Supporting Information

Ereño-Orbea et al. 10.1073/pnas.1313683110

SI Materials and Methods

Constructs and Purification of the Recombinant Proteins. The preparation of a pET28-C-hCBSOPTΔ516–525 expression construct is described in detail elsewhere (1). We used the same mutagenesis oligonucleotides to delete the residues 516–525 in our pGEX-6P1-hCBSOPT construct (2). The pET28-C-hCBSOPTΔ1–39Δ516–525 was prepared by cloning the desired sequence into pET28a vector (Novagen) in a way similar to that described previously (3). The insert was amplified by proofreading PCR using the pET28-C-hCBSOPTΔ516–525 as a DNA template and the following primers: forward, 5′-ctagCCATGGaaccgctgtggattcgtccgg, and reverse, 5′-ctagCTCGAGtttctggtcacgttcttgtgcggcc (capital letters designate NcoI and XhoI sites introduced by forward and reverse primers, respectively). Both the PCR product and vector were cleaved overnight with NcoI and XhoI (New England Biolabs), separated in 1% agarose gel, cut out, and cleaned up. Ligation was performed using the QuickLigation Kit (New England Biolabs) according to the manufacturer's recommendations. The D444N pathogenic mutation was introduced by using a QuikChange II XL mutagenesis kit (Agilent) according to the manufacturer's recommendations. The D444N mutagenesis oligonucleotides were forward,5′-aatcctgcgtgaaaaaggttttAATcaggccccgg, and reverse, 5′-ccggggcctgATTaaaaccttttcacgcaggatt (capital letters designate the codon of the mutated residue). Purification of recombinant proteins followed the protocols that we developed for various human cystathionine β-synthase (hCBS) constructs carrying either cleavable GST at the N terminus (2) or a permanent 6xHis tag at the C terminus (3) with a few modifications (1).

Biochemical Characterization of the Purified Enzymes. Protein concentrations were determined by the Bradford method (Thermo Pierce) using BSA as a standard according to the manufacturer's recommendations. Denatured proteins were separated by SDS/PAGE using 10% polyacrylamide precast gels (Mini-PROTEAN TGX; Bio-Rad); the native samples were separated in 4–15% polyacrylamide gradient precast gels (Mini-PROTEAN TGX; Bio-Rad). For visualization, the denatured gels were stained with Simple Blue (Invitrogen). Western blot analysis of crude cell lysates under denaturing or native conditions was performed as described previously (2). The CBS activity in the classical reaction was determined by a previously described radioisotope assay using [^{14}C(U)] L-serine as the labeled substrate, essentially as described elsewhere (2, 3).

Crystallization and Data Collection. The crystal structure of the full-length hCBS was determined using two different protein constructs, referred to as C-hCBSΔ516–525 and hCBSOPTΔ516–525, that differ mainly by the presence or absence of a 6xHis-tag at the C terminus. Both constructs encompass the full-length hCBS protein and lack the residues 516–525 from the long loop connecting the strands β15 and β16 that favored the crystallization. The crystals of C-hCBSOPTΔ516–525 and hCBSOPTΔ516–525 were grown by the sitting-drop vapor-diffusion method at 293 K in 96-well crystallization plates according to the protocol described previously (1). Drops consisted of 100 nL of protein solution mixed with 200 nL of the precipitant solution [30% (vol/vol) polyethylene glycol monomethyl ether 550, 0.1 M sodium citrate tribasic dehydrate (pH 5.0)] and 100 nL of the seed-stock; the protein concentration was 21 mg/mL. Single crystals were transferred to a cryoprotection solution (paraffin oil) and flash frozen in liquid nitrogen.

