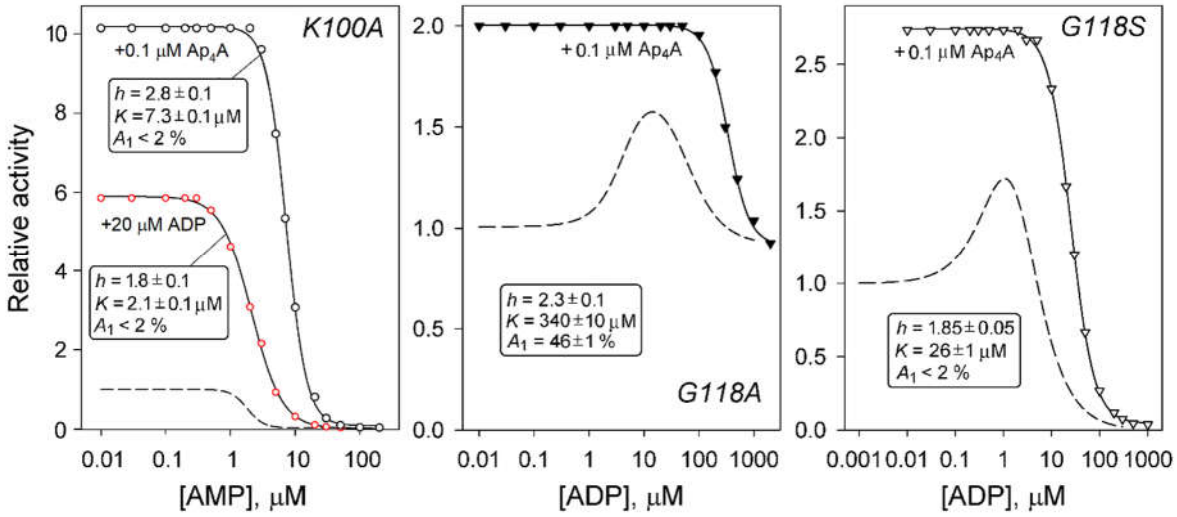


Figure 5. Competition between adenosine phosphates in selected *dh*PPase variants. The dashed lines show the activity dependence on the adenosine phosphate indicated on the abscissa and are taken from Figure 4. The solid lines with points show the same dependence measured in the presence of a fixed concentration of the adenosine phosphate indicated in the curve label. The parameter values estimated for the latter dependence using Eqn 1 are shown in a box on each panel. The activities measured in the absence of any adenosine phosphate were taken as unity for all curves in each panel.

2.5. Measurements of Adenosine Phosphate Binding by Isothermal Titration Calorimetry (ITC)

The binding stoichiometry of the adenosine phosphates was also estimated directly by isothermal titration calorimetry. In these experiments, enzyme solutions were titrated with increasing concentrations of the four ligands, and enthalpy changes were recorded after each ligand addition (Figure 6, *top*). Figure 6, *bottom* shows typical titration curves obtained for several *dh*PPase variants, and Table 4 summarizes all the information derived from the curves by fitting a simple binding equation with n binding sites per subunit. The binding constant K was another variable parameter in these fittings; however, its estimated value should be considered as only a rough average estimate because monoadenosine phosphate binding to CBS-PPase involves positive cooperativity [9]. This also refers to the $T\Delta S$ values calculated from the K values.

Most importantly, these data identified three variants with doubled binding stoichiometry n for all monoadenosine phosphates. Not surprisingly, these variants demonstrate unusual activity profiles in Figure 4. One of them (K100A) bound AMP with an unusually high positive cooperativity (Table 2), and two other (K95A and G118A) demonstrated bell-shaped profiles with ADP. Quite interestingly, two variants (G117S and G117M), which also showed bell-shaped profiles with ADP, nevertheless demonstrated "normal" binding stoichiometry for all monoadenosine phosphates. In contrast, no significant changes were detected in the value of n for Ap_4A binding to all variants (Table 4).

