



Cystathionine beta-synthase mutations: effect of mutation topology on folding and activity

Journal:	<i>Human Mutation</i>
Manuscript ID:	humu-2009-0545.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	09-Apr-2010
Complete List of Authors:	<p>Kožich, Viktor; First Faculty of Medicine-Charles University in Prague and General University Hospital in Prague, Institute of Inherited Metabolic Disorders</p> <p>Sokolova, Jitka; First Faculty of Medicine-Charles University in Prague and General University Hospital in Prague, Institute of Inherited Metabolic Disorders</p> <p>Klatovska, Veronika; First Faculty of Medicine-Charles University in Prague and General University Hospital in Prague, Institute of Inherited Metabolic Disorders</p> <p>Krijt, Jakub; First Faculty of Medicine-Charles University in Prague and General University Hospital in Prague, Institute of Inherited Metabolic Disorders</p> <p>Janošík, Miroslav; First Faculty of Medicine-Charles University in Prague and General University Hospital in Prague, Institute of Inherited Metabolic Disorders</p> <p>Jelínek, Karel; Faculty of Science-Charles University in Prague, Department of Physical and Macromolecular Chemistry</p> <p>Kraus, Jan; University of Colorado School of Medicine, Department of Pediatrics</p>
Key Words:	Homocysteine, cystathionine beta-synthase deficiency, homocystinuria, folding, misfolding, topology, mutations, activity, E.coli, expression



Cystathionine beta-synthase mutations: effect of mutation topology on folding and activity

Viktor Kožich, Jitka Sokolová, Veronika Klatovská, Jakub Krijt, Miroslav Janošík, Karel Jelínek and Jan P.Kraus

Abstract

Misfolding of mutant enzymes may play an important role in the pathogenesis of cystathionine beta-synthase (CBS) deficiency. We examined properties of a series of 27 mutant variants, which together represent 70% of known alleles observed in patients with homocystinuria due to CBS deficiency. The median amount of SDS-soluble mutant CBS polypeptides in the pellet after centrifugation of bacterial extracts was increased by 50% compared to the wild type. Moreover, mutants formed on average only 12% of tetramers and their median activity reached only 3% of the wild type enzyme. In contrast to the wild type CBS about half of mutants were not activated by S-adenosylmethionine. Expression at 18°C substantially increased the activity of five mutants in parallel with increasing the amounts of tetramers. We further analyzed the role of solvent accessibility of mutants as a determinant of their folding and activity. Buried mutations formed on average less tetramers and exhibited 23-times lower activity than the solvent exposed mutations. In summary, our results show that topology of mutations predicts in part the behavior of mutant CBS and that misfolding may be an important and frequent pathogenic mechanism in CBS deficiency.

Abbreviations

- AdoHcy, S-Adenosylhomocysteine
- AdoMet, S-Adenosylmethionine
- CBS, cystathionine beta-synthase
- PAGE, polyacrylamide gel electrophoresis
- PLP, pyridoxal 5'-phosphate
- SDS, sodium dodecylsulfate

Key words

Homocysteine, cystathionine beta-synthase deficiency, homocystinuria, folding, misfolding, topology, mutations, activity, *E.coli*, expression

Runing title

Topology of CBS mutants and misfolding

Introduction

Cystathionine beta-synthase (CBS) deficiency (OMIM 236200, also known as homocystinuria due to CBS deficiency) is a well known multisystemic inborn error of metabolism with widely varying estimates of frequency. Although the numbers of patients with clinically ascertained CBS deficiency is rather low and the prevalence is only 1:330,000 worldwide (Mudd, et al., 2001) the molecular epidemiological studies suggest an incidence of around 1:10,000 in several European populations (Gaustadnes, et al., 1999; Janosik, et al., 2009; Refsum, et al., 2004). Classical forms of this disease manifest by combination of neurological, connective tissue and vascular involvement in childhood while milder forms of the disease may be diagnosed later in adulthood with only vascular manifestation. Severity of the untreated disease is determined mostly by the underlying genetic defect and patients with milder forms of CBS deficiency usually respond to pharmacological doses of pyridoxine by a marked decrease of their total homocysteine plasma levels.

More than 150 mutant alleles have been described in patients with CBS deficiency with a representation of all types of genetic variants (<http://www.uchsc.edu/cbs/cbsdata/cbsmain.htm>). Missense mutations are the most common variants described in the *CBS* gene representing 87% of the 753 analyzed patient-derived CBS alleles as of April 2010. In a previous study we have proposed that missense CBS mutant enzymes may misfold and that this behavior may be responsible for their pathogenicity (Janosik, et al., 2001b). This observation was later confirmed by other groups and also by

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

expressing mutants in systems other than *E.coli* (for references see Table 1). The finding that misfolding and decreased activities of some mutant proteins may be corrected with the help of chemical chaperones further supports the role of misfolding in the pathogenesis of CBS deficiency (Singh, et al., 2007). Despite many years of research novel data on structural and biochemical properties of CBS mutant proteins are scattered in the literature and a comprehensive analysis of the folding-related findings is missing.

In this study, we examined systematically catalytic and folding-related structural properties of a panel of 27 mutants located in different functional domains of the CBS molecule. The panel of mutations contained among others the most prevalent mutant CBS alleles to represent well the molecular mechanisms relevant for the majority of patients suffering from CBS deficiency. Since purification of CBS mutant proteins for further studies is not universally achievable we decided to study properties of mutant proteins in crude *E.coli* extracts obviating thus any possible loss of mutants during the purification procedure. We studied the solubility of the mutant proteins in SDS, their oligomeric structure, activity, response to S-adenosylmethionine (AdoMet) and S-adenosylhomocysteine (AdoHcy) and changes of these parameters after expression at a decreased temperature. We subsequently examined whether the location of mutant residues in the three dimensional CBS structure in respect to solvent accessibility is a determinant of their properties. We show that mutations of residues buried in protein globule appear to have more severe effect on folding than those exposed to the solvent.

Methods

Nucleotide and codon numbering

Nucleotide numbering reflects the cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference GenBank sequence NM_000071.2, according to the Journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon number 1.

Molecular dynamics simulation

In silico analyses presented in this paper can be divided in two parts. The first part presents an analysis of mutant structures of catalytic domain of CBS and the second part presents analysis

of CBS regulatory domain. X-ray structure of the active core of the CBS protein (PDB code 1JBQ;(Meier, et al., 2001)) was utilized as the starting model for the catalytic domain. 1JBQ X-ray structure contains the first 413 residues of the protein and covers the whole catalytic domain. Firstly we optimized the native structure in program Modeller (<http://salilab.org/modeller>; (Sali and Blundell, 1993)) which combines empirical spatial restraints and CHARMM force field into an objective function and employs conjugate gradients and molecular dynamics with simulated annealing methods for energy optimization. Model refinement level property of the Modeller procedure was set to very low. In the next step the optimized native structure was mutated *in silico* and mutant structures were optimized in the same way.