To determine the structure of the human enzyme bearing the D444N mutation, we used two different protein constructs, referred to as C-hCBSOPTΔ516–525 D444N and C-hCBSOPTΔ1–39Δ516–525 D444N, respectively. The crystals were obtained by the hanging-drop vapor-diffusion method at 293 K in 24-well crystallization plates. Drops consisted of 200 nL of protein solution mixed with 400 nL precipitant solution [0.085 M Hepes-Na (pH 7.5), 8.5% (vol/vol) iso-propanol, 17% (wt/vol) polyethylene glycol 4000, 15% (vol/vol) glycerol anhydrous, 0.2 M ammonium citrate tribasic (pH 7.0), 0.1 M imidazole (pH 7.0), 20% (wt/vol) polyethylene glycol monomethyl ether 2000] and 200 nL of the seed-stock; the protein concentration was 25 mg/mL for both samples. The C-hCBSOPTΔ516–525 D444N crystals were directly flash frozen in liquid nitrogen. Single crystals of C-hCBSOPTΔ1–39Δ516–525 D444N were transferred to a cryoprotection solution [crystallization buffer with 30% (wt/vol) polyethylene glycol 400] and flash frozen in liquid nitrogen. All datasets used in this work were collected at European Synchrotron Radiation Facility, Grenoble beamlines ID23-1 (C-hCBSOPTΔ516–525, λ = 0.9793 Å; hCBSOPTΔ516–525, λ = 0.9725 Å; C-hCBSOPTΔ516–525 D444N, λ = 0.9763 Å) and ID29 (C-hCBSOPTΔ1–39Δ516–525 D444N, λ = 0.9790 Å) and were processed using HKL2000 (4) or XDS (5) software.

X-Ray Diffraction Data Collection, Phasing, and Refinement. The hCBS structure was determined by molecular replacement with the PHENIX program (6) using the crystal structure of the truncated 45-kDa hCBS [Protein Data Bank (PDB) code 1JBQ] as the initial search model. After several cycles of refinement using PHENIX and REFMAC5 (7), CBS domains were built manually using Coot (8). Crystallographic refinement was carried out with REFMAC5 with TLS parameters. The three other crystal structures were determined by molecular replacement with the PHENIX program using the crystal structure of C-hCBSOPTΔ516–525 as the initial search model. For refinement, PHENIX and REFMAC5 were used. The following are the Ramachandran statistics for the refined coordinates [residues in favored region (%, number of outliers)]: C-hCBSOPTΔ516–525 (96.95, 0), hCBSOPTΔ516–525 (97.71, 0), C-hCBSOPTΔ516–525 D444N (98.28, 0.1), and C-hCBSOPTΔ1–39Δ516–525 D444N (98.18, 0). The final refinement statistics are summarized in Table S1. Atomic coordinates and structure factors have been deposited in the PDB under accession codes 4L0D, 4L3V, 4L27 and 4L28.

1. Oyenarte I, et al. (2012) Purification, crystallization and preliminary crystallographic analysis of human cystathionine β-synthase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 68(Pt 11):1318–1322.
2. Majtan T, Liu L, Carpenter JF, Kraus JP (2010) Rescue of cystathionine beta-synthase (CBS) mutants with chemical chaperones: Purification and characterization of eight CBS mutant enzymes. *J Biol Chem* 285(21):15866–15873.
3. Majtan T, Kraus JP (2012) Folding and activity of mutant cystathionine β-synthase depends on the position and nature of the purification tag: Characterization of the R266K CBS mutant. *Protein Expr Purif* 82(2):317–324.
4. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
5. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125–132.
6. Adams PD, et al. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* 55(1):94–106.
7. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–367.
8. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.

