Computational and Structural Biotechnology Journal 19 (2021) 3542–3555

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage:

Structural insight into the unique conformation of cystathionine β-synthase from *Toxoplasma gondii*

Carmen Fernández-Rodríguez [a,1], Iker Oyenarte [a,1], Carolina Conter [b,1], Irene González-Recio [a], Reyes Núñez-Franco [a], Claudia Gil-Pitarch [a], Iban Quintana [c], Gonzalo Jiménez-Osés [a], Paola Dominici [b], Maria Luz Martinez-Chantar [a], Alessandra Astegno [b,*], Luis Alfonso Martínez-Cruz [a,*]

[a] *Center for Cooperative Research in Biosciences (CIC bioGUNE), Basque Research and Technology Alliance (BRTA), Bizkaia Technology Park, Building 801A, 48160 Derio, Spain*
[b] *Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy*
[c] *Physics of Surfaces and Materials Unit. TEKNIKER, Basque Research and Technology Alliance (BRTA), C/Iñaki Goenaga, 5, 20600 Eibar Gipuzkoa, Spain*

ARTICLE INFO

Article history:
Received 13 April 2021
Received in revised form 31 May 2021
Accepted 31 May 2021
Available online 6 June 2021

Keywords:
Toxoplasmosis
Toxoplasma gondii
Reverse transsulfuration
Cystathionine β-synthase
Hydrogen sulfide
Homocysteine
Pyridoxal-5′-phosphate
Crystallography
Bateman module

ABSTRACT

Cysteine plays a major role in the redox homeostasis and antioxidative defense mechanisms of many parasites of the phylum Apicomplexa. Of relevance to human health is *Toxoplasma gondii*, the causative agent of toxoplasmosis. A major route of cysteine biosynthesis in this parasite is the reverse transsulfuration pathway involving two key enzymes cystathionine β-synthase (CBS) and cystathionine γ-lyase (CGL). CBS from *T. gondii* (*Tg*CBS) catalyzes the pyridoxal-5́-phosphate-dependent condensation of homocysteine with either serine or *O*-acetylserine to produce cystathionine. The enzyme can perform alternative reactions that use homocysteine and cysteine as substrates leading to the endogenous biosynthesis of hydrogen sulfide, another key element in maintaining the intracellular redox equilibrium. In contrast with human CBS, *Tg*CBS lacks the N-terminal heme binding domain and is not responsive to S-adenosylmethionine. Herein, we describe the structure of a *Tg*CBS construct that lacks amino acid residues 466-491 and shows the same activity of the native protein. *Tg*CBS Δ466-491 was determined alone and in complex with reaction intermediates. A complementary molecular dynamics analysis revealed a unique domain organization, similar to the pathogenic mutant D444N of human CBS. Our data provides one missing piece in the structural diversity of CBSs by revealing the so far unknown three-dimensional arrangement of the CBS-type of Apicomplexa. This domain distribution is also detected in yeast and bacteria like *Pseudomonas aeruginosa*. These results pave the way for understanding the mechanisms by which *Tg*CBS regulates the intracellular redox of the parasite, and have far-reaching consequences for the functional understanding of CBSs with similar domain distribution.

1. Introduction

Toxoplasmosis is a widespread zoonotic infection caused by *Toxoplasma gondii*, an obligate intracellular protozoon belonging to the phylum Apicomplexa. It is considered one of the five neglected parasitic infections recently targeted by the Centers for Disease Control and Prevention (CDC) as priorities for public health action in the United States, where an estimated 8–22% of people are infected. A similar prevalence exists in the United Kingdom, while in Central America, South America, and continental Europe estimated infection rates are even higher, ranging from 30% to 90% [1]. While toxoplasmosis is usually asymptomatic or causes only mild symptoms in immunocompetent individuals, it may cause severe diseases in immunocompromised patients. Infection in pregnancy may be even fatal to the fetus [2]. Moreover, multiple lines of evidence indicate that toxoplasmosis is related to a higher risk of many mental health disorders, including schizophrenia [3]. Despite intense efforts, effective vaccines are practically unavailable for humans and, consequently, chemotherapy represents the mainstay of disease control and treatment. However, recommended drugs require prolonged treatments and have high rates of toxic side effects. Moreover, the treatment has become complicated by the emergence of drug resistant parasites. Thus, alternative therapeutic options and new drugs are urgently needed.

* Corresponding authors.
E-mail addresses: alessandra.astegno@univr.it (A. Astegno), amartinez@cicbiogune.es (L.A. Martínez-Cruz).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.csbj.2021.05.052

Apicomplexa parasites are highly sensitive to changes in redox environment and compounds that produce oxidative stress were found to limit the pathogenesis of *T. gondii* [4–6]. This opens opportunities for the study and evaluation of redox pathways as drug targets in these parasites. Cysteine plays an important role in redox homeostasis, being itself a potent antioxidant and a precursor of glutathione (GSH). We have recently shown that *T. gondii* possesses an intact reverse transsulfuration pathway for the generation of L-cysteine (Cys) from methionine through the two enzymes cystathionine β-synthase (CBS) and cystathionine γ-lyase (CGL) [7,8] (**Supp. Fig. S1**). CBS condenses L-serine (Ser) and L-homocysteine (Hcys) to L-cystathionine (Cth) and H_2O via a β-replacement reaction that is due to the cofactor PLP. Cystathionine is then cleaved to cysteine by a second PLP-requiring enzyme, CGL (Fig. 1**A**).

CBS is a complex modular protein present in a wide variety of organisms from all kingdoms, for which five different domain distributions (herein referred as classes A to E) and two major oligomers (dimers and tetramers) have been described (Fig. 1**B**) [9]. The simplest architecture, (class A), is found in bacteria and consists of a catalytic domain with the folding of PLP-dependent enzymes of the type II with the cofactor covalently attached to the ε-amino group of a conserved lysine via a Schiff base bond. This type of CBS protomer is the most studied structurally (PDB IDs 6AHI, 5HBG (*Helicobacter pylori*) [10], 5XW3 [11] (*Bacillus anthracis*), 5B1H, 5B1I (*Lactobacillus plantarum*) [12], 4OFX (*Coxiella burnetii*) (Fig. 1**C**) and self-assembles into active dimers from which reaction intermediates have recently been isolated [11,12]. A rare variant of this CBS, class B, occurs in *Caenorhabditis elegans* (*Ce*CBS) and contains two catalytic domains in tandem arrangement, only one of which binds PLP [9,13]. No structural data is currently available on this CBS type. The next complexity level, class C, occurs in yeast [14] and in some Apicomplexa [7] (Fig. 1**C**), but can also be found in Gram negative bacteria such as *Pseudomonas aeruginosa*, whose CBSs retain the catalytic domain and contain an additional C-terminal Bateman module. The Bateman module in turn consists of two so-called CBS motifs (CBS1, CBS2) whose function is still unknown [15,16]. The crystal structure of a truncated construct containing the catalytic core of *Saccharomyces cerevisiae* CBS bound to some reaction intermediates represents the only structural data available on this class (PDB 6C2Z, 6C4P, 6C2Q, 6C2H) [17]. The next CBS structure type, class D, is found in some insects, e.g., fruit fly [18] (PDB IDs 3PC2, 3PC3, 3PC4) and honeybee [19] (PDB ID 5OHX), and contains an N-terminal heme-binding domain of unclear function that precedes the catalytic domain. The only two crystal structures available from this CBS type [18,19] indicate that the fundamental role of the Bateman module in class D enzymes lies in stabilizing a constitutively activated dimeric assembly without playing any additional role in the regulation or activity of the enzyme. The class D CBSs exist in a sole constitutively active conformation (Fig. 1**C**). Finally, class E, the most evolved version of the enzyme, is found in mammals, including humans, and has the same domain distribution as flies and bees (Fig. 1**B**). However, in this case the Bateman module becomes more important and not only determines the degree of oligomerization of the enzyme, which includes dimers, tetramers or even larger aggregates, but also becomes an allosteric module that regulates the degree of enzyme activity upon binding of S-adenosylmethionine (AdoMet) [9,20–23] (Fig. 1**B**). Importantly, class E CBS can adopt two distinct conformations (Fig. 1**C**). The first, or *basal* conformation (PDB ID 4L0D), is less active and forms basket-shaped dimers in which the Bateman module of each subunit is placed over the entrance of the catalytic cavity of the complementary counterpart. This arrangement impedes the entry of substrates into the active site [24,25]. Binding of AdoMet to the Bateman module causes each protomer to enter the so-called *activated state* [26] (PDB IDS 4PCU, 4COO), which is fivefold more active and associates in shell-like dimers in which the autoinhibitory effect of the Bateman module is abolished. This activated conformation of mammals CBS is structurally equivalent to the activated conformer found in insects and is only maintained while AdoMet is bound to the Bateman module. Disassembly of AdoMet necessarily involves the return of the protein to its basal conformation. The CBS from *T. gondii* (*Tg*CBS) belongs to class C since it lacks the N-terminal heme-binding motif preceding the catalytic domain but possesses the Bateman module. As previously reported [7], *Tg*CBS can use both Ser and *O*-acetylserine (OAS) to produce Cth, linking *Tg*CBS to other enzymes such as the *O*-acetylserine-dependent cystathionine β-synthase (OCBS) [10] and *O*-acetylserine sulfhydrylase (OASS) that also use OAS as a substrate. Unlike *Homo sapiens* CBS (*Hs*CBS) [7], *Tg*CBS does not respond to stimulation with AdoMet, but the reasons for the observed catalytic activity without the help of the AdoMet allosteric effector were not clear so far. In addition to a role in cysteine biosynthesis, *Tg*CBS is also strongly implicated in the production of hydrogen sulfide (H_2S) by highly efficiently catalyzing the β-replacement of Cys with Hcys [7]. Thus, H_2S may act as an important gaseous modulator of oxidative stress in this obligate parasite.

Herein, we present the 3.2 Å resolution crystal structure of a modified *Tg*CBS Δ466-491 construct that, like the wild-type *Tg*CBS protein, contains the entire catalytic core and a C-terminal Bateman module. In our construct, the Bateman module lacks residues 466-491, which belong to a flexible loop that protrudes from the central β-strand of the CBS2 domain. The structure of *Tg*CBS Δ466-491 has been determined alone and in complex with the aminoacrylate intermediate formed from either Ser or Cys substrates in the β-replacement reaction. A structure-based molecular dynamics (MD) study comparing *Tg*CBS, *Hs*CBS and the pathogenic variant D444N of *Hs*CBS has allowed us to identify the key features that permit *Tg*CBS existing in a sole catalytically active basket-like conformation. A principal component analysis (PCA) of the MD trajectories has revealed a distinct dynamic behavior of the three proteins that helps explain their different activity in the absence of AdoMet. The *Tg*CBS structure represents the first structural insight into a class C CBS, and adds one missing piece into the scarce structural knowledge on the CBS enzymes which so far was limited to types A, D and E. Moreover, the insight gained from this work will help in understanding the implication of CBS in regulating the intracellular redox homeostasis in apicomplexan parasites and will facilitate the structure-guided inhibitor design.