Although a model of the regulatory domain of human CBS (residues 415-543) has been published a few years ago (Sen, et al., 2005) the X-ray structure of this part of the CBS molecule is as yet unknown. We generated our own model of this portion of the wild type CBS using the comparative structure modeling. As the initial model we used X-ray structure of the conserved CBS domain of the protein TA0289 from *Thermoplasma acidophilum* (PDB code 1PVM; (Proudfoot, et al., 2008)). Sequence of the CBS regulatory domain was aligned with the TA0289 sequence and coordinates of its native structure were derived from the 1PVM coordinates (data not shown). The native structure was then optimized and mutant structures were computed as in case of the active core. For the native as well as all computed mutated structures of both domains, the surface accessible area of amino acid residues was calculated and normalized by maximum accessible surface area for the corresponding residue type by the program Modeller (Sali and Overington, 1994). A residue was considered exposed if its accessible surface area was larger than 40 \AA^2 and/or if its relative accessible surface was larger than 9%, while considered buried otherwise (Mirkovic, et al., 2004).

Expression in *E.coli*

The wild type CBS and mutant constructs were derived from the pHCS3 expression plasmid (Kozich and Kraus, 1992). Seven mutant plasmids, namely the c.341C>T (p.A114V), c.442G>A (p.G148R), c.526G>A (p.E176K), c.1224-2A>C (r.1224_1358del; p.W409_G453del), c.1265C>T (p. P422L), c.1304T>C (p.I435T) and c.1397C>T (p.S466L) were prepared using the previously described procedure of replacing a fragment of wild-type CBS expression plasmid with the equivalent restriction fragment derived from the mutant patient-derived cDNA (Kozich and Kraus, 1992). The remaining mutations were introduced

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

into the wild-type expression plasmid by help of the GeneTailor Site-directed mutagenesis kit (Invitrogen, Carlsbad, California) according to the manufacturer’s procedure. The sequences of mutagenic primers are given in Supp. Table S1. The authenticity of all constructs was verified by dideoxy sequencing using ALF sequencer (Amersham Pharmacia Biotech, Piscataway, New Jersey).

The expression plasmids were used to transform *E.coli* DH5α cells (Gibco BRL, Carlsbad, California). The cells were grown at 37°C or 18°C in SOB (Super Optimal Broth) media with 100 µg/ml ampicilin to OD₆₀₀ of ~0.5, expression of CBS was carried out in the presence of isopropyl β-D-1-thiogalactopyranoside (final concentration 0.33 mmol/l) at 37°C or 18°C for 3 hours or overnight, respectively. Each mutant protein was expressed in at least five independent experiments on separate days (i.e. 3 and 2 expression at 37°C and 18°C, respectively). Each series of mutant proteins contained also bacteria transformed with the pKK388.1 vector lacking the CBS insert as a negative control, and the wild type CBS that was used to normalize data.

Lysates from *E.coli* cell pellets were prepared by sonication in lysis buffer (50 mmol/l Tris-Cl pH 8, 10 mmol/l EDTA, 1 mmol/l DTT, 0.1% lysozyme) with the addition of commercial Protease Inhibitor Cocktail (Sigma Aldrich, St. Luis, Missouri) according to manufacturer’s protocol. A 50 µL aliquot of each sonicate was saved for the analysis of CBS in the particulate and non-particulate fraction, the remaining sonicate was centrifuged at 15,000 g for 30 minutes at 4° C, supernatant was aliquoted and stored for subsequent analyses at -85°C. Protein concentration in the supernatant was determined by the method of Lowry with bovine serum albumin as a standard (Lowry, et al., 1951).

Optimizing extraction of aggregated CBS

We tested several methods that have been previously used by other authors to extract aggregated proteins- such as amyloid- from bacterial lysates (Carra, et al., 2005). Optimization was carried out with the wild type CBS and four mutant enzymes (p.A114V, p.V180A, p.K102N, p.I278T). We have compared the effects of the following denaturing reagents: a) 1.5% SDS in the presence of 0.1M beta-mercaptoethanol (without preincubation), b) 3% SDS in the presence of 0.1M beta-mercaptoethanol (without preincubation), c) 6M urea, 1.5% SDS and 0.1M beta-mercaptoethanol (1 hour incubation at room temperature) d) 6M urea/thiourea (v/v 2:1) containing 1.5% SDS and 0.1M beta-mercaptoethanol (1 hour incubation at room temperature) and e) 100% formic acid (30 minutes incubation at 37 °C)

1
2
3 followed by a lyophilization and dissolution in denaturing solution containing 1.5% SDS
4 and 0.1M beta-mercaptoethanol. After incubation of uncentrifuged sonicates with different
5 denaturing reagents for the indicated time the samples were boiled for 10 minutes at 100 °C
6 and analyzed by SDS-PAGE and western blotting as described below. These experiments
7 revealed, that methods b) and e) provided a higher yield of immuno-reactive CBS signal
8 compared to the other methods. Owing to a simplicity and a sufficient efficacy of method b)
9 we subsequently used the extraction of insoluble CBS from pellets and supernatants by 3%
10 SDS.
11
12
13
14
15
16
17

18 19 **Analysis of particulate and non-particulate fraction**

20 50 µL of sonicated cell extracts were centrifuged at 4°C at 15,000g for 30 minutes, the
21 supernatants were saved and the pellets were resuspended in 50µL of denaturation buffer (50
22 mmol/l Tris-Cl pH8, 0.1 mol/l beta-mercaptoethanol, 3% SDS). To solubilize the CBS present
23 in the samples 5 µL of either the resuspended particulate fraction or of the supernatant were
24 mixed with 15 µL of the denaturation buffer and boiled at 100°C for 10 minutes, subsequently
25 4 µL of electrophoresis loading buffer supplemented with 0.1 mmol/l dithiothreitol were
26 added and the solution was heated for 10 minutes at 100° prior to loading on the SDS-PAGE
27 gel. Samples of the supernatant and particulate fraction of each mutant protein were loaded
28 next to each other on the gel and subjected to electrophoresis and western blot analysis (see
29 below the section Electrophoresis and Western Blot Analysis), each gel contained also the
30 supernatant and particulate fraction of a negative control and of bacteria expressing wild type
31 CBS. The sharply demarcated signal of CBS in each fraction was quantified by
32 chemiluminescence after subtracting the background of the negative control. For each mutant
33 the signal in the supernatant and the particulate fractions were added yielding thus the total
34 CBS antigen of the respective mutant, this total signal was subsequently normalized to the
35 total signal of wild type CBS enzyme after correction for the protein concentration.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 **Electrophoresis and Western Blot Analysis**

52 Separation of CBS under denaturing conditions was carried out in 4-15% gradient native
53 polyacrylamide gels (Criterion Precast Gel, BioRad) with standard Laemmli SDS-containing
54 buffer system (Laemmli, 1970). For the analysis of CBS under native conditions cell lysates
55 containing 10 µg of total protein were separated using the same precast polyacrylamide gels
56 with standard Laemmli buffer system without SDS (Laemmli, 1970).
57
58
59
60

Proteins separated by either SDS-PAGE or native PAGE were transferred onto polyvinylidene difluoride membrane (Immobilon-P, Millipore, Billerica, Massachusetts) using semi-dry blotting transfer technique. After transfer nonspecific binding sites were blocked by overnight incubation with 5% non-fat dry milk in 1 x phosphate-buffered saline solution (PBS), pH 7.0 and 0.2% Tween 20. The membrane was subsequently incubated with immunopurified anti-CBS antibody diluted 1:5000 in 1 x PBS containing 3% bovine serum albumin for 1 hour and after a series of washes with secondary anti-rabbit IgG antibody conjugates with HRP (Pierce, Rockford, Illinois) for 30 minutes. The secondary antibody was diluted 1:30 000 in 1 x PBS/ 0.2% Tween 20 containing 5% non-fat dry milk. After a second series of washes the signal was visualised using the West Pico Super Signal system (Pierce Biotechnology, Rockford, Illinois) followed by bioimaging system ChemiGenius-Q (Syngene Inc., Frederick, Maryland) with cooled CCD camera. The amount of sharply demarcated fractions of tetramers and oligomers and of fuzzy high-molecular weight aggregates were quantified with Gene Tools software (Syngene Inc.) after the subtraction of background on an *E.coli* extract lacking CBS that was analyzed in each series.