 doi:10.20944/preprints202404.1753.v1

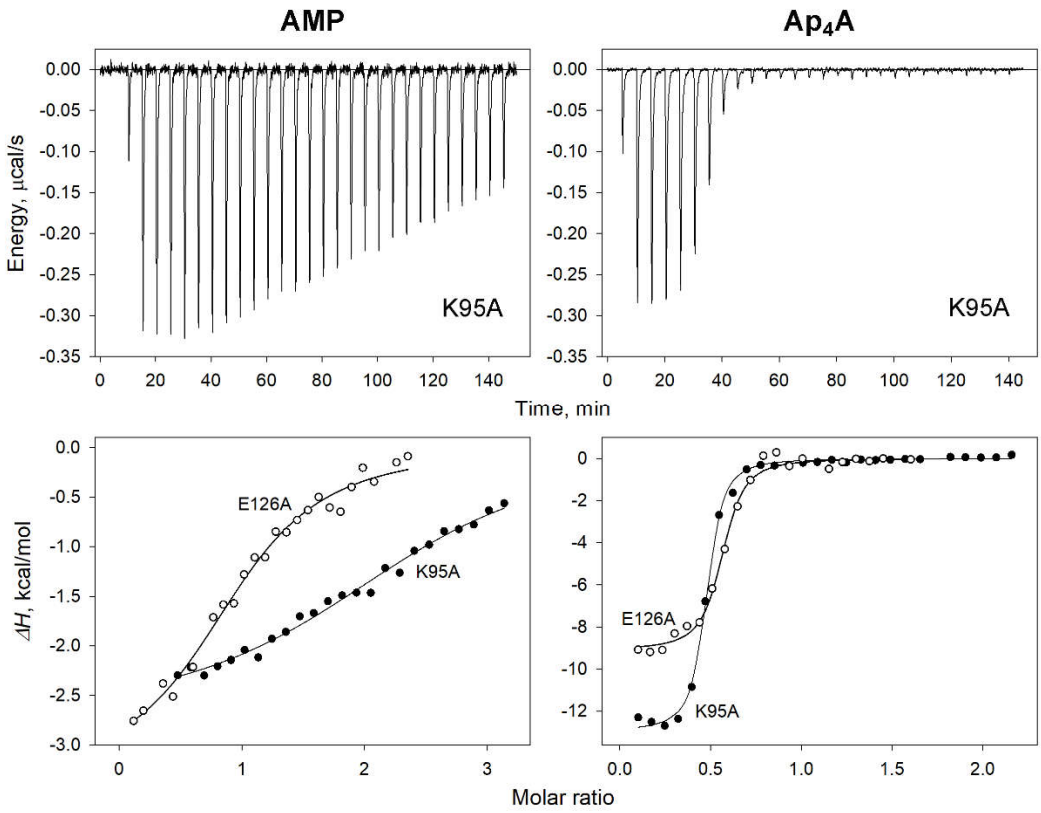


Figure 6. ITC measurements of adenosine phosphate binding to *dh*PPase variants. Top, Typical raw data for successive injections of AMP or Ap_4A into 8 µM solution of the K95A variant. Bottom, Integrated heats for selected variant titrations by AMP (left) and Ap_4A (right) after correction for dilution. The lines show the best fits of the single-binding-site model.

Table 4. The binding parameters derived from ITC titrations[a].

Enzyme variant	Adenosine phosphate	ΔH, kcal/mol	n	K_N, µM	$-T\Delta S$, kcal/mol
WT	AMP	-5.6 ± 0.5	0.79 ± 0.05	0.8 ± 0.3	-2.7 ± 0.9
K95A	AMP	-2.6 ± 0.2	**2.41 ± 0.04**	4.6 ± 0.7	-4.7 ± 0.4
K100A	AMP	-4.3 ± 0.1	**1.85 ± 0.04**	1.3 ± 0.2	-3.7 ± 0.3
V117A	AMP	-4.1 ± 0.6	0.94 ± 0.12	4.4 ± 0.8	-3.2 ± 0.8
V117T	AMP	-5.5 ± 0.2	0.94 ± 0.03	1.7 ± 0.2	-2.4 ± 0.3
G118A	AMP	-2.2 ± 0.1	**1.96 ± 0.09**	4.2 ± 0.6	-5.2 ± 0.2
G118S	AMP	-3.6 ± 0.2	1.16 ± 0.06	2.8 ± 0.4	-4.0 ± 0.3
G118M	AMP	-4.0 ± 0.1	0.98 ± 0.02	0.52 ± 0.07	-4.5 ± 0.2
G118I	AMP	4.6 ± 0.2	1.07 ± 0.03	1.6 ± 0.2	-3.3 ± 0.3
S121A	AMP	-5.1 ± 0.2	1.07 ± 0.02	1.1 ± 0.10	-3.1 ± 0.3
Y124A	AMP	-3.0 ± 0.2	1.10 ± 0.04	2.5 ± 0.2	-4.7 ± 0.2
E126A	AMP	-3.8 ± 0.2	1.05 ± 0.04	1.9 ± 0.4	-4.0 ± 0.4
WT	ADP	-5.9 ± 0.3	0.90 ± 0.04	1.0 ± 0.2	-2.3 ± 0.5
K95A	ADP	-3.6 ± 0.2	**2.1 ± 0.10**	9.3 ± 1.7	-3.2 ± 0.4
K100A	ADP	**-1.6 ± 0.2**	**2.1 ± 0.2**	6 ± 2	**-5.5 ± 0.5**
G118A	ADP	-3.3 ± 0.3	**2.1 ± 0.1**	5 ± 1	-4.0 ± 0.5
G118S	ADP	-5.8 ± 0.2	1.18 ± 0.03	1.1 ± 0.2	-2.3 ± 0.4
E126A	ADP	-5.9 ± 0.2	0.97 ± 0.02	0.6 ± 0.1	-2.6 ± 0.3
WT	ATP	-6.8 ± 0.2	0.88 ± 0.02	2.1 ± 0.3	-0.9 ± 0.3