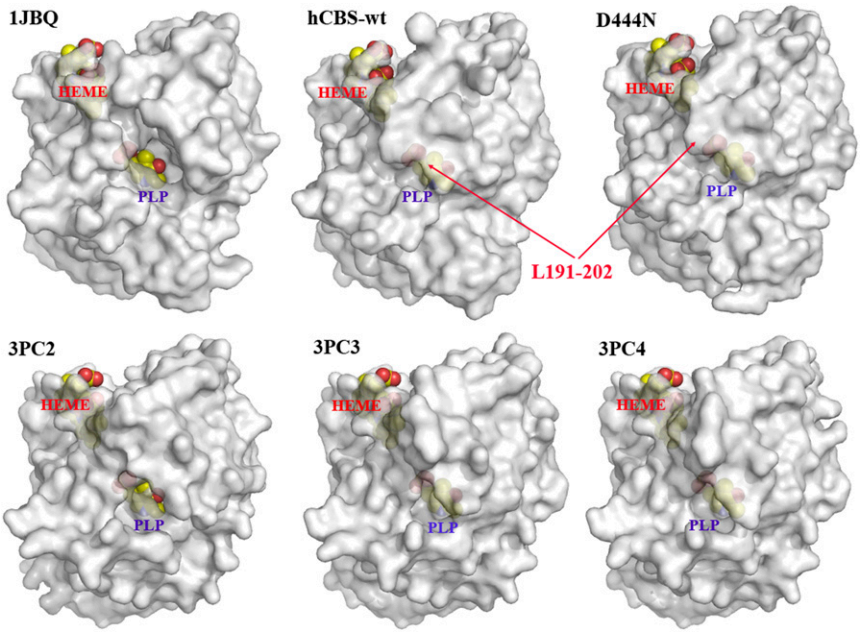


Fig. S1. Catalytic cores of different CBS enzymes. Surface representation of the catalytic cores of the truncated 45-kDa hCBS (PDB code 1JBQ), the full-length hCBS in its basal state (hCBS-wt), the D444N hCBS mutant, and *Drosophila melanogaster* CBS (dCBS) without substrates (PDB ID code 3PC2) and with bound substrates (PDB ID codes 3PC3 and 3PC4). Note that the pyridoxal-5′-phosphate (PLP) cavity is accessible (open) only in the truncated 45-kDa hCBS, which lacks the regulatory domain, and in the constitutively activated dCBS without bound substrates.