Estimation of molecular weight by Ferguson plot

Molecular weight of CBS fractions with varying mobility was estimated using a modification of the method published by Ferguson (Ferguson, 1964; Hedrick and Smith, 1968). The original method is based on electrophoresing the analysed protein together with markers in a series of polyacrylamide gels with increasing concentration of acrylamide. We have determined the relative mobilities normalised to the dye front (R_f) of commercial molecular mass markers (Blue Range, Pierce) ferritin (440 kDa), catalase (232kDa), urease (272 and 545 kDa), BSA (66 and 132 kDa) and egg albumin (45kDa) using 10 cm long 5%, 6%, 7%, 8%, 9% and 10% native polyacrylamide gels and Laemmli electrophoresis buffers without SDS (Laemmli, 1970) after staining of the gels by Coomassie blue R. The *E.coli* extracts containing mutant and wild type CBS were separated by the same series of polyacrylamide gels, the CBS fractions were visualized by western blotting and immunodetection as described above and the relative mobilities normalizes to the dye front have been determined.

The logarithms R_f of each band of the standard proteins were plotted against the acrylamide gel concentrations and slope of regression line of each standard protein was calculated. The calibration line was derived as a function of the logarithms of negative slopes and the molecular weights of the standard. The negative slopes of all oligomeric forms of wild type

CBS and of selected mutant proteins were determined by the same method and these values were used for calculation of molecular weight of the different CBS fractions. The bacterial extracts used for this analysis contained the wild type CBS, the p.K102N, p.S466L, p.E302K, p.G305R, p.G307S and p.R125Q variant proteins expressed at 18°C.

CBS enzymatic assay

The CBS activity in *E.coli* lysates was assayed by modification of the previously published procedure (Kozich and Kraus, 1992). The assay mixture contained 100 mM Tris-HCl (pH 8.6), 0.5 mg/mL BSA, 0.5 mM pyridoxal 5'-phosphate, 1mM dithiothreitol, 10 mM homocysteine, 10 mM 2,3,3-²H-labeled serine (Cambridge Isotope Laboratories) and 27.5 µl of *E.coli* extract in lysis buffer containing 75 µg of bacterial protein. The assay mixture was incubated at 37°C for 1 hour in a total volume of 50 µL. For measurement of responsivity to adenosyl compounds the assay mixture contained either 0.5 mM AdoMet or AdoHcy. The activity was determined by the LC-MS/MS measurement of the product of enzyme reaction, 3,3-²H-labeled cystathionine, using the commercially available kit for amino acid analysis (EZ:faast, Phenomenex, Torrance, USA) (Krijt et al, unpublished).

Statistical analyses

Means, standard deviations, and medians were determined using the spreadsheet Microsoft Office Excel 2003 (Microsoft, Czech Republic), linear regression analysis was carried out using the same software.

Results

Selection of mutations and in silico analysis

In order to gain clinically as well as mechanistically relevant insight into the role of misfolding in CBS deficiency we first selected a series of naturally occurring CBS mutations for further laboratory studies. The set of nine most prevalent mutations having a frequency of at least 10 alleles in the CBS Mutation Database was selected; these included the p.A114V, p.R125Q, p.E144K, p.T191M, p.R266K, p.I278T, p.G307S, p.W409_G453del and p.D444N. This set was expanded by additional 18 less frequent mutations known to be localized in different domains of the CBS protein (Meier, et al., 2001; Meier, et al., 2003). For the list of

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

mutations see Table 1, location of the mutations in the CBS active core and in the modelled carboxyterminal domain is shown in Figure 1. The final series reflects well the molecular pathology of CBS deficiency as these 27 mutations are present on about 70% of all pathogenic CBS alleles known in humans to date.

Based on our previous work (Singh, et al., 2007) we hypothesized that CBS mutations of residues buried in protein globule may have more severe effect on folding than those exposed to the solvent. After selecting the mutants we performed an *in silico* analysis with calculation of the solvent-accessible surface area. Based on previously published criteria we considered mutant residues with relative and absolute solvent-accessible surface area higher than 9% and/or 40 Å², respectively, as being solvent exposed while the rest of the mutations were considered being buried in the globule (Mirkovic, et al., 2004). Table 1 shows the predicted solvent accessibility and demonstrates that mutations belonging to these two classes are represented almost equally in the series of 27 variant proteins.

Total amount of CBS mutant proteins expressed in E.coli and their solubility

It has been shown previously that some CBS mutant proteins expressed in bacteria misassemble and aggregate (de Franchis, et al., 1999; Gaustadnes, et al., 2002; Janosik, et al., 2001b; Singh, et al., 2007; Yamanishi, et al., 2006). It is conceivable that aggregated mutant CBS polypeptides will form inclusion bodies and that a substantial fraction of misfolded CBS proteins may thus escape detection if the commonly used supernatant fraction of bacterial extracts is utilized for western blotting analysis.

To test this assumption we first optimized a method for extracting water-soluble as well as aggregated and water-insoluble CBS from bacteria expressing wild type and four mutant enzymes (for details see the Methods section). Boiling in 3% SDS was subsequently used for the entire series of 27 mutant proteins to determine the amount of CBS antigen in both the particulate (i.e. pellet after centrifugation) and non-particulate (i.e. supernatant) fractions of sonicated bacterial extracts. The sum of CBS signal in both fractions is referred to as the total SDS-soluble CBS antigen.

Using the SDS extraction we observed that all 27 mutant proteins were present in detectable amounts not only in the supernatants but more importantly also in the particulate fractions of the bacterial extracts. The presence of CBS antigen in both fractions is shown in Figure 2;

1
2
3 however, this publication gel does not allow inferring on relative proportions of the mutant
4 proteins in the supernatant and particulate fractions neither on their relation to the wild type
5 enzyme (for details see the Methods section and legend to Figure 2). The signal of total SDS-
6 soluble CBS antigen of the mutant proteins was generally somehow decreased with a median
7 of 66% of the wild type CBS (see Supp. Table S2; gels used for quantification are not shown)
8 although some variant proteins were present in increased quantities. These data suggest that
9 compared to the wild type enzyme the majority of mutant CBS proteins are less stable in
10 *E.coli*; however, we cannot exclude that some of them are also less soluble in 3% SDS.
11
12
13
14
15
16
17
18