 doi:10.20944/preprints202404.1753.v1

K95A	ATP	-4.3 ± 0.7	**2.4 ± 0.3**	6 ± 2	-2.8 ± 1.0
K100A	ATP	**-0.41 ± 0.03**	**1.9 ± 0.1**	0.4 ± 0.2	**-8.3 ± 0.5**
G118A	ATP	**-1.3 ± 0.4**	**2.0 ± 0.3**	3 ± 4	**-6.3 ± 1.7**
G118S	ATP	**-1.0 ± 0.2**	1.1 ± 0.1	0.5 ± 0.5	**-8 ± 1**
E126A	ATP	-4.1 ± 0.6	1.0 ± 0.1	4.3 ± 1.2	-3.2 ± 0.9
WT	Ap_4A	-10.3 ± 0.3	0.41 ± 0.01	N.D.[b]	N.A.
K95A	Ap_4A	-13.7 ± 0.6	0.45 ± 0.01	N.D.	N.A.
K100A	Ap_4A	-11.4 ± 0.4	0.52 ± 0.01	N.D.	N.A.
V117A	Ap_4A	-9.0 ± 0.3	0.41 ± 0.01	N.D.	N.A.
V117T	Ap_4A	-9.1 ± 0.4	0.41 ± 0.01	N.D.	N.A.
G118A	Ap_4A	-10.0 ± 0.2	0.47 ± 0.01	N.D.	N.A.
G118S	Ap_4A	-8.0 ± 0.2	0.40 ± 0.01	N.D.	N.A.
G118M	Ap_4A	-9.2 ± 0.4	0.45 ± 0.01	N.D.	N.A.
G118I	Ap_4A	-10.0 ± 0.4	0.44 ± 0.01	N.D.	N.A.
S121A	Ap_4A	-14.3 ± 0.3	0.42 ± 0.01	N.D.	N.A.
Y124A	Ap_4A	-10.0 ± 0.4	0.41 ± 0.01	N.D.	N.A.
E126A	Ap_4A	-10.7 ± 0.2	0.40 ± 0.01	N.D.	N.A.

[a] Large effects are marked by boldface. [b] N.D., not determined.

3. Discussion

CBS-PPase regulation by adenosine phosphates has six phenomenological characteristics: direction of the effect (activation or inhibition) and its size (regulation "intensity"), ligand binding stoichiometry, affinity, and cooperativity, and tetramer stability (its dissociation abolishes activity). In what follows, we consider the identities of the amino acid residues that control these characteristics (Table 5), speculate on the possible structural mechanisms of this control, and compare the obtained information with that reported for other enzymes regulated via CBS domains.

Table 5. Summary of the effects of the substitutions in the CBS1 domain on the regulatory characteristics of CBS-PPase [a].

Residue	Direction of ADP effect	Effect size	Binding stoichiometry	Binding affinity	Binding cooperativity	Tetramer stability
Lys95	++		+++			– –
Lys100	+++	+++	+++		+++	– –
Val117	+	– – –		++		+
Gly118	+++	++	+	+	+++	– –
Ser121						–
Tyr124				+		
Glu126				+		+

[a] The plus and minus signs indicate increase and decrease, respectively, in the numerical value of the characterizing parameter, where applicable, and their number indicates the relative size of the effect.

CBS-PPase is differentially regulated by adenosine phosphates—AMP and ADP inhibit it, whereas ATP, cAMP, and Ap_4A activate it [9,10,18]. Apart from the polyphosphate chain length, three residues of the CBS1 domain appear to control the direction of the effect: Lys95, Lys100, and Gly118. Of these, Lys100 appears to be the most important, as its substitution by alanine completely reversed the ADP effect from inhibition to activation, whereas in the K95A and G118A variants, ADP remained inhibitory at high concentrations (Figure 4). The remaining inhibition was, however, much less pronounced in the G118A variant than in the wild-type enzyme and simply counterbalanced the activation effect observed at low ADP concentrations (Tables 2 and 3).

Similar effects have been previously observed with this and other CBS-proteins. Thus, Arg/Ala and Asn/Ala replacements in the CBS2 domain of *dh*PPase converted ATP from an activator to an

 doi:10.20944/preprints202404.1753.v1

inhibitor [15]. Two substitutions in the CBS2 domain of *Moorella thermoacetica* CBS-PPase converted AMP from an inhibitor to an activator [19]. A similar modulation of the regulation of cystathionine β-synthase by S-adenosyl methionine [7] and AMP-dependent protein kinase by AMP and ATP has been reported [4,20]. Interestingly, the Arg residue of AMP-dependent muscle protein kinases, analogs of CBS-PPase Lys100, is the site of a natural pathogenic mutation [20–22]. In human cystathionine β-synthase, this position is also mutagenic but is occupied by Asp [23].