2. Results

***Toxoplasma gondii* CBS associates in active dimers.** Careful sequence and structural alignments between *Tg*CBS and *Hs*CBS (**Supp. Fig. S2**) revealed an unusually extended loop (residues 463-496) between the last two β-strands of the CBS2 motif in *Tg*CBS which, as observed in humans, is presumably flexible and may impede crystal growth. Accordingly, we engineered a protein construct (*Tg*CBS Δ466-491) that lacks 25 amino acid residues at the tip of this internal loop. Importantly, this modification of the peptide chain made possible the crystallization of the enzyme and improved the diffraction quality of the crystals. *Tg*CBS Δ466-491 was overexpressed in *Escherichia coli* and purified as His-tagged protein (located at the N-extreme) to greater than 95% purity as judged by the presence of a single band on SDS-PAGE (Fig. 2**A**). Through absorbance and circular dichroism (CD) studies, we ascertained that the mutant displayed no substantial differences compared to the wild type protein [7] in either the absorbance spectroscopic features (Fig. 2**B**) or protein conformation (secondary structure elements) (Fig. 2**C**). Moreover, thermal

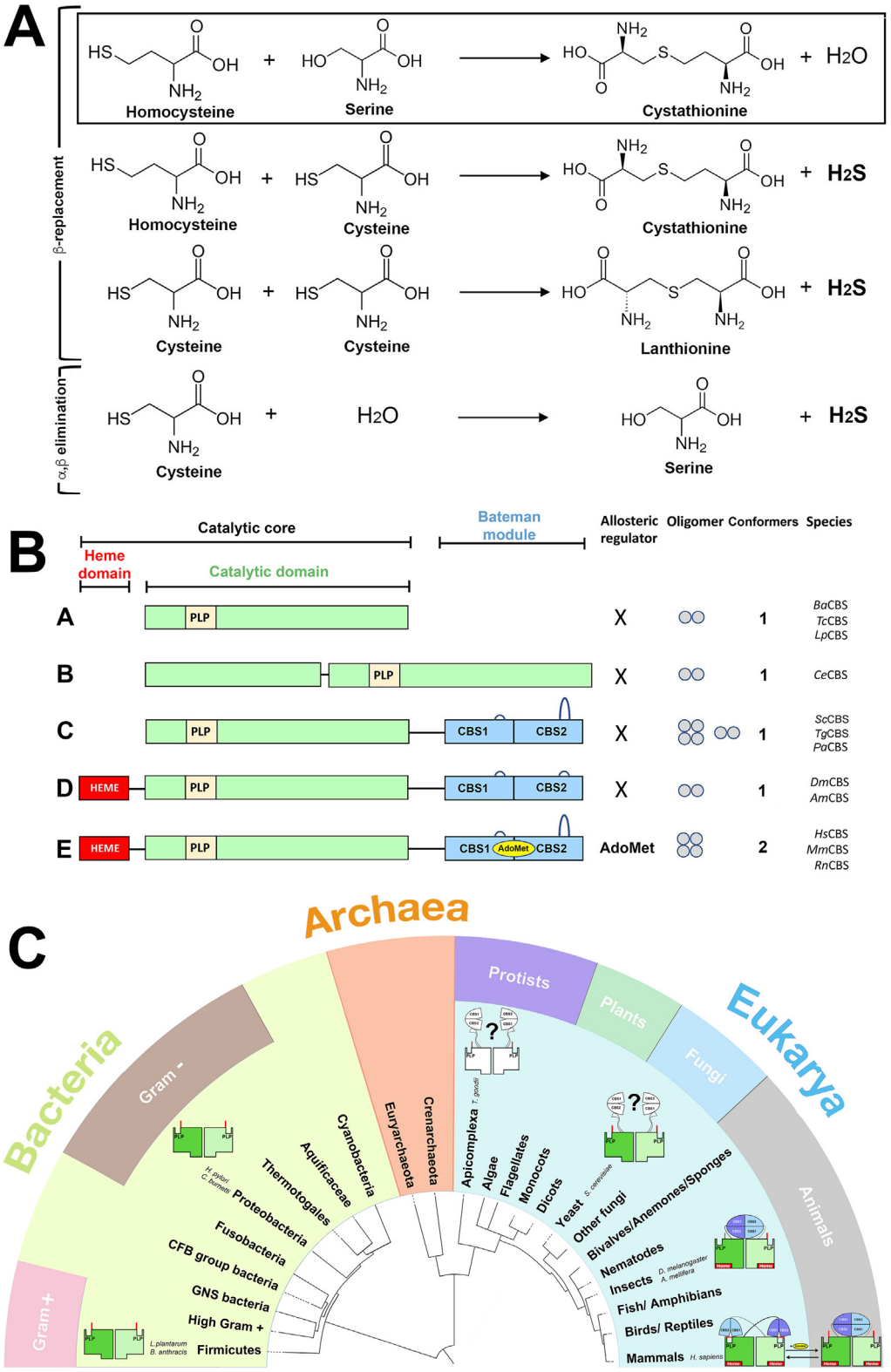


Fig. 1. CBS reactions & architectures. (**A**) Reactions catalyzed by CBS enzymes. The canonical reaction is framed. (**B**) CBS architectures. Class **A**, found in *B. anthracis* (*Ba*CBS), *L. plantarum* (*Lp*CBS) and *T. cruzi* (*Tc*CBS); Class **B**, found in *C. elegans* (*Ce*CBS); Class **C,** found in *T. gondii* (*Tg*CBS), *S. cerevisiae* (*Sc*CBS) and *P. aeruginosa* (*Pa*CBS). Class **D,** found in *D. melanogaster* (*Dm*CBS) and *Apis mellifera* (*Am*CBS). All CBSs but class C, usually show a long loop at CBS2. Class **E**, found in mammals, binds AdoMet. The small circles indicate the oligomerization degree. (**C**) Phylogenetic tree of organisms encoding the CBS enzyme. The nine available crystal structures of CBSs are indicated. Unsolved domains are in white.

stability studies by CD measurements at 222 nm resulted in similar denaturation profiles, indicating that the mutation did not affect the stability of the protein (Fig. 2**D**). The deletion of the region 466-491 did not even alter the oligomeric state of the protein since the *Tg*CBS Δ466-491 variant exists primarily as a dimer with a minor percentage of tetramer as observed for the wild type protein

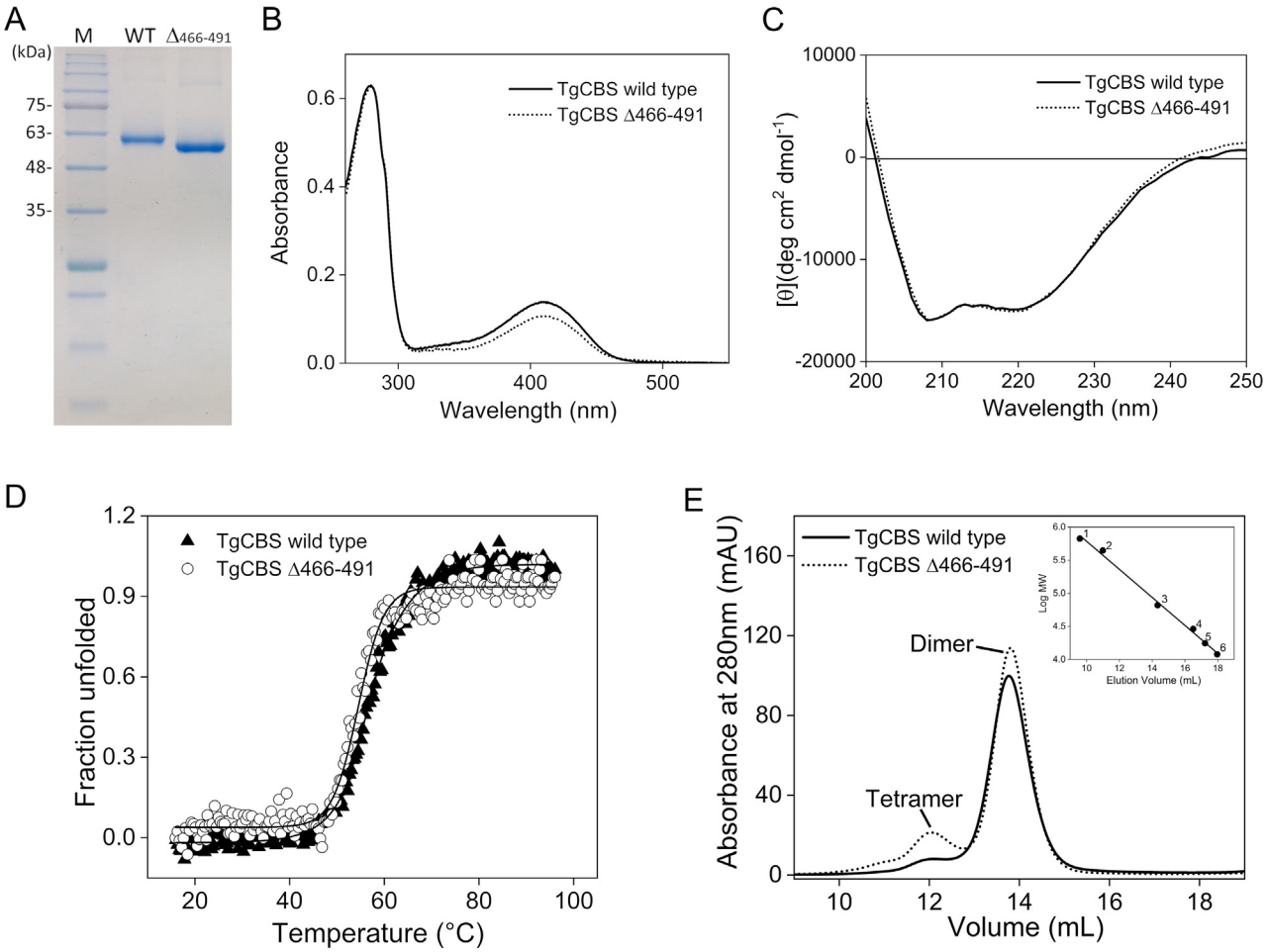


Fig. 2. Properties of *TgCBS* Δ466-491. (**A**) 12% SDS–PAGE analysis of purified recombinant *Tg*CBS wild type and *Tg*CBS Δ466-491. Lane M, protein marker. (**B**) UV–visible absorption spectra of 15 μM purified *Tg*CBS wild type (solid line) and *Tg*CBS Δ466-491 (dotted line) recorded in 20 mM sodium phosphate buffer pH 8.5. (**C**) Far-UV CD spectra of *Tg*CBS wild type (solid line) and *Tg*CBS Δ466-491 (dotted line) at 0.2 mg/mL in 20 mM sodium phosphate buffer pH 8.5. (**D**) Thermal denaturation of *Tg*CBS wild type (solid triangles) and *Tg*CBS Δ466-491 (open circles) recorded following ellipticity signal at 222 nm at 0.2 mg/mL protein concentration in 20 mM sodium phosphate buffer pH 8.5. (**E**) Gel filtration chromatography of *Tg*CBS wild type (solid line) and *Tg*CBS Δ466-491 (dotted line) at 1 mg/mL using a Superdex 200 10/300 GL column in 20 mM sodium phosphate, 150 mM NaCl buffer pH 8.5. *Inset*, calibration curve of logarithm of the molecular weight *versus* elution volume. The standard proteins used were: 1, thyroglobulin; 2, apoferritin; 3, albumin bovine serum; 4, carbonic anhydrase; 5, myoglobin; 6, cytochrome *c*.

(the percentage of molecules in dimeric and tetrameric form were estimated to be 92% and 8% for *Tg*CBS wild-type, and 82% and 18% for *Tg*CBS Δ466-491, respectively) (Fig. 2**E**).

The continuous assay for Cth production, employing cystathionine β-lyase (CBL) and lactate dehydrogenase (LDH) as coupling enzymes, was employed to determine the steady-state kinetic parameters of *Tg*CBS Δ466-491 as previously described for the wild type protein [7]. Except for a 1.4-fold difference in K_m^{Ser}, the kinetic parameters of the β-replacement reaction catalyzed by *Tg*CBS Δ466-491 and wild type proteins are identical, within their experimental error, demonstrating that the removal of the of 25-residue loop does not alter activity (**Supp. Table S1**). No response to AdoMet was observed as for the wild type protein (**Supp. Table S2**).