19 Proportion of the SDS-soluble CBS antigen which is present in the particulate fraction
20 indicates the propensity of the mutant protein to form high molecular weight aggregates and
21 inclusion bodies. The particulate fraction of the wild type CBS lysates contained only 12% of
22 antigen of the total SDS-soluble signal while more of the signal in the pellet was observed for
23 the mutant proteins (median value was 18 % of the total SDS-soluble antigen with a range of
24 5 to 50%; see Supp. Table S2). These data lend support to our hypothesis that the mutant
25 proteins in this study were in general more prone to formation of large molecular weight
26 aggregates and/or inclusion bodies. This part of our study further indicates that the previously
27 reported analyses of mutants CBS in the supernatants of centrifuged bacterial extracts
28 (Kozich, et al., 1993) may have underestimated the steady state amounts of mutant proteins
29 present in bacteria and that extraction by 3% SDS is needed to recover the insoluble
30 aggregated molecular species.
31
32
33
34
35
36
37
38
39
40
41

42 ***Impaired quaternary structure of mutant proteins as a measure of their misfolding***

43 The published data strongly support the hypothesis that CBS deficiency may be a
44 conformational disease caused by misfolding of mutant enzymes. Detailed studies of folding,
45 however, require protein purification which would be especially difficult to accomplish for
46 the most severely affected CBS enzymes. Since misfolded polypeptide chains cannot be
47 correctly assembled and tend to form aggregates we examined the quaternary structure of the
48 mutant proteins in the non-particulate fraction of crude bacterial extracts as a surrogate
49 marker of their folding status. To assess the oligomeric structure of the mutant CBS enzymes
50 we used electrophoresis in gradient polyacrylamide gels under non-denaturing conditions
51 followed by western blotting.
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Visual inspection of blots showed variable amounts of fuzzy and putatively misassembled CBS, and largely differing amounts of clearly demarcated CBS tetramers or higher order oligomers as shown in Figure 2. Subsequent quantitative analysis of the blots revealed that 88% of the total water-soluble antigen of the wild type CBS is present as tetramers/oligomers. In contrast these correctly assembled fractions were substantially decreased in our set of mutant proteins (median tetrameric/oligomeric signal was only 31% of the total water-soluble CBS antigen, range 1%-92%; data are given in Supp. Table S2). Since the CBS antigen in the non-particulate fraction of mutant proteins was also decreased to a median 60% compared to the wild type the resulting net yield of tetramers/oligomers of the mutant proteins was severely reduced (median 12% of tetramers compared to the wild type CBS). There was no clear correlation between the location of mutant proteins in different domains and their misfolding with the exception of mutant proteins in the carboxyterminal domain that were mostly assembled correctly. In summary the above analyses demonstrated that a substantial proportion of CBS mutant proteins formed severely reduced amounts of correctly assembled tetramers; these findings add further support to the notion that misfolding is an important pathogenic mechanism in CBS deficiency.

Different mobility of CBS mutant proteins in native westerns

The above mentioned western blot analysis revealed slight but consistent differences in migration distance of the putative tetramers/oligomers of many mutant proteins in comparison to the wild type CBS tetramer (see Figure 2). These mobility alterations may have resulted from charge differences, conformational changes of tetramers or even different number of subunits in the differently migrating fractions. To explore the latter hypothesis we used the Ferguson plot to determine the molecular weight of four fractions seen for wild type enzyme, and the molecular weight of fast migrating (K102N, S466L), slow migrating (E302K, G305R and G307S) and fast/slow migrating (R125Q) mutant proteins. For results see Supp. Figure S1.

The apparent molecular weights of sharply demarcated and thus presumably correctly assembled oligomer fractions seen for the wild type enzyme were 177, 398, 512 and 724 kDa. Considering the limitations of the technique employing crude extracts with subsequent western blotting and keeping in mind that CBS is known to form tetramers, the results suggest that in our electrophoretic system the wild type CBS is present predominantly in the form of tetramers, and in small amounts also as octamers, dodecamers and hexadecamers. Two

fractions of the R125Q mutant protein were predicted to be composed of tetramers and octamers, respectively. There was only a small difference in slopes and thus in estimated molecular weights of the remaining fast (162 and 166 kDa) and slow (199-213 kDa) migrating mutant proteins as demonstrated in Supp. Figure S1. Consequently, the estimated number of CBS subunits ranged between 2.7 and 3.5 per fraction which is congruent with tetrameric composition of these fractions. In summary, the results of Ferguson plot do not support a hypothesis that different mobility of mutant proteins compared to wild type CBS is caused by changes in the number of enzyme subunits. Thus the observed varying mobility of CBS mutant proteins results most likely from subtle changes in the shape and/or charge of the assembled tetramers.

Catalytic activity of mutant proteins

In order to enable measurements of low residual activities of mutant proteins we firstly developed a sensitive method for the analysis of the CBS reaction product – cystathionine. The radioactive substrate used in previous studies - ^{14}C labeled serine- was replaced by 2,3,3- ^2H -labeled serine and the amount of 3,3- ^2H -labeled cystathionine produced was determined using LC-MS/MS (Krijt et al, unpublished). Using this technique, we were able to measure reproducibly activities as low as 0.3 nmol cystathionine/hour/mg of protein, which represents about 0.2% of the average wild type CBS activity in *E.coli* extracts.

Using the sensitive new assay we observed an extreme variation of activities ranging from 0 to 245% of the wild type enzyme with a median activity of 3% of the wild type CBS (for details see Table 2). Only three mutant proteins- p.G305R, p.T262R and p.N228K- did not exhibit any measurable residual activity above 0.3 nmol/mg/hour while the remaining 24 mutant proteins had above-threshold activity in at least two of the three independent expression experiments. The activities determined in the present study are for the majority of mutant proteins congruent with the previously published values listed in Table 1. In contrast to effects on folding location of the mutant proteins appears to play a more important role in affecting their catalytic activity. There was a tendency for mutations located in the active site, heme binding pocket, and elsewhere in the active core to exhibit rather low activities while mutations inside or on the surface of the first CBS domain and at the dimer-dimer interface were in many cases as active as the wild type enzyme (for details see Table 2). These data suggest that mutations in different domains of the enzyme may have substantially variable

1
2
3 impact on its catalytic activity and that they may not affect folding and activity
4 simultaneously.
5
6
7

8
9 ***Effect of AdoMet and AdoHcy on the catalytic activity of mutant proteins***

10 To get an insight into the possible pathophysiological and therapeutic implications of AdoHcy
11 and AdoMet in CBS deficiency we studied the effect of these two compounds on the activity
12 of mutant proteins in non-particulate fractions of bacterial extracts. Blood concentration of
13 AdoHcy is increased in patients with CBS deficiency (Orendac, et al., 2004) due to its reverse
14 synthesis from homocysteine and adenosine with the help of AdoHcy hydrolase, an enzyme
15 which can be *in vitro* blocked with different inhibitors. AdoMet is an allosteric activator of
16 CBS that can possibly stimulate the residual activity of mutant proteins and that can be
17 administered as a drug in the form of tosylate.
18
19
20
21
22
23
24
25

26 Activity of the wild type CBS enzyme increased 3.9x and 1.3x in the presence of 0.5mM
27 AdoMet and 0.5mM AdoHcy, respectively (see Table 2). Since activation of CBS by AdoHcy
28 has not yet been reported we explored whether it could not have been caused by
29 contamination of this compound by AdoMet. However, LC-MS/MS analysis showed that the
30 AdoHcy standard contained less than 1% of AdoMet indicating that the 1.3x increase in wild
31 type CBS activity in crude bacterial extracts can be attributed to AdoHcy itself.
32
33
34
35
36
37
38