In the absence of the structure of CBS-PPase complex with ADP, the structure of the regulatory part with bound AMP (Figure 2B) provides some clue to the reversal of the ADP effect in the K100A variant. In the wild-type enzyme, the pyrophosphate moiety of ADP likely also interacts with Lys100 NH_3^+ group, and its removal should shift ADP upward, to a nearby Ser279 hydroxyl group (Arg278 guanidino group in *dh*PPase). This shift is expected to convert the CBS2 domain into "open" conformation found in the complex with the activator Ap_4A, wherein such transition is mediated by conservative RYSN /RYRN loop encompassing Ser279/Arg278 [13]. This transition and, hence, activation cannot occur with AMP because its short phosphate chain cannot reach Ser279/Arg278 but is permitted with ATP, which activates CBS-PPase.

The dual activation/inhibition effect of ADP on the G118A/S/M variants needs a different explanation. Notably, there is a correlation between the ADP effect on activity and residue 118 sidechain size: both canonical sites are inhibitory if this residue is Gly, one site becomes activating in the Gly/Ala, Gly/Ser, and Gly/Met variants, and both sites are activating in the Gly/Ile variant. The presence of a residue with a side chain in this position may dictate the adenosine moiety to bind in the conformation adopted in the activated Ap_4A complex in Figure 2B, allowing a similar partial "opening" of the CBS2 domain via the RYRN loop. However, because of electrostatic repulsion, the second ADP molecule cannot adopt the same conformation and triggers conversion of both ADP molecules into the "AMP" conformation with concomitant inhibition. This conversion is apparently prevented by the large side chain in the G118I variant, making ADP activator at both sites.

Interestingly, the two Ala variants that demonstrated bell-shaped activation/inhibition profiles with ADP exhibited a twofold greater ADP-binding stoichiometry (Table 4), indicating the appearance of an additional binding site. This site accommodated all mono-adenosine phosphates and was likely the pseudosymmetry-related S2 site, "silent" in the wild-type enzyme. Notably, a binding stoichiometry greater than one per Bateman module has also been reported for four CBS2 domain variants of *M. thermoacetica* CBS-PPase [19] and authentic CBS domain-containing proteins [24–27].

The localization of the additional binding site was confirmed by AMP docking experiments with an AlphFold2-generated model of the regulatory part of the K100A variant of *dh*PPase (Figure S1). AMP docked to both 'canonical" S1 site occupied in Figure 2B and "new" S2 site in this variant (Figure S1), with the values of the scoring function being 7.0 and 7.2 kcal/mol, respectively, i e., very similar. Notably, S2 site is functional in other CBS-proteins. Thus, both sites are occupied by adenosine derivative in the crystal structures of IMP dehydrogenase [3] and AMP-dependent protein kinase [27,28]. Moreover, the structures of cystathionine β-synthase [29] and chloride channel ClC-5 [30] contains the regulating adenosine derivative in only S2 site.

One could have speculated that ADP activates the variant *dh*PPase by binding to this "new" site and inhibits it by binding to the "canonical" site found in the wild-type enzyme (Figure 2B). This explanation was, however, ruled out by the bell-shaped dependence of activity on [ADP] for the G118S and G118M variants (Figure 4) that have no additional ADP-binding site (Table 4). In the latter variants, the observed activation and inhibition clearly result from ADP binding to the canonical sites—ADP activates the enzyme when bound to one site but inhibits when bound to both sites of the Bateman module, as discussed above. Consistent with this interpretation, the Hill coefficient is close to unity for both the activation and inhibition parts of the activity profiles for the G118S and G118M variants. In the presence of Ap_4A, ADP only inhibits these variants because ADP displaces Ap_4A molecule (a better activator) occupying both canonical sites.

Because of the inherent cooperativity and the variable value of the Hill coefficient, the *K* value should be considered as only a rough estimate of the binding affinity in the case of mono-adenosine

 doi:10.20944/preprints202404.1753.v1

phosphates. In contrast, Ap_4A binding is non-cooperative, and the determined *K* values do report the binding affinities of the *dh*PPase variants for this ligand. The largest effects on Ap_4A binding were observed upon Val117 substitution—increasing residue polarity (V to T substitution) stabilized the complex, whereas decreasing side chain volume (V to A substitution) destabilized the complex (Table 2). These findings emphasize the role of subunit interaction in CBS-PPase regulation. The fact that Val117 forms a hydrophobic interaction with the same residue of the partner subunit apparently doubled the effects of the Val117 substitutions. Val117 was also important for mono-adenosine phosphate binding, and the effects of its substitutions were opposite for complexes with AMP or ADP (destabilization) and ATP (stabilization).

The effects of E126A substitution were qualitatively similar. It should be noted that the *h* values were similar for the wild-type enzyme and the Val117 and E126A variants, justifying direct comparisons of the *K* values. Tyr124 also contacts its counterpart Tyr124′ of the neighboring subunit and apparently stabilized binding of only mono-adenosine phosphates in terms of *K*, the effect being larger for AMP and ADP (Table 2).