Overall Structure of TgCBS. The *Tg*CBS Δ466-491 construct yielded crystals belonging to space groups $P3_1$ and $P3_121$, and diffracting X-rays from 3.15 to 3.6 Å resolution (**Supp. Table S3**). In agreement with its behavior in solution, *Tg*CBS Δ466-491 forms symmetric dimers in the crystals (Fig. 3). Each protomer is composed of two independent blocks, consisting of a N-terminal catalytic domain (residues 1–323) and a C-terminal Bateman module (residues 353–514), tethered by a long linker (residues 324–352) that contains two α-helices. At first glance, the most striking characteristic is not the intrinsic folding of each of these domains, but their relative orientation, which unexpectedly reproduces the arrangement observed in the *basal conformation* of the human enzyme (Fig. 3) [22,23]. In such conformation the enzyme self-associates in dimers that adopt a basket-shaped domain-swapped symmetrical structure where the catalytic core of each protomer interacts with both the catalytic core and the regulatory domain of the complementary monomer (Fig. 3**A**). Meanwhile, the regulatory domain is strategically positioned above the entrance of the catalytic crevice of the complementary monomer, thus hampering the access of substrates and retaining the enzyme in an apparent inactive state. A priori, we found these findings surprising considering the inability of *Tg*CBS to bind or to be regulated by AdoMet.

Structurally, the catalytic domain of *Tg*CBS presents the overall fold of the β-family of the PLP-dependent enzymes and maintains its secondary elements basically unaltered with respect to the equivalent region in the human (Fig. 3**B, D**) [20,23,24], fly [18] and honeybee enzymes (Fig. 3**C**), [19]. The catalytic domain is composed of fourteen α-helices and two β-sheets consisting of four (β3–β6) and six (β1–β2, and β7–β10) strands, respectively (Fig. 3**A**). An additional β-strand precedes the last α-helix (α14) of the domain. A main difference with the human and insect pro-

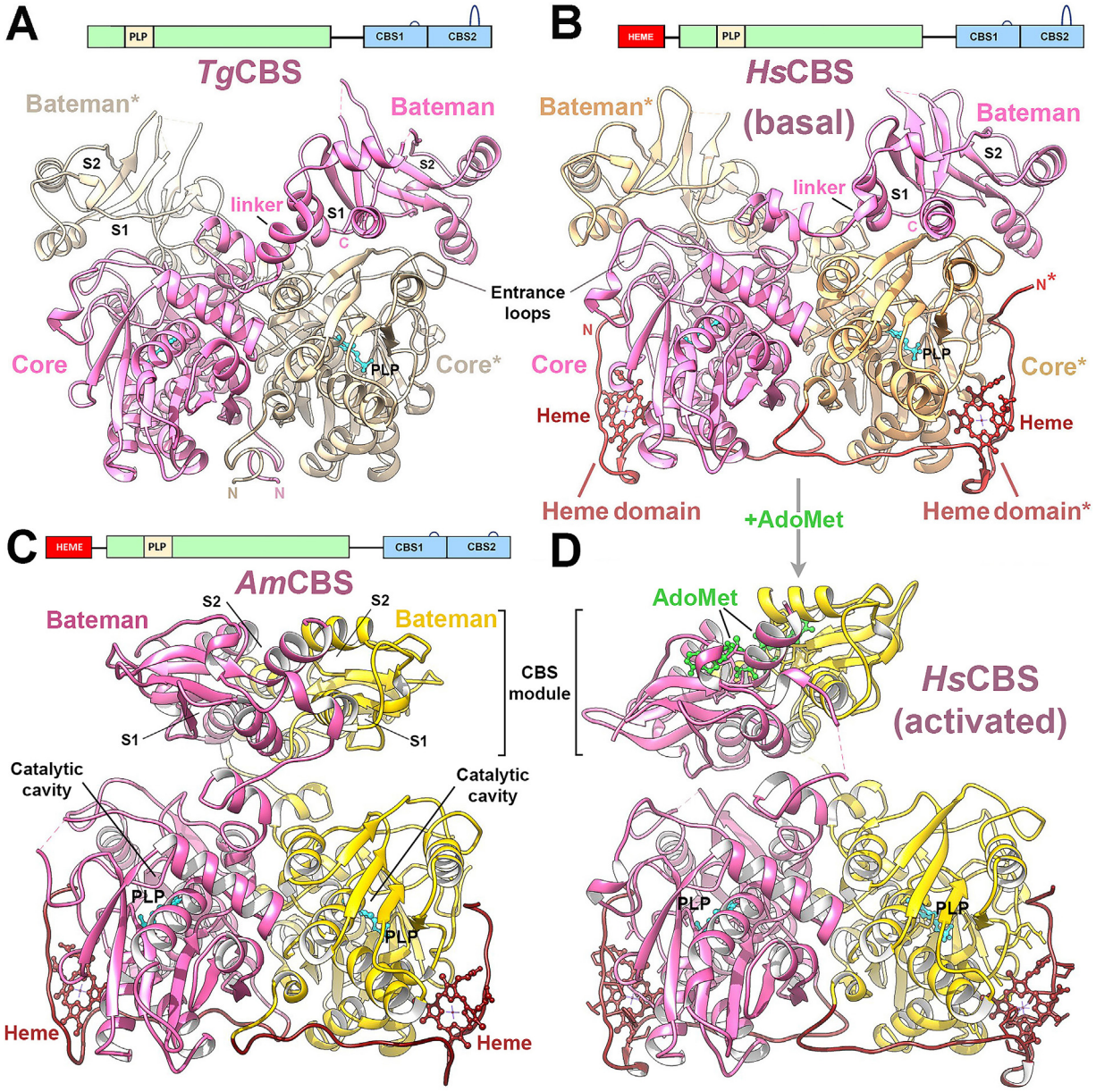
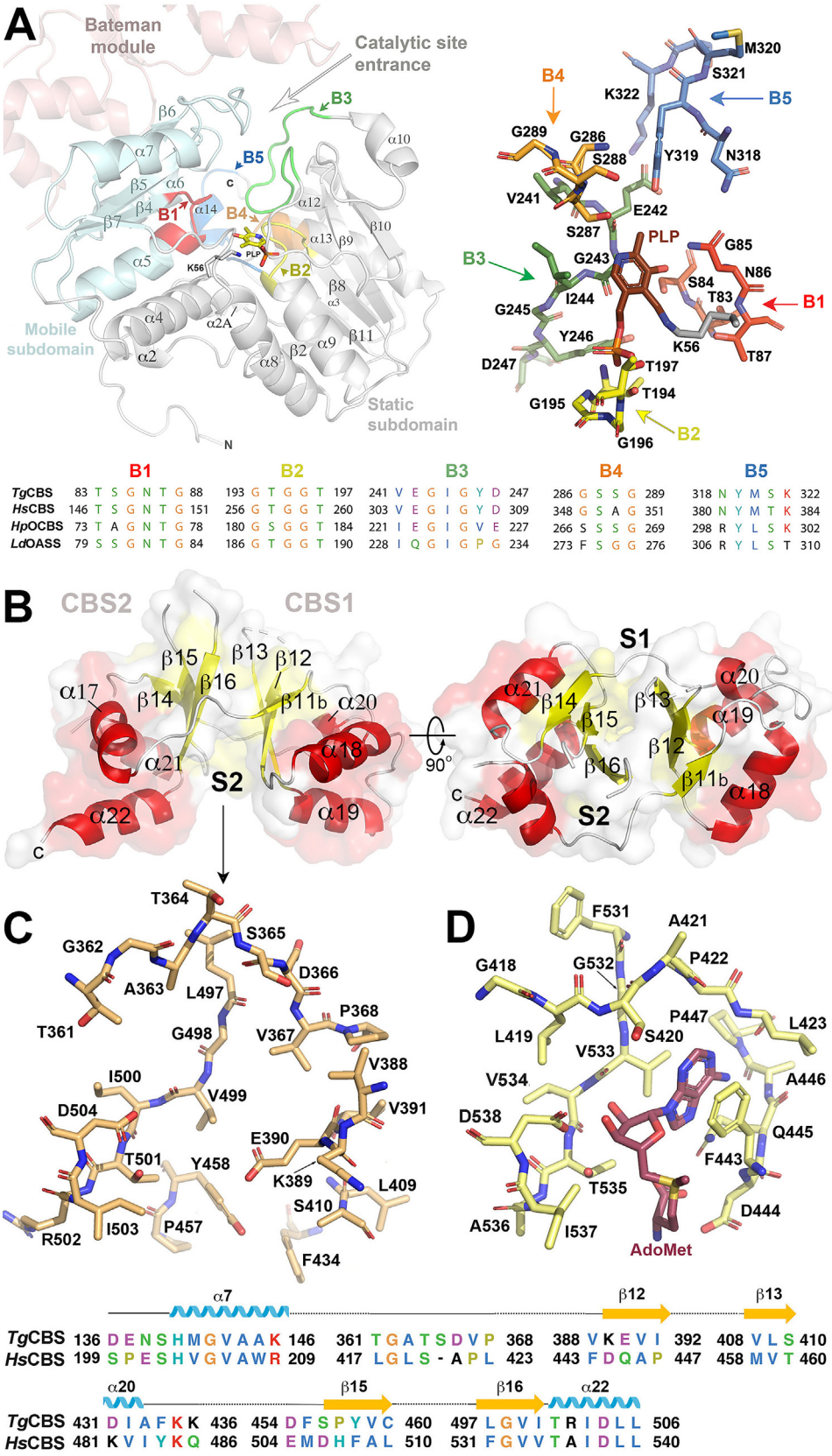


Fig. 3. Structure of *Tg*CBS. (**A**) Structure of the basket-shaped *Tg*CBS dimer (PDB ID code 6XWL). The complementary subunits are colored in *yellow* and *pink*, respectively. (**B**) Basket-shaped *Hs*CBS dimer (*basal state*) (PDB ID 4L0D) [24]. The heme-binding domain is in *red*. (**C**) Constitutively activated *sea bollard-shaped* dimer from *Apis mellifera* CBS (*Am*CBS) (PDB ID code 5OHX) [19]. (**D**) AdoMet-bound *sea bollard*-dimer of *Hs*CBS (*activated state*) (PDB ID 4PCU) [26]. The Bateman modules associate forming a disk-like *CBS module*. Heme, PLP and AdoMet are in *red*, *cyan* and *green* sticks, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

teins, but similar to the yeast enzyme [9], is the absence of a heme-binding domain (Figs. 1, 3) [7]. Two distinct blocks are clearly distinguishable in the catalytic domain (Fig. 4**A**). The larger one contains two segments of residues distanced in the peptide chain (residues 10–55 and 163–322) and ranges the so-called *static subdomain*. This block behaves as a rigid body during the catalysis in other CBSs and related enzymes, and represents the region that most interacts with other protein domains [19]. In *Tg*CBS, it contributes significantly to maintaining the integrity of the dimer by providing hydrophobic residues (c.a L18, V27, M29, I89, L93, V97, I117, C120, L121, I277, L282, M320, I324) at the interface between complementary subunits. The overall architecture of the static subdomain is mostly conserved when compared with the wild type human enzyme. A least square superposition of the Cα-positions of this region between *Tg*CBS and *Hs*CBS yields a root mean square deviation of only 0.48 Å and 0.58 Å at the monomer and dimer level, respectively. Importantly, this static block configures half the entrance to the catalytic site (Fig. 4**A**) [17,18] and contains four out of the five consensus sequences that participate in the interaction with the different CBS substrates (blocks B2, B3, B4 and B5 in Fig. 4**A**) [10]. The main differences with the human protein are in the region comprising residues 232–238 of block B3 (Fig. 4**A**), which is more directed towards the catalytic cleft in *Tg*CBS, due to a somewhat longer helix α10 (residues 228–234). The second subdomain of the catalytic core is the *mobile subdomain* [19] which is smaller in size (residues 75–157) and is intercalated in the static block, to which is linked by two loops that connect helix α4 with strand β4, and strand β7 with helix α8, respectively (Fig. 4**A**). The mobile subdomain provides the second half of the walls of the catalytic cavity, and acts as a cap that limits the access of substrates into the narrow channel where the PLP molecule is covalently attached via a Schiff base bond to the ε-amino group of lysine