39 We classified without formal statistical testing the response of mutant proteins to AdoMet and
40 AdoHcy in this simple screening system as follows: clear activation similar to the wild type
41 enzyme, clear inhibition, or absence of expected activation. About one half of the mutant
42 enzymes in the panel exhibited an unusual response to either one of the adenosyl compounds:
43 the p.E176K was inhibited by AdoMet, the p.E302K was inhibited by AdoHcy, and the
44 p.H65R, p.T191M and p.C165Y appeared to be inhibited by both adenosyl compounds. In
45 addition, mutations p.E302K, p.W409_G453del, p.I435T, p.L539S, p.E144K, p.G148R,
46 p.G307S, p.I278T and p.S466L were not activated by AdoMet to an extent that was usually
47 observed for the wild type CBS (see Table 2). These data show that responses of about half of
48 mutant proteins to AdoMet and AdoHcy may differ substantially from those of the wild type
49 CBS. The relevance of these findings for regulation of the activity of mutant proteins and
50 possibly for treatment of CBS deficiency is further explored in the Discussion section.
51
52
53
54
55
56
57
58
59
60

Effect of low temperature on properties of mutant proteins

Low temperature during expression has been shown to enhance folding of various mutant protein proteins that are prone to misfolding. To examine whether any of the CBS mutations under study may be rescued by more folding-permissive conditions we expressed all 27 mutant proteins at 18°C. Compared to expression at 37°C the lower temperature has not dramatically changed the total amount of SDS-soluble mutant CBS enzymes and only slightly decreased the proportion of signal in the particulate fraction (median of 18% at 37°C versus 15% of total signal at 18°C; see Supp. Tables S2 and S3). However, the median amount of correctly folded oligomers in the water-soluble non-particulate fraction detected on native-PAGE increased about 2-times from 31% of total signal at 37°C to 68% of total signal at 18°C (Supp. Table S2 and S3). Several mutant proteins responded to low expression temperature with a dramatic (i.e. more than approximately 3-fold) increase of the proportion of tetramers compared to the proportion of tetramers/oligomers of the wild type CBS, namely the p.R266K, p.R369C, p.R125Q, p.P78R, p.V180A and p.I435T. However, this folding promoting effect of low temperature was not universal and some mutant proteins produced conversely less tetramer. All these observations indicate that for some mutant proteins more folding-permissive conditions facilitate the attainment of correct tertiary structure which subsequently increases the assembly of tetramers.

In addition, the low expression temperature increased the median catalytic activity of mutant proteins from 3% of wild type at 37°C to a median of 13% of wild type at 18°C (see Table 2). The mutant proteins p.R266K, p.R125Q, p.R369C, p.V180A and p.P78R exhibited more than approximately 2-fold increase in activity relative to the wild type which correlated well with the increased amount of tetramer. These five mutant proteins may be considered true folding mutants as their residual activity directly correlates with the amount of tetramer. In summary, the lower expression temperature substantially increased activity of five mutant proteins further supporting the hypothesis that deficient CBS activity may be in some instances caused by misfolding that may be alleviated by folding permissive conditions.

Solvent exposure of mutant residues as a determinant of folding and activity

We hypothesized that topology of mutations in respect to solvent exposure/buriedness may have an impact on folding and assembly of mutant CBS molecules. Therefore we compared in the next step the typical properties of CBS enzymes carrying buried versus solvent accessible mutations by comparing median values in Table 2 and in Supp. Table S2 and S3. We have observed similar stability of these two classes of mutant proteins in *E.coli* as evidenced by

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

identical amounts of total SDS-soluble antigen. However, the propensity of buried mutations to correctly fold and assemble is impaired as clearly demonstrated by the lower proportion of tetrameric/oligomeric fractions in the total CBS antigen in native gels with a median of only 14% of total antigen for buried mutations versus 36% for the solvent accessible ones. More importantly the buried mutations are about 23-times less active with median activity of only 0.4% of the wild type CBS versus 9.4% of the wild type enzyme for solvent exposed mutations. In addition, the buried mutations are not activated by AdoMet (the median activation of buried mutations of 1.0x contrasts with the 2.5x activation of solvent exposed mutation) and seem to be inhibited by AdoHcy (median activation of buried mutations is 0.8x while it is 1.3x for solvent exposed mutations). These differences between buried and solvent exposed mutations appear to be further augmented by the expression at 18°C. In summary these comparisons strongly suggest that buried mutations may lead to much severe impairment of CBS structure and function than the solvent exposed ones.

Expression of several mutant proteins at 18°C resulted in an increase in their catalytic activity as well as in the amount of tetramers/oligomers; we therefore proposed that these variant proteins may be considered true folding mutations. To evaluate this phenomenon in a more systematic way we have plotted the normalized activity of mutant proteins against the normalized amount of tetramers/oligomers using the combined data from five independent expression experiments. Supp. Figure S2 demonstrates that among the solvent exposed mutations there are indeed examples of such increase of activities mirroring the increased amounts of tetramers (e.g. p.A114V, p.H65R, p.K102N, p.P78R, p.R125Q and p.V180A). However, changes in amounts of tetramers of buried mutations were not usually accompanied by corresponding differences in activities. That led us to speculate that the solvent exposure of mutations may be indeed an important determinant of their propensity to misfold and to lose the catalytic activity. We have next performed a regression analysis in which we combined data for all buried and all solvent exposed mutant proteins, respectively. This analysis has demonstrated that the activity of solvent exposed mutations correlates quite strongly with the amount of oligomers ($r^2 = 0.53$; see Figure 3, Panel B) and suggests that for this type of mutations enhanced folding achieved by e.g. low temperature or chaperones may result in increased residual activity. In contrast the mutations buried in the enzyme globule did not exhibit such properties as there was only a weak correlation between the activity and the amount of oligomers ($r^2 = 0.11$; see Figure 3, panel A). The above described regression

analysis suggests that solvent exposed CBS mutations may be more likely to respond to interventions aimed at correcting their misfolding.

Discussion

This paper focused on testing in a systematic way the hypothesis that misfolding plays an important role in CBS deficiency. To examine the hypothesis we selected a representative set of mutants located in different functional domains of the CBS molecule. Several lines of evidence suggest that misfolding is an important pathogenic mechanism in CBS deficiency: (i) the amount of SDS-soluble mutant CBS polypeptides in particulate fraction was increased by 50% compared to the wild type possibly due to inclusion body formation, (ii) the net yield of mutant tetramers/oligomers in the non-particulate water-soluble fractions was severely reduced to only 12% of the wild type enzyme, and (iii) more folding permissive conditions (i.e. expression at 18°C) led to a doubling of tetramer amounts and partial restoration of activity. The mutant proteins studied in our paper represent about 70% of known patient derived CBS alleles, and 18 of these 27 mutations clearly result into moderate to severe misfolding of the CBS enzyme (i.e. formation of less than 25% of tetramer of the wild type). This is an important finding which is relevant to 362 of 753 mutant alleles deposited in the CBS Mutation Database as of April 2010; in other words, about one half of all known chromosomes affected by pathogenic mutations code for mutant polypeptides that misfold and misassemble. Taken together all these data suggest that abnormal folding is an important and common molecular mechanism in CBS deficiency.