Wild-type *dh*PPase binds four molecules of the mono-adenosine phosphate ligands per tetramer in a positively cooperative manner, with a Hill coefficient between 1 and 2 (References [9,10] and Table 2). In the limiting case of positive binding cooperativity, the Hill coefficient could reach 4, which is at least twice as high. A likely explanation of this difference is that the binding sites interact within the four-CBS-domain structure shown in Figure 2B, but not with the distantly located sites of two other Bateman modules (Figure 1B). The lack of cooperativity in the Ap_4A binding by the wild-type and all *dh*PPase variants (Table 2) provides strong support for this idea. Each Ap_4A molecule binds to two adjacent Bateman modules (Figure 2B) in the subunit pairs a1-a3 and a2-a4. In other words, the binding sites appear to be structurally and functionally organized in pairs and interact with each other only within these otherwise independent pairs.

The K100A variant demonstrated remarkably high cooperativity in AMP binding (h = 3,5). Consistent with the concept of "two non-interacting pairs of interacting regulatory sites", this substitution unmasked an additional binding site in each subunit (Table 4), increasing to four the total number of sites in two interacting Bateman modules. AMP binding to any of these sites markedly increases the binding affinity of the other sites; thus, at equilibrium, all Bateman modules are predominantly in a ligand-free form or in a complex with four AMP molecules, with a low content of intermediate forms.

Why did cooperativity not increase with ADP and ATP in the K100A variant and with any adenosine phosphate in the K95A and G118A variants (Table 2) despite the increased binding stoichiometry (Table 4)? Additional phosphate groups apparently constitute a steric constraint for ligand binding to the additional site, making it less favorable in terms of the free energy of the system. As a result, sequential binding of ligand molecules becomes less favorable and results in the accumulation of appreciable amounts of intermediate complexes and, hence, decreased cooperativity. The lack of a significant increase in the cooperativity upon the K95A and G118A substitutions indicates the importance of the ionic pair formed by Lys95 and the conformational freedom around Gly118 for the interaction between four binding sites in four adjacent CBS domains. In the G118S and G118M variants, cooperativity is completely lost because of the increased steric constraints imposed by bulky side chains, further signifying the importance of Gly118 for site interaction in the Bateman module pairs.

As six of the seven substituted residues are involved in subunit contacts, most substitutions decreased tetramer stability to approximately the same extent (K_d increased from 0.7 to 4.0-6.8 uM). Because partial dissociation of the tetramer can occur under physiological conditions and the dissociated form (presumably dimer) is inactive [11], this phenomenon may contribute to CBS-PPase regulation. In two variants (V117T and Y124A), the tetramer retained its stability, whereas V117A and E126A substitutions stabilized it (Table 1). The smaller size of the side chain in the substituting Ala and, hence, a closer subunit contact in the V117A variant apparently explains its stronger subunit interaction. The effect of the E126A substitution was unexpected because Glu126 and Glu126′ form intersubunit contacts with Lys95′ and Lys95, respectively, whose substitution by Ala destabilized the

 doi:10.20944/preprints202404.1753.v1

tetramer (Table 1). A similar asymmetry was noted above in the effects of Lys95 and Glu126 substitutions on the activity modulation by ADP. This asymmetrical behavior of the ion pair constituents may be rationalized in terms of formation of a surrogate Asp130′—Lys95 ion pair in the E126A variant. This is, however, only a hypothesis and needs to be verified by structural studies.

The regulating adenosine phosphates can change the activity of the wild-type *dh*PPase from 3.7 % (AMP bound) to 300 % (Ap_4A bound) (Table 2), i.e., 80-fold. All substitutions of most residues increased this dynamic range to 250−600, whereas the Val117 substitutions decreased it to 2.2−4.4. The same applies to the AMP/ATP pair because the activating effects of ATP and Ap_4A are very similar and differ 2−3-fold with only K100A and G118M variants, decreasing the dynamic range by the same factor. Because PPase activity controls the concentration of pyrophosphate, the key regulator of biosynthesis [31], it would be interesting to determine in future studies the effects of the CBS-PPase mutation on the growth characteristics of their host bacterial species.

In summary, these findings reveal the important role of the CBS1 domain in CBS-PPase regulation by adenosine derivatives. By modifying specific amino acid residues in this domain, one can control the functionality of two binding sites for adenosine derivatives in Bateman module, increase the size of their regulating effect, reverse it, and make more sensitive to changes in regulating ligand concentration. The already available information makes CBS domains promising transmissible blocks for engineering proteins sensitive to the distribution of various adenosine phosphates, i.e. cell energy status. However, the implementation of this idea will require a broader understanding of the regulation mechanism associated with CBS domains. Most importantly, further exploration should be undertaken to investigate the way by which the regulating signal reaches the distantly located active site.