	B1			B2			B3			B4			B5		
*Tg*CBS	83	T S G N T G	88	193	G T G G T	197	241	V E G I G Y D	247	286	G S S G	289	318	N Y M S K	322
*Hs*CBS	146	T S G N T G	151	256	G T G G T	260	303	V E G I G Y D	309	348	G S A G	351	380	N Y M T K	384
*Hp*OCBS	73	T A G N T G	78	180	G S G G T	184	221	I E G I G V E	227	266	S S S G	269	298	R Y L S K	302
*Ld*OASS	79	S S G N T G	84	186	G T G G T	190	228	I Q G I G P G	234	273	F S G G	276	306	R Y L S T	310

*Tg*CBS	136	D E N S H M G V A A K	146	361	T G A T S D V P	368	388	V K E V I	392	408	V L S	410
*Hs*CBS	199	S P E S H V G V A W R	209	417	L G L S - A P L	423	443	F D Q A P	447	458	M V T	460

*Tg*CBS	431	D I A F K K	436	454	D F S P Y V C	460	497	L G V I T R I D L L	506
*Hs*CBS	481	K V I Y K Q	486	504	E M D H F A L	510	531	F G V V T A I D L L	540

Fig. 4. *Tg*CBS domains. (**A**, *left*) The catalytic domain of *Tg*CBS includes the *mobile* (*cyan*) and the *static* (*grey*) subdomains. The Bateman module is in *light red*. (**A**, *right*) Residues configuring the catalytic cavity. PLP (*brown*) is covalently bound to K56. (**A**, *bottom*) Sequence alignment of blocks configuring the active site. Block B1 (*asparagine loop, red*), stabilizes the aminoacrylate intermediate and contains a conserved serine (S84) in all CBSs and OASS enzymes. Block B2 (*phosphate loop, yellow*) anchors the phosphate moiety of PLP. Block B3 (*green*) is thought to interact with the second substrate, Hcys, and participates in the nucleophilic attack on aminoacrylate, in the regeneration of the internal aldimine, and in the closure of the active site [41]. Block B4 (*orange*) stabilizes the pyridine ring of PLP. Block B5 (*blue*) interacts with the lip of the active site. (**B**) Bateman module of *Tg*CBS. S1 and S2, are the main cavities. (**C**) Residues within cavity S2 in *Tg*CBS; (**D**) Residues in site S2 of *Hs*CBS. AdoMet is in *wine* sticks. *Bottom*: Sequence alignment of residues at cavity S2 in *Tg*CBS and *Hs*CBS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(K56) [27]. Several H-bonds anchor the PLP molecule to the protein matrix and orient the cofactor appropriately within the cavity. Among them are those formed between the nitrogen of the pyridine and the Oγ of residue S287, and between the 3́-hydroxyl group of PLP and the $N_{\delta 2}$ of residue N86 (Fig. 4**A**). Importantly, N86 is coplanar with the pyridine ring of PLP to allow the appropriate ring tilt upon transaldimination [23]. On the opposite side of PLP, the phosphate moiety interacts with the so-called *phosphate binding loop* (block B2 in Fig. 4**A**, residues 193–197), which is located between strand β8 and helix α9. This loop includes two important threonine residues (T194 and T197) as well as three glycines that form a network of H-bond interactions that ensure the correct orientation of PLP. Mutation of the equivalent threonines in *Hs*CBS is known to cause a loss of CBS activity, and impairs carbonylation of the heme moiety [22]. Importantly, in *Tg*CBS the mobile subdomain collapses towards the catalytic cleft pushed by the Bateman module, which is placed just above (Fig. 4**A**). A similar arrangement is observed in the basal state of the human enzyme [24,25] (Fig. 3**B**), where the loops L145–148, L171–174, and L191–202, defining the entrance of the catalytic site, are sandwiched between the core and the regulatory module, thus impairing free access of substrates into the PLP site. Finally, the C-terminal domain of *Tg*CBS comprises a Bateman module (~20 kDa) made up of two CBS motifs (CBS1, residues 354–425; CBS2, residues 426–514) that are connected to each other by a linker of eight amino acid residues (421–428). Both, CBS1 and CBS2 show the canonical βαββα sequence of secondary elements usually found in these types of motif [15], and contact each other via their three-stranded β-sheets. As observed in humans [22,23], the Bateman module of *Tg*CBS presents two symmetrical cavities (S1, S2) (Fig. 4**B**) of which only S2 is exposed to solvent. Site S2 conserves some of the residues that would be necessary to stabilize one AdoMet molecule inside (c.a. T501 and D504 that would interact with the ribose ring of the nucleotide, or I503 and V499 that would help in accommodating the methionyl group of AdoMet), but misses other key features to host AdoMet. Among them is a longer peptide segment comprising residues 361–368 (417–423 in *Hs*CBS, Fig. 4**D**), that stands out preceding the strand β11b, and displaces residue T364 away from the center of the S2 cavity, thus preventing its potential interaction with the ribose ring of AdoMet. The extension of this segment distorts the hydrophobic pocket otherwise required to accommodate and orient the adenine ring of AdoMet inside the cavity [26]. Additionally, a negatively charged glutamate at position 390 invades the position that the AdoMet adenine ring would potentially occupy (Fig. 4**C, D**). All these features disable *Tg*CBS from binding AdoMet.

TgCBS exists in a unique basal-like conformation. A careful comparative analysis of the available CBS three-dimensional data has led us to propose that the basal-type folding of *Tg*CBS represents its sole conformational state. We base our conclusion on three key structural features. The first one is the inability of *Tg*CBS to bind AdoMet at the site S2 of the Bateman module (Fig. 4**C**), and thus to suffer an AdoMet-induced relative rotation of its two CBS motifs. In allosterically regulated CBS enzymes as *Hs*CBS, such rotation weakens the interactions that maintain the Bateman module anchored to the catalytic core just above the entrance of the catalytic site, and this makes the enzyme progress towards the activated state [26]. It is well established that the migration of the Bateman module allows the free access of substrates into the PLP cleft [26]. Nevertheless, such displacement of the Bateman module can only occur if its contacts with the catalytic core are not too tight. If this were the case, both domains would remain permanently anchored to each other (as it happens in *Tg*CBS). The second relevant feature refers to the peptide linker connecting the Bateman module with the catalytic core, which needs to be sufficiently long and flexible as to facilitate the migration of the regulatory domain to its final location in the activated state. Our structures show that, although the length of the *Tg*CBS linker is sufficient, the interactions that it maintains with the Bateman module and with the core are stronger than in *Hs*CBS (Fig. 5**A, B**). These contacts significantly enhance the stability of the basket-like conformation vs other possible arrangements, and occur thanks to a specific turn of the interdomain linker in the parasite enzyme, around residues 342–344, that allows positioning helix α16 closer to helices α6, α21 and α22 and to strands β5 and β6, than in the human protein (Fig. 5**B**). This peculiar location of helix α16 in *Tg*CBS results in the formation of a multiple network of salt bridges participated by residues K341, E344 and R345 (α16), K119 (α6), D454 (α21), R502 (α22), R101 (β5), and E124 (β6). These attractive electrostatic interactions complement the effect of the hydrophobic pocket formed by residues F349 from helix α16 and residues L505 and L509 from helix α22, and keep the interdomain linker of *Tg*CBS firmly attached to both the catalytic core and the Bateman module, making the displacement of the latter highly unlikely.

Finally, the third requirement in an allosterically regulated CBS enzyme to reach the activated state consists in being able to couple its Bateman module with the equivalent region from the complementary subunit to form a flat disk-shaped structure, known as "*CBS module*". Once formed, this assembly locates far from the entrance of the catalytic cavity, thus allowing free access of all substrates (**Supp. Fig. S3**) [26]. The formation of the CBS module necessarily requires the previous rotation of the CBS motifs mentioned above, which only occurs upon binding of one AdoMet molecule to the S2 site (Fig. 3) [26]. The formation of the CBS module also involves establishing new interactions between the interfacial helices of the CBS domains (α18, α19 from CBS1; α21, α22 from CBS2) (Fig. 5**C-F**), (**Supp. Fig. S3**). It is worth mentioning that in all known CBS enzymes, the CBS module is antiparallel (the two Bateman modules are oriented in a head-to-tail manner) [18,19,26]. In the said arrangement, the CBS1 of the first subunit interacts with the CBS2 motif of the complementary monomer and vice versa (**Supp. Fig. S3**). Obviously, if the residues at the interfacial positions do not favor the corresponding interactions, the CBS module cannot be assembled. Said this, we found that *Tg*CBS does not fulfil the third condition for activation. The independent structural superimposition of the CBS1 and CBS2 domains of *Tg*CBS and *Hs*CBS or *Tg*CBS and *Am*CBS are good (rmsd 0.902 /0.816 and 1.147/0.719, respectively), but significantly worsens when the structural alignment is performed comparing the entire Bateman module of *Tg*CBS with the corresponding region from the AdoMet-bound activated form of the human enzyme (rmsd = 1.37) or to the constitutively activated *Am*CBS (rmsd = 4.55). In its native conformation, the Bateman module of *Tg*CBS is not compatible with the assembly of a CBS module, as its interfacial helices would not face those of its complementary counterpart (**Supp. Fig. S3**). Self-assembly of Bateman modules to form a CBS module of *Tg*CBS would cause clashes between some residues (c.a K387) and a bound adenosine derivative (Fig. 5**C**). In such hypothetical CBS module, some favorable hydrophobic interactions could be potentially established between residues of the interfacial helices α18, α19, α21 and α22 of complementary CBS1 and CBS2 motifs (c.a L446, V449, L453 from α21; Q419, A415 from helix α19), but would also be impaired by steric clashes and/or repulsive forces between bulky or charged residues of the complementary subunits (c.a H507/H384) (Fig. 5**C**). Put together, all these findings strongly suggest that the basal-like fold of *Tg*CBS represents its unique conformation and that the enzyme does not progress to a second activated form structurally similar to the AdoMet-bound form found in mammals [26] and the constitutively active species of insects [18,19].