An important finding of our study is the more important role of solvent accessibility of mutations rather than their location in enzyme domains in predicting properties of mutant enzymes. In general, buried mutations exhibited much more profound detrimental effect on enzyme folding and activity compared to the solvent accessible mutations. Appearance of sufficient amounts of tetramers in native gels did not predict enzyme activity for the buried mutants, however, increased tetramer formation in surface mutants elicited by low expression temperature led to an increase in residual activity. The plot of activity versus amount of tetramer (see Figures 3 and Supp. Figure S2) shows a good correlation for solvent exposed mutants implying that enzyme activity of these genetic variants is determined in part by the

amount of tetramers and in part by the precise location of the mutant residue and the character of the change.

In this study we observed clearly abnormal response to AdoMet and AdoHcy for about one half of mutant proteins: (i) several previously characterized mutants in the COOH-terminal regulatory region were hyperactive and were not further activated by AdoMet (Dawson, et al., 1997; Kluijtmans, et al., 1999; Kluijtmans, et al., 1996); these mutants were previously shown to have similar conformation to that of the activated wild type CBS with bound AdoMet, (ii) in contrast, another class of mutant proteins with low activity was either not stimulated or was even inhibited by AdoMet, and (iii) several mutant proteins were inhibited by AdoHcy. It has been shown previously that whole blood or plasma concentrations of AdoHcy are grossly elevated and that the AdoMet levels are either normal or increased in patients with CBS deficiency (Heil, et al., 2007; Orendac, et al., 2003). It is at present unknown whether the disturbed AdoMet and AdoHcy metabolism in CBS deficiency may modulate the residual activity of mutant enzymes and whether these observations may have any therapeutic implications.

The major limitation of our work relates to the different intracellular milieu of prokaryotic and eukaryotic cells, and especially to the differences in folding machineries and in protein surveillance/degradation systems. Therefore the conclusions from our study have to be taken with care and have to be validated by analyses of mutant proteins that will be expressed in eukaryotic cells or obtained from tissues of animal models of CBS deficiency. Despite these limitations, utility of the simple bacterial expression systems has been appreciated for their ability to demonstrate the inherent predisposition of other mutant enzymes such as phenylalanine hydroxyls towards aberrant folding and misassembly (Waters, 2003).

The above described observations may lead to the development of new strategies for treating patients with CBS deficiency. The most important implication is the potential of chaperones to rescue to a varying extent the misfolding and decreased activity of a substantial proportion of mutant proteins. We have shown previously in a selected set of mutant proteins that chemical chaperones may restore biological function of CBS mutants (Singh, et al., 2007). In addition, we examined the effect of six chemical chaperones or CBS ligands on a set of 27 mutant proteins presented in this paper and we have observed an effect of some of these compounds (Kopecka, et al., in press). However, the practical utility of chemical chaperones

for the treatment of misfolding disorders is modest and a more promising modality would be the use of a specific pharmacological CBS chaperone that awaits its discovery. The second group of treatment options implied theoretically from our present study are the AdoHcy hydrolase inhibitors that may decrease the concentrations of AdoHcy and alleviate thus inhibitory effect of AdoHcy on some mutant proteins, or the administration of AdoMet tosylate to stimulate the residual activity of other mutant CBS enzymes. AdoMet tosylate in doses of several hundreds of milligrams per day has been used in treatment of liver disease, myopathies and depression without reports on adverse effects (Mato, et al., 2001). Whether or not chaperones and adenosyl compounds will be of any use in patients with CBS deficiency and will improve quality of their lives is presently unknown and needs to be explored in the future.

Acknowledgement

The authors would like to thank Ms. Alena Dutá and Kateřina Raková for technical help and Aleš Hnízda, MSc. and Jana Kopecká, MSc. for critical reading of the manuscript. This work was supported by the Wellcome Trust International Senior Research Fellowship in Biomedical Science in Central Europe (Reg. No. 070255/Z/03/Z), institutional support was provided by the Research Project from the Ministry of Education of the Czech Republic (Reg. No. MSM0021620806).

References Endnote

- Carra S, Sivilotti M, Chavez Zobel AT, Lambert H, Landry J. 2005. HspB8, a small heat shock protein mutated in human neuromuscular disorders, has in vivo chaperone activity in cultured cells. *Hum Mol Genet* 14(12):1659-69.
- Chen S, Ito M, Saijo T, Naito E, Kuroda Y. 1999. Molecular genetic analysis of pyridoxine-nonresponsive homocystinuric siblings with different blood methionine levels during the neonatal period. *J Med Invest* 46(3-4):186-91.
- Chen X, Wang L, Fazlieva R, Kruger WD. 2006. Contrasting behaviors of mutant cystathionine beta-synthase enzymes associated with pyridoxine response. *Hum Mutat* 27(5):474-82.
- Dawson PA, Cox AJ, Emmerson BT, Dudman NP, Kraus JP, Gordon RB. 1997. Characterisation of five missense mutations in the cystathionine beta-synthase gene from three patients with B6-nonresponsive homocystinuria. *Eur J Hum Genet* 5(1):15-21.
- de Franchis R, Kozich V, McInnes RR, Kraus JP. 1994. Identical genotypes in siblings with different homocystinuric phenotypes: identification of three mutations in cystathionine beta-synthase using an improved bacterial expression system. *Hum Mol Genet* 3(7):1103-8.
- de Franchis R, Kraus E, Kozich V, Sebastio G, Kraus JP. 1999. Four novel mutations in the cystathionine beta-synthase gene: effect of a second linked mutation on the severity of the homocystinuric phenotype. *Hum Mutat* 13(6):453-7.
- Evande R, Blom H, Boers GH, Banerjee R. 2002. Alleviation of intrasteric inhibition by the pathogenic activation domain mutation, D444N, in human cystathionine beta-synthase. *Biochemistry* 41(39):11832-7.
- Ferguson KA. 1964. Starch-Gel Electrophoresis--Application to the Classification of Pituitary Proteins and Polypeptides. *Metabolism* 13:SUPPL:985-1002.
- Gaustadnes M, Ingerslev J, Rutiger N. 1999. Prevalence of congenital homocystinuria in Denmark. *N Engl J Med* 340(19):1513.
- Gaustadnes M, Wilcken B, Oliveriusova J, McGill J, Fletcher J, Kraus JP, Wilcken DE. 2002. The molecular basis of cystathionine beta-synthase deficiency in Australian patients: genotype-phenotype correlations and response to treatment. *Hum Mutat* 20(2):117-26.
- Gordon RB, Cox AJ, Dawson PA, Emmerson BT, Kraus JP, Dudman NP. 1998. Mutational analysis of the cystathionine beta-synthase gene: a splicing mutation, two missense mutations and an insertion in patients with homocystinuria. *Mutations in brief no. 120*. Online. *Hum Mutat* 11(4):332.
- Gupta S, Wang L, Hua X, Krijt J, Kozich V, Kruger WD. 2008. Cystathionine beta-synthase p.S466L mutation causes hyperhomocysteinemia in mice. *Hum Mutat* 29(8):1048-54.
- Hedrick JL, Smith AJ. 1968. Size and charge isomer separation and estimation of molecular weights of proteins by disc gel electrophoresis. *Arch Biochem Biophys* 126(1):155-64.
- Heil SG, Riksen NP, Boers GH, Smulders Y, Blom HJ. 2007. DNA methylation status is not impaired in treated cystathionine beta-synthase (CBS) deficient patients. *Mol Genet Metab* 91(1):55-60.
- Hu FL, Gu Z, Kozich V, Kraus JP, Ramesh V, Shih VE. 1993. Molecular basis of cystathionine beta-synthase deficiency in pyridoxine responsive and nonresponsive homocystinuria. *Hum Mol Genet* 2(11):1857-60.
- Janosik M, Kery V, Gaustadnes M, Maclean KN, Kraus JP. 2001a. Regulation of human cystathionine beta-synthase by S-adenosyl-L-methionine: evidence for two