4. Materials and Methods

4.1. Materials

Wild-type *dh*PPase (UniProtKB: B8FP42) and its variants were produced in *E. coli* BL21 cells transformed with the pET-42b vector (Novagen) carrying the corresponding genes. All mutations in the CBS1 domain part of the *dh*PPase gene were performed using overlap extension PCR with Phusion DNA polymerase. The forward and reverse primers are listed in Table S1. The protein isolation procedure [11] included cell disruption by freezing/thawing, ion exchange chromatography on DEAE Toyopearl 650M, and size exclusion chromatography on Superdex 200, with absorbance monitoring at 280 nm. The final protein preparations were stored frozen in the elution buffer used at the gel filtration step (0.1 M Mops/KOH, pH 7.2, 2 mM $MgCl_2$, 0.1 mM $CoCl_2$, and 150 mM KCl). The purity of the isolated proteins, as estimated by SDS–PAGE [32] with Coomassie staining, was > 90%. Protein concentrations in milligrams per milliliter were determined spectrophotometrically using the extinction coefficient $A_{280}^{0.1\%}$ calculated from the amino acid composition with ProtParam of 0.477 for wild-type *dh*PPase and most of its variants and 0.455 for the Y124A variant. Molar concentrations were calculated in terms of the subunit using a subunit molecular mass of 60.5 kDa.

P1,P4-Diadenosine 5′-polyphosphate (Ap_4A, ammonium salt), AMP (free acid), ADP (di-monocyclohexylammonium salt), and ATP (di-sodium salt) were obtained from Sigma-Aldrich. The concentrations of nucleotide stock solutions were estimated by measuring absorbance at 259 nm (ε = 15,400 $M^{-1}cm^{-1}$ for the mono-adenosine phosphates and 30,800 $M^{-1}cm^{-1}$ for Ap_4A).

4.2. Enzyme Activity Assay

The initial rates of PPi hydrolysis were measured using a continuous P_i assay [33]. The assay medium contained 0.1 M Mops-KOH, pH 7.2, 5.23 mM $MgCl_2$, 140 μM PP_i, (corresponding to 50 μM $MgPP_i$ complex). Mg^{2+} complexation with AMP, ADP, and ATP was considered. The reaction was initiated by adding 0.1–10 nM enzyme and continued for 2–3 min at 25 °C. Rate values were obtained from the initial slopes of the P_i accumulation curves. Enzyme concentration in the assay was varied to obtain similar P_i accumulation rates in all cases, especially at low enzyme activities. Rate values (v)

 doi:10.20944/preprints202404.1753.v1

were subsequently normalized to the same enzyme concentration and weighed according to $1/v^2$ in the non-linear regression analysis.

4.3. Isothermal Titration Calorimetry (ITC)

Heat production upon nucleotide binding to dhPPase and its variants was measured at 25 °C using a VP-iTC calorimeter (MicroCal Ltd.). Enzyme and adenine nucleotide solutions were prepared on 0.1 M Mops/KOH buffer (pH 7.2) containing 2 mM $MgCl_2$, 0.1 mM $CoCl_2$, and 150 mM KCl. Titrations were performed by successive 10-µl injections of 100–300 µM AMP, ADP, ATP, or 33–60 µM Ap_4A solution into 1.4 ml of 8–10 µM protein solutions. The interval between injections was 5 min. The measured heat values were corrected for ligand dilution effects. The ITC data were analyzed using a MicroCal ITC subroutine in Origin 7.0 using a single-binding-site model.

4.4. Structure Modeling and Docking

The three-dimensional structure of the regulatory part of dhPPase (residues 68–303) was predicted from its amino acid sequence using AlphaFold2 (version 2.3.0) [34] with default parameter settings. All calculations were performed using an Nvidia RTX A5000 graphical card. Twenty-five models generated (five models per prediction mod) were ranked according to their iptm+ptm score, and the best model (score = 0.899) was selected.

AMP molecules were docked into the modeled structure using AutoDock Vina program [35] (ver. 1.2.5.) with a grid size of 70x70x70 $Å^3$. Twenty best models were selected for scoring in each calculation run. Final AMP positions agreed within 0.7 Å rmsd in three independent docking experiments. All structural visualizations were produced using UCSF Chimera [36].

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Table S1: The primers used for the site-directed mutagenesis of the CBS1 domain in *dh*PPase; Figure S1: Docking of AMP into the modeled structure of the dimeric regulatory part of *dh*PPase K100A variant.

Author Contributions: conceptualization, V.A.A.; methodology, all authors; validation, all authors; formal analysis, V.A.A.; investigation, V.A.A. and E.A.K.; writing—original draft preparation, V.A.A.; writing—review and editing, A.A.B.; visualization, V.A.A. and A.A.B.; supervision, V.A.A.; funding acquisition, V.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation (research project 22-74-00031).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data are available on request from the corresponding author.