TgCBS is catalytically active maintaining a basal-like overall fold. At first sight, the basal-like architecture of the crystallized

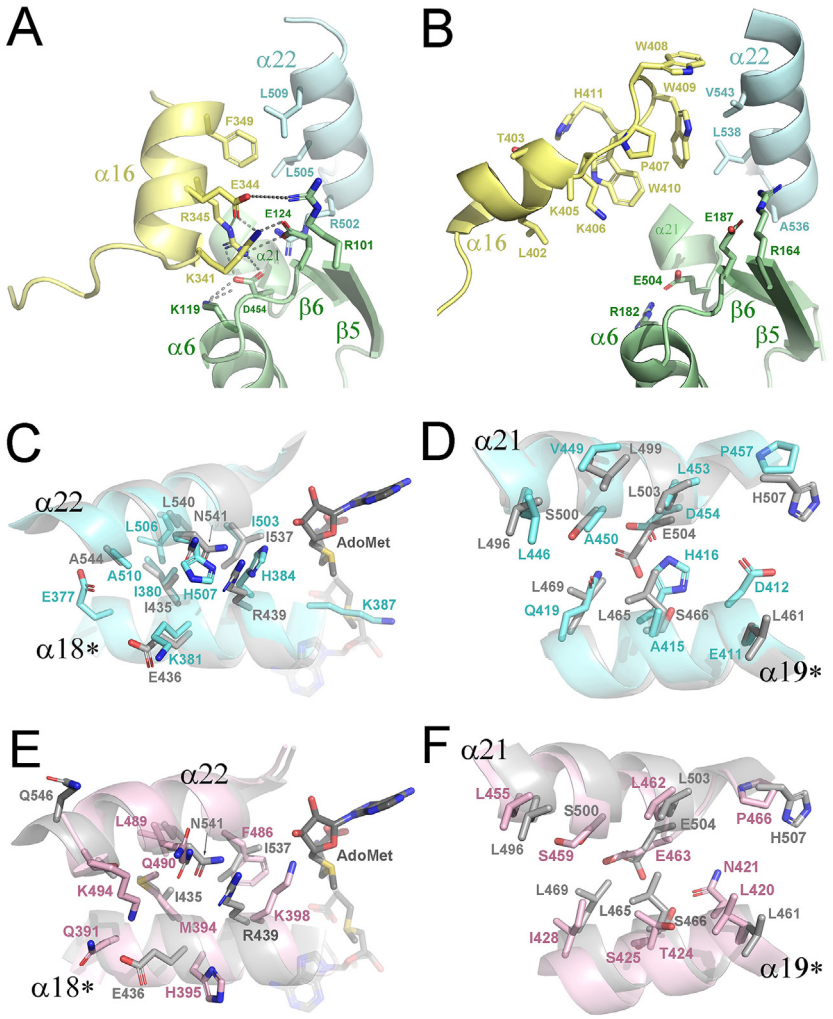


Fig. 5. (**A**) **Interdomain interactions in *Tg*CBS, *Hs*CBS and *Am*CBS.** Interactions between the interdomain linker (*yellow*), the catalytic core (*green*) and the Bateman module (*cyan*), in *Tg*CBS (**A**) and in *Hs*CBS (**B**). Interface residues in a potential CBS module of *Tg*CBS (*cyan*) (**C**, **D**), and *Am*CBS (*pink*) (**E**, **F**) compared to that found in *Hs*CBS (*grey*) [37]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

*Tg*CBS and its closed active site would appear to reflect an inactive or poorly active enzyme. However, detailed 3D-alignments of *Tg*CBS with *Hs*CBS revealed a slight displacement of the complementary Bateman modules towards the central cavity existing between subunits (**Supp Fig. S4**). This shift also affects helices α18 and α22, and occurs without significant changes in the catalytic core (**Supp. Fig. S4A**). Interestingly, a similar shift was formerly observed in the pathogenic D444N mutant of *Hs*CBS (**Supp. Fig. S4B**) [24], which shows an approximately two-fold increase in basal activity and impaired response to AdoMet stimulation as compared with the wild type enzyme [28]. Taking into account that crystallographic data offer an average spatial-time snapshot of the protein structure captured in a local energy minimum under specific buffered conditions, we found reasonable to postulate that, as formerly observed in the D444N variant [24], the mentioned displacement of the Bateman module provides a greater dynamic freedom to the mobile subdomain of the core domain of *Tg*CBS that facilitates the access of the substrates to the interior of its catalytic site. To evaluate this hypothesis, we explored the conformational behavior of *Tg*CBS, *Hs*CBS and the D444N human variant through microsecond molecular dynamics(μs-MD) simulations (**Supp. Movies S1-S3**). To identify the most relevant conformational transitions occurring in each system, the dimensionality reduction technique of principal component analysis (PCA) was applied (**Supp. Movies S4-S6**). The analysis made it possible to decompose the MD trajectory followed by each protein and revealed three distinct structural changes that mostly affect the orientation of the Bateman modules. In the first one, that we have called *longitudinal displacement*, the Bateman module moves along its longest axis, which runs from the CBS1 to the CBS2 domain (Fig. 6). In the second one, or *lateral displacement*, the shift takes place perpendicularly to the previous movement and occurs along the shortest axis that passes between domains CBS1 and CBS2. Finally, the third change involves a rotation of CBS1 with respect to CBS2 around an imaginary axis that runs parallel to the central β-sheets of the CBS domains and that is perpendicular to the two axis mentioned before (Fig. 6). Interestingly, the *rotational change* is specific of the human enzyme and is similar to that induced by AdoMet binding at site S2 site in human CBS [26]. These findings suggest that the effect of AdoMet consists essentially in increasing and stabilizing one of the conformational populations already existing in the native protein, thus promoting the activation of the enzyme. The PCA analysis also revealed that each of the three proteins analyzed undergo two main the structural rearrangements, which occur in different percentages (Fig. 6). The different combination of movements in each protein has well-differentiated effects and determines the number and shape of tunnels that gives access into the corresponding active sites. We found that mutant D444N displays a wider channel cluster than the wild-type protein, and that *Tg*CBS behaves similarly to the D444N mutant (Fig. 6, (**Supp. Movies**).

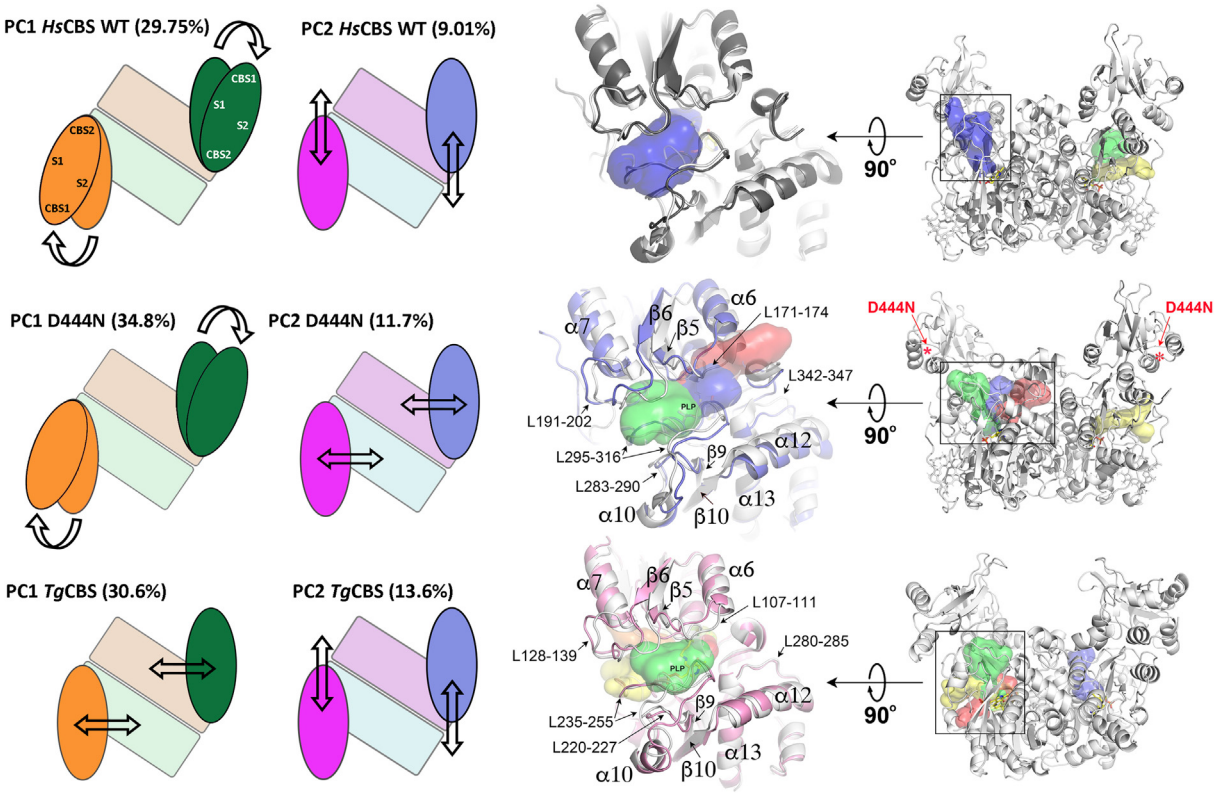


Fig. 6. MD analysis of *Tg*CBS, *Hs*CBS and mutant D444N*Hs*CBS-Δ516-525. (*Right*) The Bateman modules are in opaque ribbons. Heme and PLP are in sticks. (**C-E**) Structural overlap of the loops defining the entrance of the catalytic site at the narrower (*light grey*) and the wider amplitude points of the MD simulation in (*up*) basal wild-type *Hs*CBS, (*middle*) mutant D444N*Hs*CBS-Δ516-525 and (*bottom*) *Tg*CBS. *Hs*CBS, D444N*Hs*CBS-Δ516-525 and *Tg*CBS at the wider amplitude are colored in *black*, *marine* and *pink*, respectively. The clusters of tunnels formed along the MD simulation at the corresponding catalytic sites are shown in different colors. (*Left*) Main displacements suffered by the Bateman module of wt-*Hs*CBS (*up*), mutant D444N (*middle*) and *Tg*CBS (*bottom*) detected after PCA analysis of the MD trajectories. The figure schematizes a top view of the enzyme dimer. The Bateman module and the catalytic core are represented as an oval and a rectangle, respectively. Complementary subunits are colored in *green*/*orange* in PC1 and in *magenta*/*blue* in PC2, respectively. The corresponding displacements are indicated with arrows. Location of CBS1, CBS2, site S1 and site S2 within the Bateman module are also indicated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Intriguingly, the channel clusters were not symmetrically distributed between the two complementary subunits in any of the analyzed dimers, thus suggesting synergistic behavior between the two active sites. This helps explain the different degree of activity observed for each enzyme in the absence of AdoMet. The dynamic behavior of the Bateman module is also reflected in the degree of aperture of the entrance loops defining the access to the catalytic cavity in each protein. To define the aperture, we measured the angle between the α-carbons of residues K441, S387 of the first subunit, and K481 of the second subunit of the human enzyme along the MD trajectories. The equivalent angle between residues T386, S325 and D431 of *Tg*CBS was analyzed. As hypothesized, we found that mutant D444N and *Tg*CBS exhibited a significantly wider opening than the wild type human enzyme (Fig. 6). The impact of conformational dynamics on substrate accessibility to PLP resulted in the appearance of different transient access tunnels along the MD trajectories in each protein (Fig. 6).

To confirm that the crystallized enzyme is indeed catalytically active, we also crystallized *Tg*CBS with each of two substrates, Ser and Cys. In agreement with the MD studies, the Polder omit maps [29] showed additional electron density near the PLP molecule, consistent with the synthesis of the aminoacrylate intermediate. Of note, formation of the external aldimine in *Tg*CBS did not result in major distortions in the overall protein fold, and only caused subtle reorientations in the side chain of lysine K56, otherwise associated with PLP in the apoenzyme, and threonine T87, that rotates slightly to avoid clashes with the carboxylate moiety of aminoacrylate (Fig. 7). These small changes contrast with the behavior of other CBSs, such as the fly enzyme, where three loops (residues 119–122; 144–148 and 165–176) are displaced towards the PLP molecule upon addition of substrates. Formation of the aminoacrylate intermediate in the *Tg*CBS crystals is consistent with the behavior of the protein *in vitro*, where the addition of Ser or OAS results in the disappearance of the 412 nm-peak, representing the internal aldimine in the ketoenamine form, and the appearance of a major band centered at 460 nm in both absorbance and CD spectra, that is attributed indirectly to the aminoacrylate reaction intermediate (Fig. 7) [7].