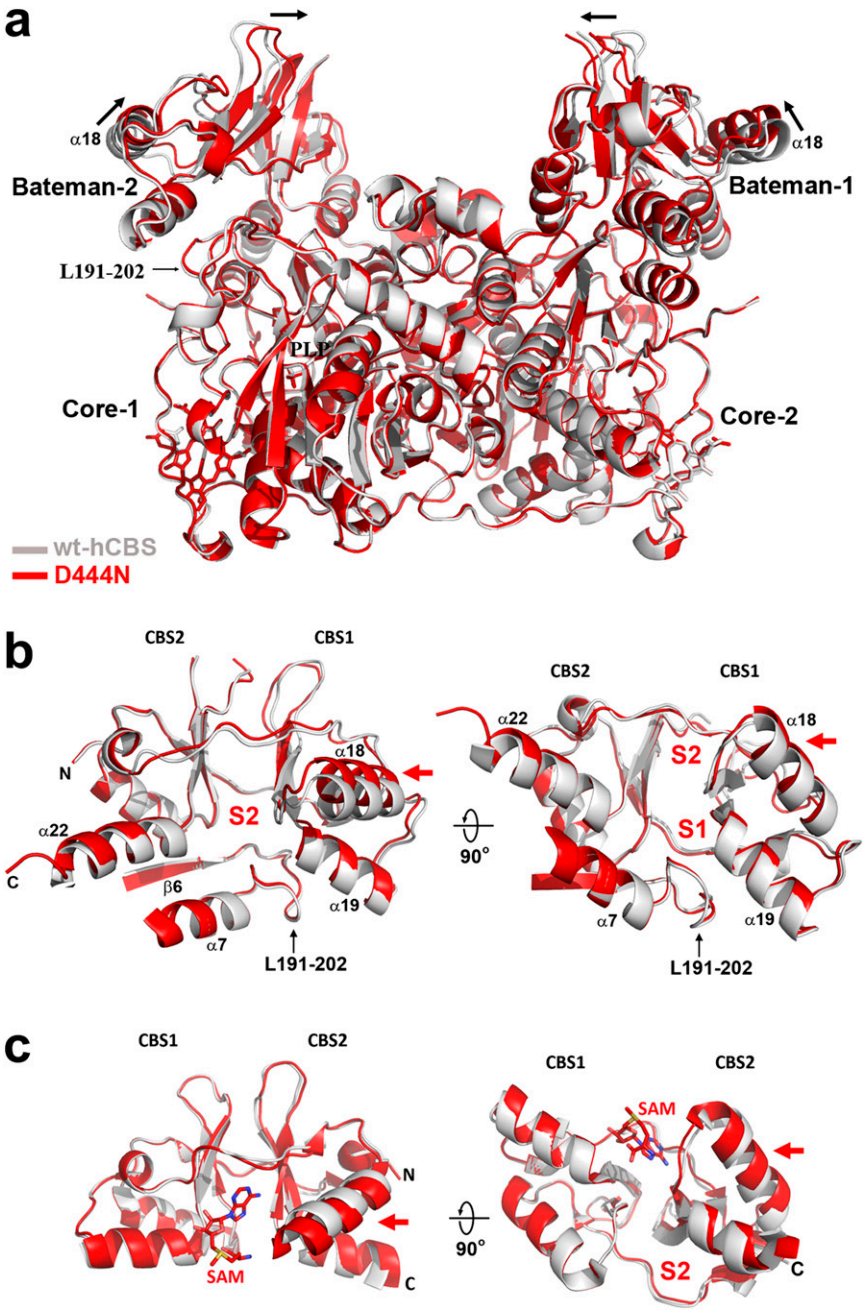


Fig. S2. (*A* and *B*) Crystal structure of wild-type hCBS and the pathogenic D444N mutant. (*A*) The main difference between the native hCBS (gray) and the D444N mutant (red) dimers is a slight displacement of the Bateman modules, accompanied by a shift of helix α18 (designated by arrows). (*B*) Structural superimposition of the Bateman modules of both proteins in two different views. (*C*) Effect of S-adenosyl-L-methionine (AdoMet) binding at the site S1 of the Bateman module of the archaeal MJ0100 (1). The apo- and holo-forms of MJ0100 are represented in gray and red, respectively. Note that in hCBS the D444N mutation induces an effect similar to that induced by the addition of AdoMet to MJ0100.

1. Lucas M, et al. (2010) Binding of S-methyl-5′-thioadenosine and S-adenosyl-L-methionine to protein MJ0100 triggers an open-to-closed conformational change in its CBS motif pair. *J Mol Biol* 396(3):800–820.

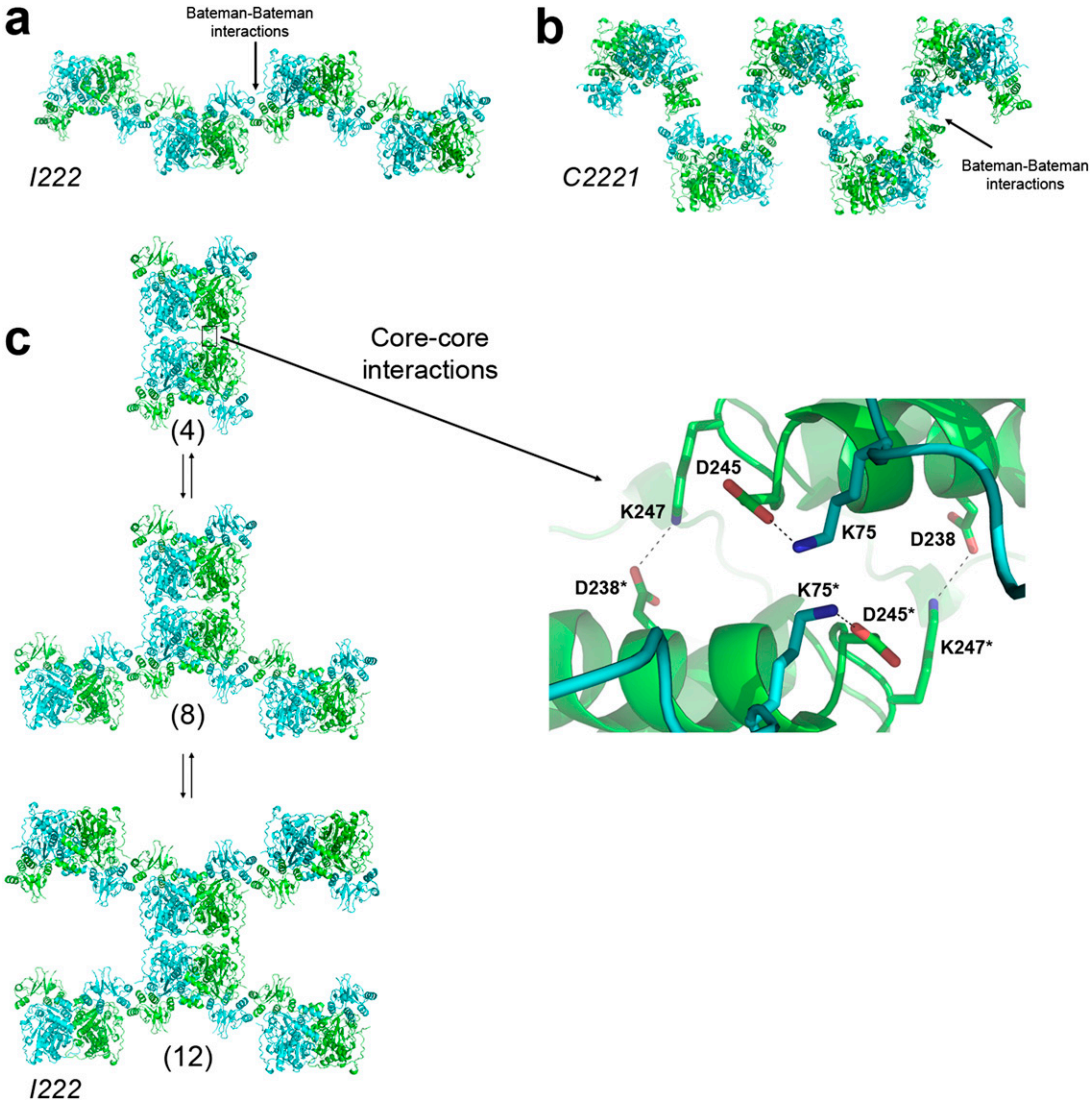
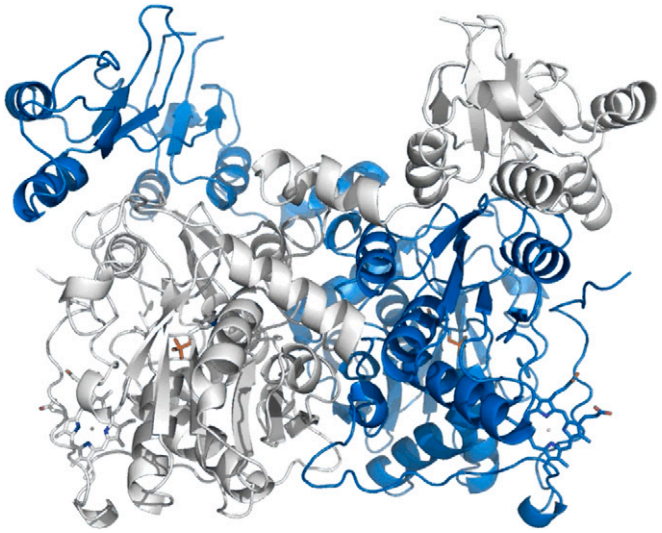