Acknowledgments: We thank Anu Salminen for the *dh*PPase gene and Vera A. Aksenova for help in engineering the genetic constructs. The ÄKTA Purifier chromatographic system used for protein isolation became available in the framework of the Moscow State University Development Program PNR 5.13.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bateman, A. The structure of a domain common to Archaebacteria and the homocystinuria disease protein. Trends Biochem. Sci. **1997**, *22*, 12–13.
2. Scott, J.W.; Hawley, S.A.; Green, K.A.; Anis, M.; Stewart, G.; Scullion, G.A.; Norman, D.G.; Hardie, D.G. CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. J. Clin. Invest. **2004**, *113*, 274–284.
3. Labesse, G.; Alexandre, T.; Vaupre, L.; Salard-Arnaud, I.; Him, J.L.; Raynal, B.; Bron, P.; Munier-Lehmann, H. MgATP regulates allostery and fiber formation in IMPDHs. Structure **2013**, *21*, 975–985.
4. Adams, J.; Chen, Z.P.; Van Denderen, B.J.; Morton, C.J.; Parker, M.W.; Witters, L.A.; Stapleton, D.; Kemp, B.E. Intrasteric control of AMPK via the c1 subunit AMP allosteric regulatory site. Protein Sci. **2004**, *13*, 155–165
5. Fernandez-Justel, D.; Marcos-Alcalde, I.; Abascal, F.; Vidana, N.; Gomez-Puertas, P.; Jimenez, A.; Revuelta, J.L.; Buey, R.M. Diversity of mechanisms to control bacterial GTP homeostasis by the mutually exclusive binding of adenine and guanine nucleotides to IMP dehydrogenase. Protein Sci. **2022**, *31*, e4314-e4314.

6. Jämsen, J.; Tuominen, H.; Salminen, A.; Belogurov, G.A.; Magretova, N.N.; Baykov, A.A.; Lahti, R. A CBS domain-containing pyrophosphatase of Moorella thermoacetica is regulated by adenine nucleotides. Biochem. J. **2007**, *408*, 327–333.
7. Kery, V.; Poneleit, L.; Kraus, J. Trypsin cleavage of human cystathionine β-synthase into an evolutionarily conserved active core: structural and functional consequences. Arch. Biochem. Biophys. **1998**, *355*, 222–232.
8. Baykov, A.A.; Anashkin, V.A.; Salminen, A.; Lahti, R. Inorganic pyrophosphatases of Family II—two decades after their discovery, FEBS Lett. **2017**, *591*, 3225−3234.
9. Salminen, A.; Anashkin, V.A.; Lahti, M.; Tuominen, H.K.; Lahti, R.; Baykov, A.A. Cystathionine β-synthase (CBS) domains confer multiple forms of Mg2+-dependent cooperativity to family II pyrophosphatases. J. Biol. Chem. **2014**, *289*, 22865–22876.
10. Anashkin, V.A.; Salminen, A.; Tuominen, H.K.; Orlov, V.N.; Lahti, R.; Baykov, A.A. Cystathionine β-synthase (CBS) domain-containing pyrophosphatase as a target for diadenosine polyphosphates in bacteria. J. Biol. Chem. **2015**, *290*, 27594-27603.
11. Anashkin, V.A.; Salminen, A.; Orlov, V.N.; Lahti, R.; Baykov, A.A. The tetrameric structure of nucleotide-regulated pyrophosphatase and its modulation by deletion mutagenesis and ligand binding. Arch. Biochem. Biophys. **2022**, *692*, 108537.
12. Zamakhov, I.M.; Anashkin,V.A.; Moiseenko, A.V.; Orlov, V.N.; Sokolova, O.S.; Baykov, A.A. The structure and nucleotide-binding characteristics of regulated CBS domain-containing pyrophosphatase without one catalytic domain. Int. J. Mol. Sci. **2023**, *24*, 17160.
13. Tuominen, H.; Salminen, A.; Oksanen, E.; Jämsen, J.; Heikkilä, O.; Lehtio, L.; Magretova, N.N.; Goldman, A.; Baykov, A.A.; Lahti, R. 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. **2010**, *398*, 400–413.
14. Anashkin, V.A.; Orlov, V.N.; Lahti, R.; Baykov, A.A. An arginine residue involved in allosteric regulation of cystathionine β-synthase (CBS) domain-containing pyrophosphatase. Arch. Biochem. Biophys. **2019**, *662*, 40–48.
15. Anashkin, V.A.; Anu, S.; Osipova, E.; Kurilova, S.A.; Deltsov, I.D.; Reijo, L.; Baykov, A.A. Residue network involved in the allosteric regulation of cystathionine β-synthase domain-containing pyrophosphatase by adenine nucleotides. ACS Omega **2019**, *4*, 15549–15559.
16. Ereno-Orbea, J.; Oyenarte, I.; Martìnez-Cruz, L.A. CBS domains: ligand binding sites and conformational variability. Arch. Biochem. Biophys. **2013**, *540*, 70–81.
17. Day, P.; Sharff, A.; Parra, L.; Cleasby, A.; Williams, M.; Hörer, S.; Nar, H.; Redemann, N.; Tickle, I.; and Yon, J. Structure of a CBS domain pair from the regulatory γ1 subunit of human AMPK in complex with AMP and ZMP. Acta Crystallogr., Sect. D: Biol. Crystallogr. **2007**, *63*, 587–596.
18. Anashkin, V.A.; Aksenova, V.A.; Vorobyeva, N.N.; Baykov, A.A. Roles of nucleotide substructures in the regulation of cystathionine β-synthase domain-containing pyrophosphatase. Biochim. Biophys. Acta - General Subjects, **2019**, *1863*, 1263-1269.
19. Jämsen, J.; Tuominen, H.; Baykov, A. A.; Lahti, R. Mutational analysis of residues in the regulatory CBS domains of Moorella thermoacetica pyrophosphatase corresponding to disease-related residues of human proteins. Biochem. J. **2011**, *435*, 497–504.
20. Barnes, B.R.; Marklund, S.; Steiler, T.L.; Walter, M.; Hjälm, G.; Amarger, V.; Mahlapuu, M.; Leng, Y.; Johansson, C.; Galuska, D.; Lindgren, K.; Åbrink, M.; Stapleton, D.; Zierath, J.R.; Andersson, L. The 5′-AMP-activated protein kinase γ3 isoform has a key role in carbohydrate and lipid metabolism in glycolytic skeletal muscle. J. Bio.l Chem. **2004**, *279*, 38441−38447.
21. Gollob, M.H.; Green, M.S.; Tang, A.S.; Gollob, T.; Karibe, A.; Ali Hassan, A.S.; Ahmad, F.; Lozado, R.; Shah, G.; Fananapazir, L. et al. Identification of a gene responsible for familial Wolff–Parkinson–White syndrome. N. Engl. J. Med. **2001**, *344*, 1823–1831.
22. Costford, S.; Kavaslar, N.; Ahituv, N.; Chaudhry, S.; Schackwitz, W.; Dent, R.; Pennacchio, L.; McPherson, R.; and Harper, M. Gain-of-function R225W mutation in human AMPKγ3 causing increased glycogen and decreased triglyceride in skeletal muscle. PLoS ONE, **2007**, *2*, e903.
23. Kluijtmans, L.A.; Boers, G.H.; Stevens, E.M.; Renier, W.O.; Kraus, J.P.; Trijbels, F.J.; Van Den Heuvel, L.P.; and Blom, H.J. Defective cystathionineβ-synthase regulation by S-adenosylmethionine in a partially pyridoxine responsive homocystinuria patient. J. Clin. Inves.t **1996**, *98*, 285–289.
24. 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.
25. Jin, X., Townley, R., and Shapiro, L. (2007) Structural insight into AMPK regulation: ADP comes into play. Structure 15, 1285–1295.