3. Discussion

For decades, the limited structural data on the CBS enzyme has hampered the comparative study of its regulatory mechanisms across various species. Among the five distinct domain distributions identified for the enzyme (Fig. 1**B**), those containing the C-terminal Bateman module are the ones raising more incognitas. Indeed, such a Bateman module is present in the human enzyme, but also in the CBS of a variety of pathogens (such as *T. gondii* or *P. aeruginosa*), where it exerts different roles that are not yet understood. In mammals, the Bateman module is essential for the allosteric regulation of the enzyme after binding of AdoMet. In contrast, in other organisms such as the fruit fly or honeybee, it appears to have no function other than that of ensuring dimeric assemblies of constitutively activated enzymes. A confident prediction of the role of the intriguing Bateman domain and its position in overall protein folding is not possible at present, as it represents the least conserved region in all CBSs.

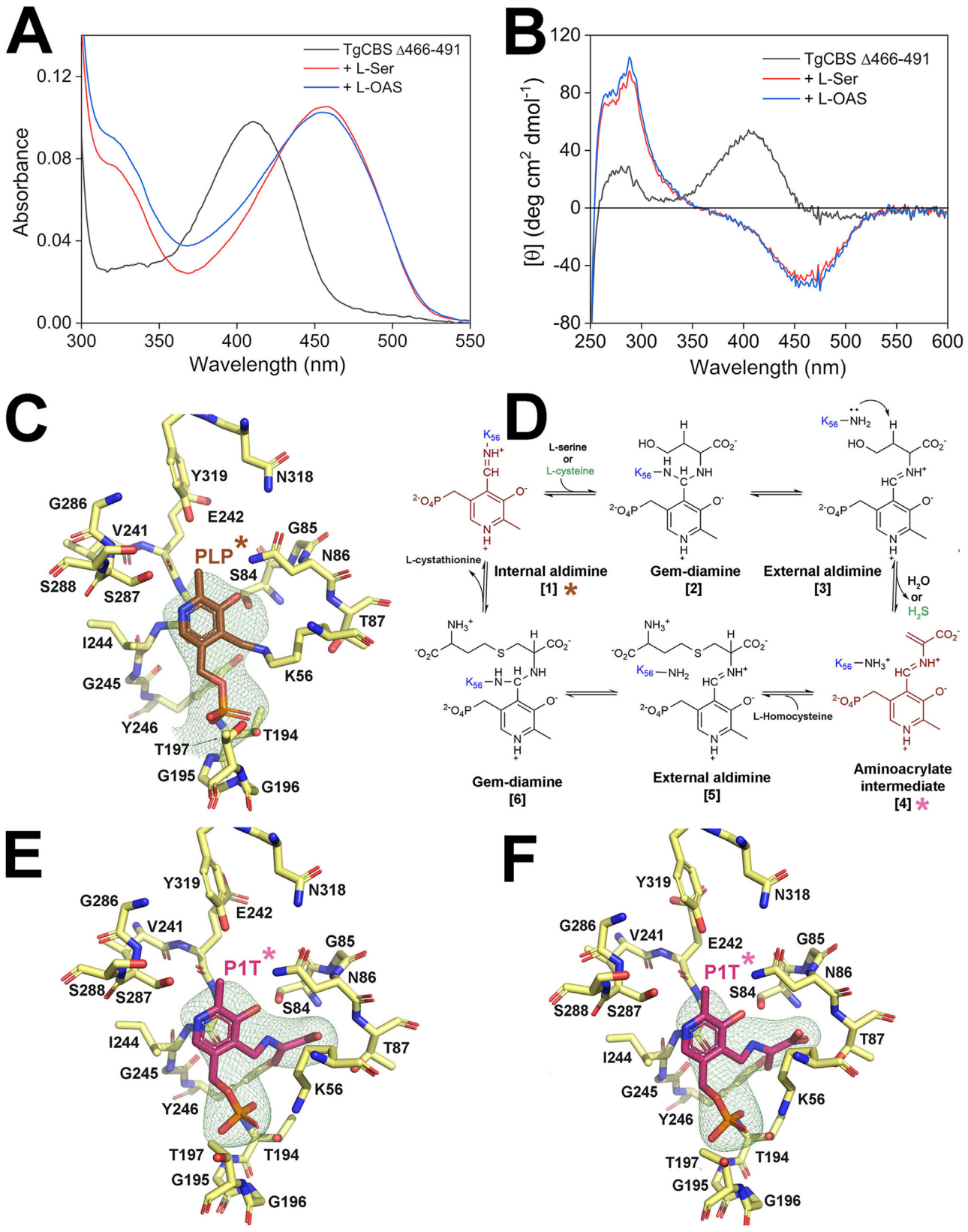


Fig. 7. *Tg*CBS is active. (*Top*) Spectra of *Tg*CBS Δ466-491 with substrates. (**A**) Absorbance spectra of 15 μM *Tg*CBS Δ466-491 alone (*black* line) and upon addition of 10 mM Ser (*red* line) or 10 mM OAS (*blue* line). **(B)** CD spectra of 1 mg/mL *Tg*CBS Δ466-491 alone (*black* line) and upon addition of 10 mM Ser (*red* line) or 10 mM OAS (*blue* line). (**C**) Fo-Fc Polder omit map around PLP calculated in the absence of substrates. (**D**) Mechanism of Cth formation from Ser (or Cys) and Hcys (adapted from [18]). (**C**, **E**, **F**) Zoom views of the *Tg*CBS catalytic site. The Fo-Fc Polder omit maps (contoured at 3σ) represent positive electron density around PLP or P1T for *Tg*CBS (**C**), and 24 h soaked crystals in a buffer containing 10 mM Ser (**E**) or Cys (**F**). PLP and P1T are pyridoxal phosphate and aminoacrylate intermediate, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In developmental terms, it seems established that the dimeric assemblies of CBS can exist in two distinct conformations, mostly recognized by the orientation of their Bateman modules. The first is a basket-shaped, and less active conformation (*basal conformation*), (Fig. 3**B**) while the second is a marine mooring bollard-shaped and significantly more active conformation (*activated conformation*) (Fig. 3**D**). The presence of the two conformations has been documented only in human CBS, where the transition from the first to the second conformation necessarily requires binding of the allosteric effector AdoMet at site S2 of the Bateman module. In less evolved organisms, for example the fruit fly or the honeybee, only a single constitutively activated type has been distinguished [18,19] (class D in Fig. 1), (Fig. 3**C**). This activated fold is structurally equivalent to the AdoMet-bound CBS form of mammals [26]. Intriguingly, the basket-like conformation had not been seen before in any other organism, and its presence in *Tg*CBS as a unique active species was completely unexpected (Fig. 3**A**).

Our structural and biochemical data showing the formation of a stable aminoacrylate intermediate upon addition of either Ser or Cys to the crystals, as expected for a PLP-catalyzed β-replacement reaction, confirms that the crystallized *Tg*CBS conformation corresponds to an active form of the enzyme (Fig. 7), and have enabled the identification of key distinguishing features that constitutively hold *Tg*CBS in a basket type fold and likewise, the features that would impede, or alternatively promote, the progression of other CBSs towards bollard-shaped activated states. These features comprise (i) the capacity (or inability) of the Bateman module to host AdoMet at site S2 (Fig. 3) [26]; (ii) the arrangement of interaction networks between the interdomain linker and the Bateman module that hinder (or not) the displacement of the latter (Fig. 5); and (iii) the capacity (or incapacity) of the Bateman module to frame an antiparallel CBS module (**Supp. Fig. S3**).

These results, by addressing the first proof of a basket-shaped conformation in the CBS enzyme of a lower eukaryote, reveal important structural details of class C CBSs, shared by Apicomplexa and yeast, therefore expanding the structural knowledge on CBS enzymes as well as their structure–activity relationship. In particular, in light of our new information, we propose that CBSs containing a Bateman module can be grouped into at least three categories: (i) *active basal-like species* (c.a *Tg*CBS) without the capacity to develop conformationally to higher activity bollard-like states; (ii) *constitutively active and non-allosterically regulated* CBS enzymes, similar to those found in *D. melanogaster* [18] and *A. mellifera* [19]; (iii) *allosterically activated* enzymes (ca. *Hs*CBS), the most advanced CBSs that can embrace two different conformations, one in which the enzyme is essentially less active (*basal*) and a second one (*activated*) with high activity that must be reached with the assistance of an allosteric molecule, for example, AdoMet, which briefly advances the change (Fig. 3).

Importantly, the comparative MD and subsequent PCA of *Tg*CBS, *Hs*CBS and the pathogenic mutant D444N help unravel what features determine the accessibility of substrate into the catalytic site, and provide clues on the ability of each enzyme to prioritize the wide variety of reactions described so far. The access into the catalytic site is not straightforwardly seen in the static, low-energy, and highly packed crystallographic structures, where in some cases the entrance loops may appear dumped towards the catalytic cavity, as observed in *Hs*CBS and *Tg*CBS crystals. The entrance tunnels are disclosed when thermal motions and solvent effects are simulated with μs-MD (Fig. 6), (**Supp. Movies**). Our MD investigation clarifies the catalytic activity shown by *Tg*CBS and reveals that inconspicuous rotations and translations of the Bateman module are adequate to make fleeting passages that associate the catalytic cavity with the outside. The MD analysis has likewise uncovered the dynamics of these enzymes and gives a reasoning to the basal activity of wild type *Hs*CBS in the absence of AdoMet [9], which was not very much clarified from the proteins crystallized so far, that showed completely shut active sites, clearly conflicting with such basal activity. Of note, the size and the quantity of these passages are wider in *Tg*CBS and in the homocystinuria causing mutant D444N [24,28] than in wild type *Hs*CBS, thus explaining their higher activity levels. Moreover, the absence of the N-terminal heme-binding domain provides an additional passage access to the catalytic site in the parasite enzyme that is not observed in the human counterpart. This tunnel is framed by helices α8 and α9 and loops β7-α8 and α10-α11, and opens a door for therapeutic intervention through the design of molecules that specifically inhibit *Tg*CBS without affecting the human host.

Moreover, our MD and PCA of the *Hs*CBS dimer strongly suggest that the main role of AdoMet as allosteric activator consists essentially in stabilizing one of the conformational populations already existing in the native protein, more concretely that resulting from the arrangement described in PC1 in Fig. 6. We postulate that such population is more prompt to evolve towards the final activated state [26]. Our PCA analysis also revealed that *Tg*CBS, *Hs*CBS and mutant D444N undergo two principal structural rearrangements identified, which occur in different percentages in each protein (Fig. 6). The combination of the PCA analysis with the reported activity values of the analyzed enzymes in the absence of AdoMet, reveal a different contribution of the identified structural changes in the catalytic activity. The lower basal activity of *Hs*CBS with respect to the D444N variant is likely due to the longitudinal displacement of the Bateman module in *Hs*CBS. On the contrary, the lateral displacement is more efficient in favoring the aperture of the active site in D444N. Interestingly, the rotational displacement of the CBS domains appears to be exclusive of the human protein, suggesting that this structural transformation is key to progress from the basal towards the activated conformational state in the mammalian CBS. Instead, the lateral displacement of the Bateman module is observed in *Tg*CBS and in the D444N mutant, but not in wild type *Hs*CBS (Fig. 6), (**Supp. Movies**).

Finally, we have found that the different combinations of movements in the analyzed proteins have well-differentiated effects in the number and shape of tunnels that give access to the corresponding active sites. Our MD analysis shows that *Tg*CBS and mutant D444N display a wider channel cluster than wild-type *Hs*CBS. Intriguingly, the channel clusters are not symmetrically distributed between the two complementary subunits in any of the analyzed dimers. This suggests a synergistic behavior between the two active sites.