- catalytically active conformations involving an autoinhibitory domain in the C-terminal region. *Biochemistry* 40(35):10625-33.
- Janosik M, Oliveriusova J, Janosikova B, Sokolova J, Kraus E, Kraus JP, Kozich V. 2001b. Impaired heme binding and aggregation of mutant cystathionine beta-synthase subunits in homocystinuria. *Am J Hum Genet* 68(6):1506-13.
- Janosik M, Sokolova J, Janosikova B, Krijt J, Klatovska V, Kozich V. 2009. Birth prevalence of homocystinuria in Central Europe: frequency and pathogenicity of mutation c.1105C>T (p.R369C) in the cystathionine beta-synthase gene. *J Pediatr* 154(3):431-7.
- Katsushima F, Oliveriusova J, Sakamoto O, Ohura T, Kondo Y, Iinuma K, Kraus E, Stouracova R, Kraus JP. 2006. Expression study of mutant cystathionine beta-synthase found in Japanese patients with homocystinuria. *Mol Genet Metab* 87(4):323-8.
- Kim CE, Gallagher PM, Guttormsen AB, Refsum H, Ueland PM, Ose L, Folling I, Whitehead AS, Tsai MY, Kruger WD. 1997. Functional modeling of vitamin responsiveness in yeast: a common pyridoxine-responsive cystathionine beta-synthase mutation in homocystinuria. *Hum Mol Genet* 6(13):2213-21.
- Kluijtmans LA, Boers GH, Kraus JP, van den Heuvel LP, Cruysberg JR, Trijbels FJ, Blom H. 1998. The molecular basis of homocystinuria due to cystathionine beta-synthase deficiency in Dutch homocystinuria patients. Effect of CBS genotyp on biochemical and clinical phenotype, and on response upon treatment. PhD thesis. The Netherlands: University of Nijmegen.
- Kluijtmans LA, Boers GH, Kraus JP, van den Heuvel LP, Cruysberg JR, Trijbels FJ, Blom HJ. 1999. The molecular basis of cystathionine beta-synthase deficiency in Dutch patients with homocystinuria: effect of CBS genotype on biochemical and clinical phenotype and on response to treatment. *Am J Hum Genet* 65(1):59-67.
- Kluijtmans LA, Boers GH, Stevens EM, Renier WO, Kraus JP, Trijbels FJ, van den Heuvel LP, Blom HJ. 1996. Defective cystathionine beta-synthase regulation by S-adenosylmethionine in a partially pyridoxine responsive homocystinuria patient. *J Clin Invest* 98(2):285-9.
- Kopecka J, Krijt J, Rakova K, Kozich V. In press. Restoring assembly and activity of cystathionine beta-synthase mutants by ligands and chemical chaperones. *J Inherit Metab Dis*. DOI 10.1007/s10545-010-9087-5.
- Kozich V, de Franchis R, Kraus JP. 1993. Molecular defect in a patient with pyridoxine-responsive homocystinuria. *Hum Mol Genet* 2(6):815-6.
- Kozich V, Kraus JP. 1992. Screening for mutations by expressing patient cDNA segments in *E. coli*: homocystinuria due to cystathionine beta-synthase deficiency. *Hum Mutat* 1(2):113-23.
- Kruger WD, Wang L, Jhee KH, Singh RH, Elsas LJ, 2nd. 2003. Cystathionine beta-synthase deficiency in Georgia (USA): correlation of clinical and biochemical phenotype with genotype. *Hum Mutat* 22(6):434-41.
- Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227(5259):680-5.
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. 1951. Protein measurement with the Folin phenol reagent. *J Biol Chem* 193(1):265-75.
- Maclean KN, Gaustadnes M, Oliveriusova J, Janosik M, Kraus E, Kozich V, Kery V, Skovby F, Rudiger N, Ingerslev J and others. 2002. High homocysteine and thrombosis without connective tissue disorders are associated with a novel class of cystathionine beta-synthase (CBS) mutations. *Hum Mutat* 19(6):641-55.
- Marble M, Geraghty MT, de Franchis R, Kraus JP, Valle D. 1994. Characterization of a cystathionine beta-synthase allele with three mutations in cis in a patient with B6 nonresponsive homocystinuria. *Hum Mol Genet* 3(10):1883-6.

- Mato JM, Avila MA, Corrales FJ. 2001. Biosynthesis of S-adenosylmethionin. In: Carmel R, Jacobsen DW, editors. *Homocystein in Health and Disease*. Cambridge. p 47-62.
- Meier M, Janosik M, Kery V, Kraus JP, Burkhard P. 2001. Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *Embo J* 20(15):3910-6.
- Meier M, Oliveriusova J, Kraus JP, Burkhard P. 2003. Structural insights into mutations of cystathionine beta-synthase. *Biochim Biophys Acta* 1647(1-2):206-13.
- Mirkovic N, Marti-Renom MA, Weber BL, Sali A, Monteiro AN. 2004. Structure-based assessment of missense mutations in human BRCA1: implications for breast and ovarian cancer predisposition. *Cancer Res* 64(11):3790-7.
- Mudd SH, Levy HL, Kraus JP. 2001. Disorders of transsulfuration. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. *The Metabolic and Molecular Bases of Inherited Disease*. 8th edition. New York: McGraw-Hill p2007-2056.
- Ojha S, Wu J, LoBrutto R, Banerjee R. 2002. Effects of heme ligand mutations including a pathogenic variant, H65R, on the properties of human cystathionine beta-synthase. *Biochemistry* 41(14):4649-54.
- Orendae M, Pronicka E, Kubalska J, Janosik M, Sokolova J, Linnebank M, Koch HG, Kozich V. 2004. Identification and functional analysis of two novel mutations in the CBS gene in Polish patients with homocystinuria. *Hum Mutat* 23(6):631.
- Orendac M, Zeman J, Stabler SP, Allen RH, Kraus JP, Bodamer O, Stockler-Ipsiroglu S, Kvasnicka J, Kozich V. 2003. Homocystinuria due to cystathionine beta-synthase deficiency: novel biochemical findings and treatment efficacy. *J Inherit Metab Dis* 26(8):761-73.
- Proudfoot M, Sanders SA, Singer A, Zhang R, Brown G, Binkowski A, Xu L, Lukin JA, Murzin AG, Joachimiak A and others. 2008. Biochemical and structural characterization of a novel family of cystathionine beta-synthase domain proteins fused to a Zn ribbon-like domain. *J Mol Biol* 375(1):301-15.
- Refsum H, Fredriksen A, Meyer K, Ueland PM, Kase BF. 2004. Birth prevalence of homocystinuria. *J Pediatr* 144(6):830-2.
- Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234(3):779-815.
- Sali A, Overington JP. 1994. Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Sci* 3(9):1582-96.
- Sebastio G, Sperandio MP, Panico M, de Franchis R, Kraus JP, Andria G. 1995. The molecular basis of homocystinuria due to cystathionine beta-synthase deficiency in Italian families, and report of four novel mutations. *Am J Hum Genet* 56(6):1324-33.
- Sen S, Yu J, Yamanishi M, Schellhorn D, Banerjee R. 2005. Mapping peptides correlated with transmission of intrasteric inhibition and allosteric activation in human cystathionine beta-synthase. *Biochemistry* 44(43):14210-6.
- Shan X, Dunbrack RL, Jr., Christopher SA, Kruger WD. 2001. Mutations in the regulatory domain of cystathionine beta synthase can functionally suppress patient-derived mutations in cis. *Hum Mol Genet* 10(6):635-43.
- Singh LR, Chen X, Kozich V, Kruger WD. 2007. Chemical chaperone rescue of mutant human cystathionine beta-synthase. *Mol Genet Metab* 91(4):335-42.
- Urreizti R, Asteggiano C, Cozar M, Frank N, Vilaseca MA, Grinberg D, Balcells S. 2006. Functional assays testing pathogenicity of 14 cystathionine-beta synthase mutations. *Hum Mutat* 27(2):211.
- Wang L, Chen X, Tang B, Hua X, Klein-Szanto A, Kruger WD. 2005. Expression of mutant human cystathionine beta-synthase rescues neonatal lethality but not homocystinuria in a mouse model. *Hum Mol Genet* 14(15):2201-8.