 doi:10.20944/preprints202404.1753.v1

26. Xiao, B., Sanders, M. J., Underwood, E., Heath, R., Mayer, F., Carmena, D., Jing, C., Walker, P. A., Eccleston, J. F., Haire, et al. (2011) Structure of mammalian AMPK and its regulation by ADP. Nature 472, 230–233.
27. Gomez-García, I., Oyenarte, I., and Martínez-Cruz, L. (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.
28. Chen, L.; Wang, J.; Zhang, Y.Y.; Neumann, D.; Schlattner, U.; Wang, Z.-X.; and Wu, J.-W. AMP-activated protein kinase undergoes nucleotide-dependent conformational changes. Nat. Struct. Mo.l Biol. **2012**, *19*, 716–718.
29. Ereño-Orbea, J.; Majtan, T.; Oyenarte, I.; Kraus, J.P.; Martínez-Cruz, L.A. Structural insight into the molecular mechanism of allosteric activation of human cystathionine β-synthase by S-adenosylmethionine. Proc. Natl Acad. Sci. USA, **2014**, *111*, E3845-E3852.
30. Meyer, S.; Savaresi, S.; Forster, I.; and Dutzler, R. Nucleotide recognition by the cytoplasmic domain of the human chloride transporter ClC-5. Na.t Struct. Mol. Biol. **2007**, *14*, 60–67.
31. Heinonen, J.K. Biological Role of Inorganic Pyrophosphate; Kluwer Academic Publishers: London, UK, 2001.
32. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **1970**, *227*, 680–685.
33. Baykov, A.A.; Avaeva, S.M. A simple and sensitive apparatus for continuous monitoring of orthophosphate in the presence of acid-labile compounds. Anal. Biochem. **1981**, *116*, 1–4.
34. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; ... Hassabis, D. Highly accurate protein structure prediction with AlphaFold. Nature, **2021**, *596*, 583−589.
35. Eberhardt, J.; Santos-Martins, D.; Tillack, A.F.; and Forli, S. AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings. J. Chem. Inf. Model. **2021**, *61*, 3891-3898.
36. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera: A visualization system for exploratory research and analysis. J. Comput. Chem. **2004**, *25*, 1605–1612.).