Taking everything into account, our structures of *Tg*CBS have disentangled the unique domain organization of this pivotal enzyme in the metabolism of sulfur amino acids in *T. gondii*. Because CBS represents an exceptionally important therapeutic target, this new information paves the way for the rational design of drugs that can modulate the activity of CBS and establishes the framework for a therapeutic intervention through the transsulfuration pathway in different organisms (c.a bacterial pathogens) [30–32]. Moreover, since Toxoplasma is an early branching eukaryote, the study of CBS enzyme could provide useful information about the early evolution of transsulfuration routes and H_2S signaling.

4. Materials and methods

Unless stated otherwise, all chemicals were purchased from Sigma. *Tg*CBS construct lacking the region 466-PSTKKQAGMGEHER AKISLRKAGNSR-491 (*Tg*CBS Δ466-491) was produced by site specific mutagenesis on the pET21a-TgCBS construct, using the QuikChange® site-directed mutagenesis kit (Agilent Technologies), according to the manufacturer's recommendations for multiple-site deletion in a single step. The primers used to introduce the

deletion were as follows: forward (5′-cgtttgtgtgatggacgagaag gaatgcccacatttcctgg) and reverse (5′-ccaggaaatgtgggcattccttctcgtc catcacacaaacg). The verified plasmid was transformed into *E. coli* Rosetta (DE3) expression host cells (Novagen). The conditions for expression and purification of the mutant were as described for the wild type protein [26]. The oligomeric state of the *Tg*CBS construct was determined by gel filtration analysis employing a GE Healthcare Superdex 200 10/300 GL column in 20 mM sodium phosphate buffer pH 8.5, 150 mM NaCl and 0.1 mM DTT as described elsewhere [7,33]. A calibration curve was constructed using the GE Healthcare high molecular weight gel filtration calibration kit, following protocols in [33] The CBS activity in the canonical reactions was determined by a previously described continuous assay for Cth production [7], which employs recombinant cystathionine β-lyase (CBL) from *C. diphtheriae,* produced in our lab [34,35] and lactate dehydrogenase (LDH) from rabbit muscle (Sigma-Aldrich) as coupling enzymes. The activity assay was performed using the dimeric fraction of the enzyme isolated by gel filtration. The kinetic parameters for the β-replacement reaction of *Tg*CBS (condensation of Ser and Hcys, and condensation of OAS and Hcys) were calculated, via global analysis in which both substrates are varied simultaneously, from the fit of the data to the following equation:

$$\frac{v}{E} = \frac{k_{cat} * [SA] * [SB]}{K_m^{SB} * [SA] + K_m^{SA} * [SB] * \left(1 + \frac{[SB]}{K_i^{SB}}\right) + [SA] * [SB]}$$

in which v is the initial velocity, E is the concentration of the enzyme, SA is the concentration of the first substrate, SB the concentration of the second substrate, k_{cat} and K_m are the catalytic and the Michaelis–Menten constants, respectively. Substrate inhibition at high concentration of Hcys was observed, thus the equation included a K_i^{SB} value, which represents the inhibition constant for substrate inhibition by Hcys.

Kinetic data are presented as the average ± standard error of the mean (SEM) of three to five repetitions using at least three independently purified protein batches. Data fitting was carried out with OriginPro8 (OriginLab) software (Version2008. OriginLab Corporation, MA, USA).

Absorption spectra were collected on a Jasco-V560 UV–Vis spectrophotometer in 20 mM sodium phosphate pH 8.5. Far-UV CD spectra (250–200 nm) were recorded on a Jasco J-1500 spectropolarimeter as previously described [7,8]. Thermal denaturation profiles were collected by measuring CD signal at 222 nm in a temperature range from 20 to 90 °C (scan rate 1.5 °C/min). Protein concentration was 0.2 mg/mL and measurements were performed using quartz cuvettes with a path length of 0.1 cm in 20 mM sodium phosphate pH 8.5.

For crystallization, the enzyme was buffer exchanged into 50 mM HEPES-NaOH, 150 mM NaCl, 0.1 mM DTT pH 7.5, flash-frozen in liquid N_2 and stored at −80 °C until use. The crystallization trials were carried out by the vapor-diffusion technique in a sitting drop format in 96-well MRC crystallization plates (Molecular Dimensions). The screening plates were set up automatically with a MOSQUITO nanodispenser (SPTLabtech), and incubated at a constant temperature of 293 K. Drops consisted of 200 nL of protein solution (at a concentration of 15 mg mL^{-1}) mixed with 200 nL of precipitant solution. The reservoir volume was 50 μL. The best crystals of *Tg*CBS Δ466-491 were grown at a protein concentration of 21 mg mL^{-1}, using the hanging-drop method at 293 K in 24-well VDX plates (Hampton Research) with drops consisting in 1 μL of protein solution mixed with 1 μL of precipitant solution (8% wt/vol PEG3350 and 0.1 M MES pH 5.7). To obtain the protein-aminoacrylate complexes, the substrate-free *Tg*CBS crystals were soaked overnight in a solution containing 8% wt/vol PEG3350 and 0.1 M MES pH 5.7, and either Ser or Cys (10 mM). Single crystals were subsequently cryoprotected (20% glycerol) and flash frozen in liquid N_2 [36]. All X-rays datasets were collected at Synchrotron beamlines XALOC (ALBA), and I03/I24 (DIAMOND, UK), and were processed with XDS [37], Pointless [38], Aimless [39], CCP4 [40] and Staraniso [41] from autoPROC [42]. Anisotropy correction was applied to the *Tg*CBS Δ466-491 and *Tg*CBS Δ466-491 + cysteine datasets, being the first more affected, with an spherical completeness of 79.8% (28.7% in the last resolution sphere) if processed isotropically. The structures were determined by molecular replacement methods with PHASER [43] from PHENIX [44] using the structure of the truncated 45-kDa *Hs*CBS (PDB ID code 1JBQ) [23] as the initial search model. Refinement was done with Phenix.refine [45]. Model was built with Coot [46]. Figures were done with Pymol (The PyMOL Molecular Graphics System, Version 2.2.3, Schrödinger, LLC) and UCSF Chimera (version 1.13.1; http://www.rbvi.ucsf.edu/chimera). Sequence alignments were done with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and represented with Chimera. The crystal characteristics and refinement statistics are in **Supp. Table S3**.

Molecular modelling and dynamics. The initial structure for the truncated *Hs*CBS (D444N*Hs*CBS-Δ516-525) variant and *Tg*CBS were built by homology modelling (Schrödinger Suite) using the x-ray structure of the wild-type homodimer at 2.0 Å resolution (PDB ID 4COO for *Hs*CBS) [25] and the structure reported in this study (*Tg*CBS), respectively as templates, with the assistance of PyMol software. Residues Lys119 in both homodimer chains were covalently bound to the PLP cofactors in the form of enzyme aldimines, and parameters for these modified residues were obtained using the *antechamber* module of Amber using the *gaff2* force field and with partial charges set to fit the electrostatic potential generated with HF/6-31G(d) using the RESP [47] method. The charges were calculated according to the Merz-Singh-Kollman scheme using Gaussian 16 (http://gaussian.com/). Heme groups were modelled using all-atom parameters developed by Giammona *et. al.* (D. A. Giammona, Ph.D. thesis, University of California, Davis (1984)). Molecular dynamics (MD) simulations for each homodimer were run with the Amber 18 suite (http://ambermd.org/), using the *ff14SB* [48] and *gaff2* [49] force fields. Initial structures were neutralized with either Na^+ or Cl^- ions and set at the center of a cubic TIP3P water [50] box with a buffering distance between solute and box of 10 Å. For each complex, we followed a two-stage geometry optimization approach: the first stage minimizes only the positions of solvent molecules and ions, and the second stage is an unrestrained minimization of all the atoms in the simulation cell. The systems were then heated by incrementing the temperature from 0 to 300 K under a constant pressure of 1 atm and periodic boundary conditions. Harmonic restraints of 10 kcal mol^{-1} were applied to the solute, under the Andersen temperature coupling scheme. The time step was kept at 1 fs during the heating stages, allowing potential inhomogeneities to self-adjust. Water molecules were treated with the SHAKE algorithm [51] such that the angle between the hydrogen atoms is kept fixed through the simulations. Long-range electrostatic effects were modelled using the particle mesh Ewald method [52]. A 8 Å cutoff was applied to Lennard-Jones interactions. Each system was equilibrated for 2 ns with a 2 fs time step at a constant volume and temperature of 300 K. Production was run as a 2,000 ns NVT trajectory at 300 K with a time step of 2 fs using the Andersen thermostat. The PCA were analyzed through Bio3d program [53]. Representative snapshots of the production trajectories were obtained using the *cpptraj* module of Amber and rendered with PyMol. Tunnels and channels in the protein structures were analyzed using Caver 3 software (http://www.caver.cz/) and visualized using PyMol.

Acknowledgements

This work was supported by Spanish Ministerio de Ciencia e Innovación (MICINN), Grants BFU2010-17857 and PID2019-109055RB-I00, Spanish Ministry of Economy and Competitiveness Grants BFU2013-47531-R and BFU2016-77408-R and BIOEF/EiTB MARATOIA BIO16/ER/035 to L. A. M.-C. IK4-TEKNIKER and CIC bioGUNE funded a PhD fellowship to CF-R, RTI2018-099592-B-C22 to GJO and PhD fellowship to RNF. We thank MINECO for the Severo Ochoa Excellence Accreditation (SEV-2016-0644) and a PhD fellowship (REF BES-2017-080435) awarded to IG-R. This research was also supported by departmental funds provided by the Italian Ministry of Research and Education (FUR2018, FUR2019) to A. A. and P. D., and in part by the Italian MIUR-PRIN 2017 grant No. 2017ZBBYNC to A. A. We thank the Centro Piattaforme Tecnologiche of the University of Verona for providing access to the CD spectropolarimeter.

Declaration of Competing Interest

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.05.052.