Fig. S3. Different mechanisms of hCBS aggregation. The image reflects packing contacts found in the crystals that may cause different degrees of protein aggregation in solution. The protein dimer forms polymeric chains that interact through their Bateman modules following two distinct patterns. In the C-hCBSOPTΔ516–525 (space group I222) crystals, the hCBS dimers interact through the α-helices of their CBS1 motifs (*A*), as well as via salt bridges between the residues K75, D238, D245, and K247 of their core domains (*C*). On the other hand, the C-hCBSOPTΔ516–525 (space group C2221) crystals and also the C-hCBSOPTΔ516–525 D444N (space group $P2_12_12_1$) crystals contain polymeric chains of hCBS dimers interacting through their Bateman modules (*B*). In this case, the protein contacts include both hydrophobic and polar interactions that affect the flexible loops connecting the central β-strands of the Bateman modules as well as the long loops preceding the first helix of CBS2 motifs.

Table S1. Data collection and refinement

Parameter	C-hCBSOPTΔ516–525	hCBSOPTΔ516–525	C-hCBSOPTΔ516–525 D444N	C-hCBSOPTΔ1–39Δ516–525 D444N
PDB code	4L0D	4L3V	4L28	4L27
Data collection				
Space group	I222	$C222_1$	$P2_12_12_1$	$P6_5$
Cell dimensions				
a, b, c, Å	124.365, 136.203, 169.307	227.75, 342.655, 107.253	109.498, 131.098, 207.064	190.454, 190.454, 140.847
α, β,γ, °	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 120
Resolution, Å	50–3.0 (3.1–3.0)	48–3.6 (3.8–3.6)	48–2.6 (3.5–3.4)	57–3.4 (3.5–3.4)
R_{meas}	0.09 (0.45)	0.31 (1.51)	0.36 (0.93)	0.112 (1.30)
I/ σ I	11.6 (2.2)	7.25 (2.28)	7.62 (2.26)	15.71 (1.99)
Completeness, %	91.6 (90.0)	99.55 (72.47)	99.24 (95.92)	99.72 (97.25)
Redundancy	3.7 (3.6)	2.44 (2.37)	14.3 (14.5)	8.61 (8.27)
$CC_{1/2}$, %	99.6 (66.8)	97.8 (51.2)	109.498, 131.098, 207.064	99.9 (61.8)
Refinement				
Resolution, Å	50–3.0	48–3.6	48–3.4	57–3.4
No. reflections	54,802	46,253	41,267	40,186
R_{work}/R_{free}	0.2391/0.2806	0.2159/0.2302	0.2114/0.2283	0.2037/0.2405
No. atoms				
Protein	7,617	11,546	15,271	15,147
Ligand (PLP/HEME)	30/86	45/129	60/172	60/172
B-factors				
Protein	99.8	55.2	84.0	56.9
Ligand (PLP/HEME)	75.2/110.0	38.4/45.4	67.9/81.7	35.1/62.3
rms deviations				
Bond lengths, Å	0.007	0.007	0.008	0.015
Bond angles, °	1.083	1.158	1.090	1.124