References

[1] Aguirre AA, Longcore T, Barbieri M, Dabritz H, Hill D, Klein PN, et al. The one health approach to toxoplasmosis: epidemiology, control, and prevention strategies. EcoHealth 2019.
[2] Wang ZD, Liu HH, Ma ZX, Ma HY, Li ZY, Bin YZ, et al. Toxoplasma gondii infection in immunocompromised patients: A systematic review and meta-analysis. Front Microbiol. 2017;8(MAR):1–12.
[3] Flegr J, Horáček J. Negative effects of latent toxoplasmosis on mental health. Front Psychiatry 2020;10.
[4] Charvat RA, Arrizabalaga G. Oxidative stress generated during monensin treatment contributes to altered Toxoplasma gondii mitochondrial function. Sci Rep [Internet]. 2016;6(September 2015):1–17. Available from: http://dx.doi.org/10.1038/srep22997.
[5] Strobl JS, Seibert CW, Li Y, Nagarkatti R, Mitchell SM, Rosypal AC, et al. Inhibition of toxoplasma gondii and plasmodium falciparum infections in vitro by NSC3852, a redox active antiproliferative and tumor cell differentiation agent. J Parasitol 2009;95(1):215–23.
[6] Bosch SS, Kronenberger T, Meissner KA, Zimbres FM, Stegehake D, Izui NM, et al. Oxidative stress control by apicomplexan parasites. Biomed Res Int 2015.
[7] Conter C, Fruncillo S, Fernández-rodríguez C, Martinez-cruz LA, Astegno A. Cystathionine beta synthase is involved in cysteine biosynthesis and H 2 S generation in Toxoplasma gondii. Sci Rep 2020;10(1).
[8] Maresi E, Janson G, Fruncillo S, Paiardini A, Vallone R, Dominici P, et al. Functional Characterization and Structure-Guided Mutational Analysis of the Transsulfuration Enzyme Cystathionine γ-Lyase from Toxoplasma gondii. Int J Mol Sci. 2018 Jul 20;19(7):2111. Available from: http://www.mdpi.com/1422-0067/19/7/2111.
[9] Majtan T, Pey AL, Fernández R, Fernández JA, Martínez-Cruz LA, Kraus JP. Domain Organization, Catalysis and Regulation of Eukaryotic Cystathionine Beta-Synthases. Uversky VN, editor. PLoS One [Internet]. 2014;9(8):e105290. Available from: https://dx.plos.org/10.1371/journal.pone.0105290.
[10] Devi S, Tarique KF, Ali MF, Abdul Rehman SA, Gourinath S. Identification and characterization of Helicobacter pylori O-acetylserine-dependent cystathionine β-synthase, a distinct member of the PLP-II family Available from. Mol Microbiol [Internet]. 2019;112(2):718–39. https://onlinelibrary.wiley.com/doi/abs/10.1111/mmi.14315.
[11] Devi S, Abdul Rehman SA, Tarique KF, Gourinath S. Structural characterization and functional analysis of cystathionine β-synthase: an enzyme involved in the reverse transsulfuration pathway of Bacillus anthracis Available from. FEBS J [Internet]. 2017;284(22):3862–80. https://onlinelibrary.wiley.com/doi/abs/10.1111/febs.14273.
[12] Matoba Y, Yoshida T, Izuhara-Kihara H, Noda M, Sugiyama M. Crystallographic and mutational analyses of cystathionine β-synthase in the H 2 S-synthetic gene cluster in Lactobacillus plantarum Available from. Protein Sci [Internet]. 2017;26(4):763–83. http://doi.wiley.com/10.1002/pro.3123.
[13] Vozdek R, Hnízda A, Krijt J, Kostrouchová M, Kožich V. Novel structural arrangement of nematode cystathionine β-synthases: characterization of Caenorhabditis elegans CBS-1. Biochem J. 2012;443(2):535–47.Available from: https://portlandpress.com/biochemj/article/443/2/535/46107/Novel-structural-arrangement-of-nematode.
[14] Jhee K-H, McPhie P, Miles EW. Yeast cystathionine β-synthase is a pyridoxal phosphate enzyme but, unlike the human enzyme, is not a heme protein Available from. J Biol Chem [Internet]. 2000;275(16):11541–4. http://www.jbc.org/lookup/doi/10.1074/jbc.C000056200.
[15] Bateman A. The structure of a domain common to archaebacteria and the homocystinuria disease protein Available from. Trends Biochem Sci [Internet]. 1997;22(1):12–3. https://linkinghub.elsevier.com/retrieve/pii/S0968000496300467.
[16] Baykov AA, Tuominen HK, Lahti R. The CBS domain: A protein module with an emerging prominent role in regulation. ACS Chem Biol 2011;6(11):1156–63.
[17] Tu Y, Kreinbring CA, Hill M, Liu C, Petsko GA, McCune CD, et al. Crystal Structures of Cystathionine β-Synthase from Saccharomyces cerevisiae : One Enzymatic Step at a Time Available from. Biochemistry [Internet]. 2018;57(22):3134–45. https://pubs.acs.org/doi/10.1021/acs.biochem.8b00092.
[18] Koutmos M, Kabil O, Smith JL, Banerjee R. Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine -synthase Available from. Proc Natl Acad Sci [Internet]. 2010;107(49):20958–63. http://www.pnas.org/cgi/doi/10.1073/pnas.1011448107.
[19] Giménez-Mascarell P, Majtan T, Oyenarte I, Ereño-Orbea J, Majtan J, Klaudiny J, et al. Crystal structure of cystathionine β-synthase from honeybee Apis mellifera Available from. J Struct Biol. 2018;202(1):82–93. https://linkinghub.elsevier.com/retrieve/pii/S1047847717302319.
[20] Taoka S, Lepore BW, Kabil Ö, Ojha S, Ringe D, Banerjee R. Human Cystathionine β-Synthase Is a Heme Sensor Protein. Evidence That the Redox Sensor Is Heme and Not the Vicinal Cysteines in the CXXC Motif Seen in the Crystal Structure of the Truncated Enzyme. Biochemistry [Internet]. 2002 Aug;41(33):10454–61. Available from: https://pubs.acs.org/doi/10.1021/bi026052d.
[21] Janošík M, Oliveriusová J, Janošíková B, Sokolová J, Kraus E, Kraus JP, et al. Impaired Heme Binding and Aggregation of Mutant Cystathionine β-Synthase Subunits in Homocystinuria. Am J Hum Genet [Internet].2001;68(6):1506-13. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929707610623.
[22] Weeks CL, Singh S, Madzelan P, Banerjee R, Spiro TG. Heme regulation of human cystathionine β-synthase activity: Insights from fluorescence and Raman spectroscopy. J Am Chem Soc 2009;131(35):12809–16.
[23] Meier M. Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein Available from. EMBO J [Internet]. 2001;20(15):3910–6. http://emboj.embopress.org/cgi/doi/10.1093/emboj/20.15.3910.
[24] Ereno-Orbea J, Majtan T, Oyenarte I, Kraus JP, Martinez-Cruz LA. Structural basis of regulation and oligomerization of human cystathionine -synthase, the central enzyme of transsulfuration. Proc Natl Acad Sci [Internet]. 2013 Oct 1;110(40):E3790–9. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1313683110.
[25] McCorvie TJ, Kopec J, Hyung S-J, Fitzpatrick F, Feng X, Termine D, et al. Inter-domain Communication of Human Cystathionine β-Synthase. J Biol Chem [Internet]. 2014 Dec 26;289(52):36018–30. Available from: http://www.jbc.org/lookup/doi/10.1074/jbc.M114.610782.
[26] Ereño-Orbea J, Majtan T, Oyenarte I, Kraus JP, Martínez-Cruz LA. Structural insight into the molecular mechanism of allosteric activation of human cystathionine β-synthase by S -adenosylmethionine. Proc Natl Acad Sci [Internet]. 2014 Sep 16;111(37):E3845–52. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1414545111.
[27] Christen P, Mehta PK. From cofactor to enzymes. The molecular evolution of pyridoxal-5?-phosphate-dependent enzymes. Chem Rec [Internet]. 2001;1(6):436–47. Available from: http://doi.wiley.com/10.1002/tcr.10005.
[28] Evande R, Blom H, Boers GHJ, Banerjee R. Alleviation of Intrasteric Inhibition by the Pathogenic Activation Domain Mutation, D444N, in Human Cystathionine β-Synthase Available from. Biochemistry [Internet]. 2002;41(39):11832–7. https://pubs.acs.org/doi/10.1021/bi026248d.
[29] Liebschner D, Afonine PV, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, et al. Polder maps: Improving OMIT maps by excluding bulk solvent. Acta Crystallogr Sect D Struct Biol. 2017;73(2):148–57.
[30] Shatalin K, Shatalina E, Mironov A, Nudler E. H2S: A universal defense against antibiotics in bacteria. Science 2011;334(6058):986–90.
[31] Pal VK, Bandyopadhyay P, Singh A. Hydrogen sulfide in physiology and pathogenesis of bacteria and viruses. IUBMB Life 2018;70:393–410.
[32] Saini V, Chinta KC, Reddy VP, Glasgow JN, Stein A, Lamprecht DA, et al. Hydrogen sulfide stimulates Mycobacterium tuberculosis respiration, growth and pathogenesis. Nat Commun [Internet]. 2020;11(1):1–17. Available from: http://dx.doi.org/10.1038/s41467-019-14132-y.
[33] Astegno A, Capitani G, Dominici P. Functional roles of the hexamer organization of plant glutamate decarboxylase Available from. Biochim Biophys Acta - Proteins Proteomics [Internet]. 2015;1854(9):1229–37. https://linkinghub.elsevier.com/retrieve/pii/S1570963915000126.
[34] Astegno A, Giorgetti A, Allegrini A, Cellini B, Dominici P. Characterization of C-S lyase from C. diphtheriae: A possible target for new antimicrobial drugs. Biomed Res Int 2013;701536.
[35] Astegno A, Allegrini A, Piccoli S, Giorgetti A, Dominici P. Role of active-site residues Tyr55 and Tyr114 in catalysis and substrate specificity of Corynebacterium diphtheriae C-S lyase. Proteins Struct Funct Bioinforma 2015;83(1):78–90.
[36] Oyenarte I, Majtan T, Ereño J, Corral-Rodríguez MA, Kraus JP, Martínez-Cruz LA. Purification, crystallization and preliminary crystallographic analysis of

human cystathionine β-synthase Available from. Acta Crystallogr Sect F Struct Biol Cryst Commun [Internet]. 2012;68(11):1318–22. http://www.ncbi.nlm.nih.gov/pubmed/23143240.
[37] Kabsch WXDS. Acta Crystallogr Sect D Biol Crystallogr 2010;66:125–32.
[38] Evans P. Scaling and assessment of data quality. Acta Crystallogr Sect D Biol Crystallogr 2006;62(1):72–82.
[39] Evans PR, Murshudov GN. How good are my data and what is the resolution?. Acta Crystallogr Sect D Biol Crystallogr 2013;69(7):1204–14.
[40] Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D Biol Crystallogr 2011;67(4):235–42.
[41] Tickle I. STARANISO : use of a WebGL-based 3D interactive graphical display to represent and visualise data quality metrics for anisotropic macromolecular diffraction data. Acta Crystallogr Sect A Found Adv 2019;75(a2):e162.
[42] Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, et al. Data processing and analysis with the autoPROC toolbox Available from. Acta Crystallogr Sect D Biol Crystallogr [Internet]. 2011;67(4):293–302. http://www.ncbi.nlm.nih.gov/pubmed/4954368.
[43] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40(4):658–74.
[44] Adams PD, Afonine PV, Bunkóczi G, Chen VB, Echols N, Headd JJ, et al. The Phenix software for automated determination of macromolecular structures Available from. Methods [Internet]. 2011;55(1):94–106. https://linkinghub.elsevier.com/retrieve/pii/S1046202311001319.
[45] Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr Sect D Biol Crystallogr 2012;68(4):352–67.
[46] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot Available from. Acta Crystallogr Sect D Biol Crystallogr [Internet]. 2010;66(4):486–501. http://scripts.iucr.org/cgi-bin/paper?S0907444910007493.
[47] Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model Available from. J Phys Chem [Internet]. 1993;97(40):10269–80. https://pubs.acs.org/doi/abs/10.1021/j100142a004.
[48] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB Available from. J Chem Theory Comput [Internet]. 2015;11(8):3696–713. http://www.ncbi.nlm.nih.gov/pubmed/26574453.
[49] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field Available from. J Comput Chem [Internet]. 2004;25(9):1157–74. http://doi.wiley.com/10.1002/jcc.20035.
[50] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water Available from. J Chem Phys [Internet]. 1983;79(2):926–35. http://aip.scitation.org/doi/10.1063/1.445869.
[51] Miyamoto S, Kollman PA. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models Available from. J Comput Chem [Internet]. 1992;13(8):952–62. http://doi.wiley.com/10.1002/jcc.540130805.
[52] Darden T, York D, Pedersen L. Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems Available from. J Chem Phys [Internet]. 1993;98(12):10089–92. http://aip.scitation.org/doi/10.1063/1.464397.
[53] Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: An R package for the comparative analysis of protein structures. Bioinformatics 2006;22(21):2695–6.