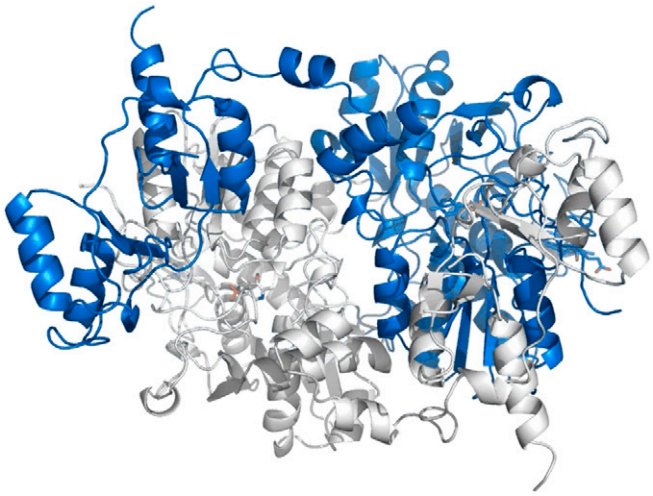
One crystal was used for each data set. Values in parentheses are for highest-resolution shell.



Movie S1. The unlikely hCBS-to-dCBS activation mechanism (view 1): an impossible transition from the basal form of hCBS to the constitutively activated conformation reported for dCBS (1). As shown, this mechanism of activation is not possible because it would involve severe clashes between the Bateman modules of the complementary subunits.

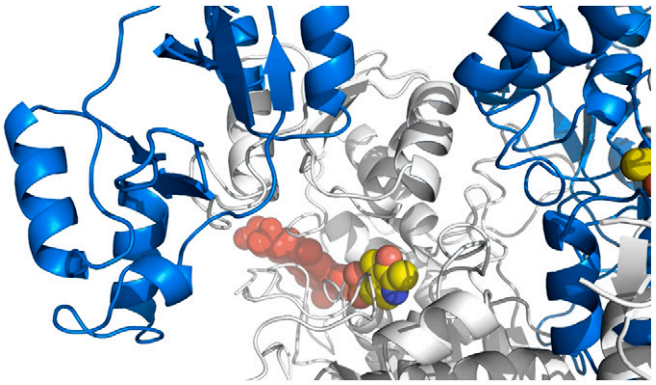
1. Koutmos M, Kabil O, Smith JL, Banerjee R (2010) Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine beta-synthase. *Proc Natl Acad Sci USA* 107(49):20958–20963.

Movie S1



Movie S2. The unlikely hCBS-to-dCBS activation mechanism (view 2): a view from the top (at 90°) of the structures shown in Movie S1.

Movie S2



Movie S3. Proposed activation of hCBS by AdoMet (ribbon representation): the modeled mechanism of hCBS activation upon AdoMet binding to the Bateman module (in blue). The heme group is represented by red spheres. PLP is represented by multicolored spheres. Binding of AdoMet to the cavity S2 of the Bateman module triggers a conformational change in this domain, similar to that reported for the AdoMet-binding archaeal protein MJ0100 (1). This structural change favors the displacement of three loops (L145–148, L171–174, and L191–202) located at the entrance of the PLP cavity from a closed to an open conformation and allows unrestricted access of substrates into the PLP cavity.

1. Lucas M, et al. (2010) Binding of S-methyl-5′-thioadenosine and S-adenosyl-L-methionine to protein MJ0100 triggers an open-to-closed conformational change in its CBS motif pair. *J Mol Biol* 396(3):800–820.

Movie S3