- 1
2
3 Waters PJ. 2003. How PAH gene mutations cause hyper-phenylalaninemia and why
4 mechanism matters: insights from in vitro expression. Hum Mutat 21(4):357-69.
5 Yamanishi M, Kabil O, Sen S, Banerjee R. 2006. Structural insights into pathogenic
6 mutations in heme-dependent cystathionine-beta-synthase. J Inorg Biochem
7 100(12):1988-95.
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE LEGENDS

Figure 1. Location of studied mutations in three dimensional CBS structure. Panel A. Mutations located at the surface of the active core are shown in magenta, PLP is shown in yellow. Panel B. Mutations buried within the active core are indicated in green, heme and PLP molecules are shown in red and yellow, respectively. Panel C. Mutations located at the surface of the carboxyterminal regulatory CBS domain are shown in red.

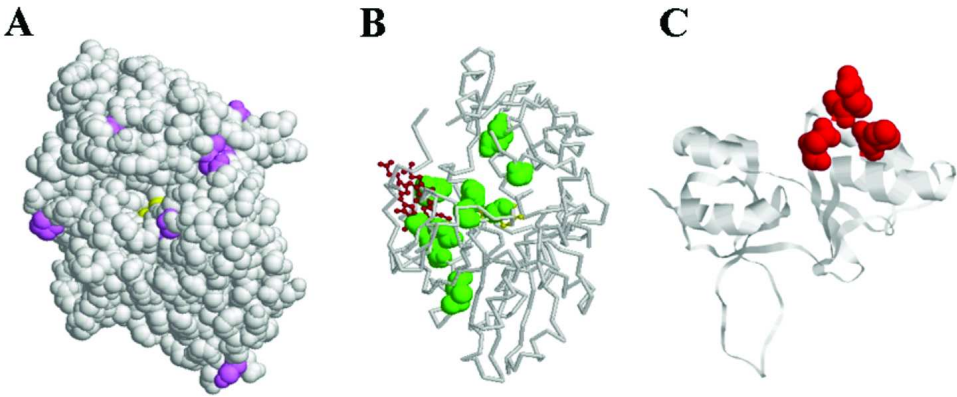
Figure 2. Amounts of mutant CBS antigen in fractions of bacterial extracts. Top part of panels. The amount of SDS-soluble CBS antigen in particulate (Pel, pellet) and non-particulate (Sup, supernatant) fractions of bacterial extracts obtained by centrifugation was determined by SDS-PAGE and western blotting as described in Methods. The blots in this figure are only shown to demonstrate the presence of CBS antigen in both fractions and do not reliably reflect the proportion between the fractions; the particulate fractions and the supernatant fractions were loaded on separate gels while in the experimental blots used for quantification these two fractions of each mutant were loaded next to each other on the same gel (data not shown). Only a single representative sample from 2-3 independent expression experiments has been loaded for this publication gel. **Bottom part of panels.** The quaternary structure of CBS mutants was assessed in the water-soluble non-particulate fraction of bacterial extracts by electrophoresis under native conditions followed by western blotting; sharply demarcated fractions are tetramers and higher oligomers. Only a single representative sample from 2-3 independent expression experiments has been loaded for this publication gel.

Panel A, mutant proteins expressed at 37°C; **Panel B**, results of expression of mutant proteins at 18°C; **Solvent exposed mutations**, mutations with accessible surface area larger than 40 Å² and/or with relative accessible surface larger than 9%, mutations p.C165Y, p.P422L and

p.S466L belong to buried mutations; **Buried mutations**, remaining mutations from the panel are grouped here, mutation p.H65R belongs to solvent exposed mutations.

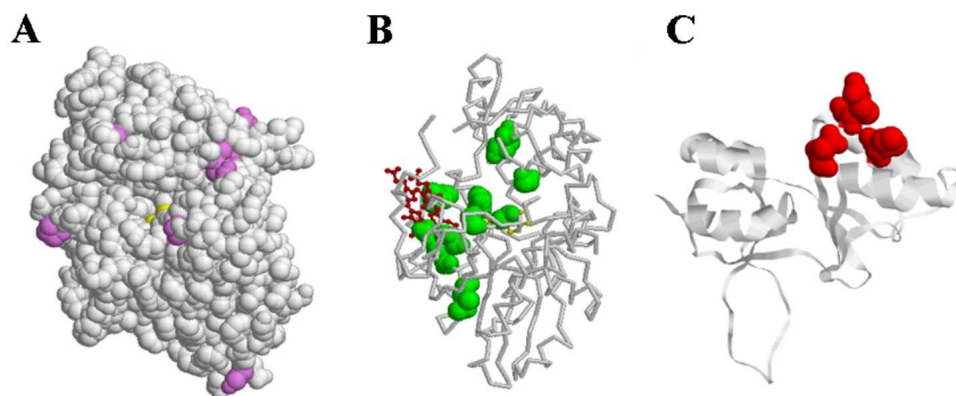
WT, bacterial extracts containing wild type CBS; pKK 388.1, extracts of bacteria transformed with the empty pKK 388.1 plasmid lacking any CBS

Figure 3. Solvent exposure and misassembly of mutants. Data on normalized activities and tetramer amounts shown in Supp. Figure S2 are here combined separately for all solvent exposed and all buried mutations; O and Δ indicate data for expression at 37°C and 18°C, respectively. Linear regression analysis shows much stronger correlation between activity and the amount of tetramers for solvent exposed mutations than for the buried ones.



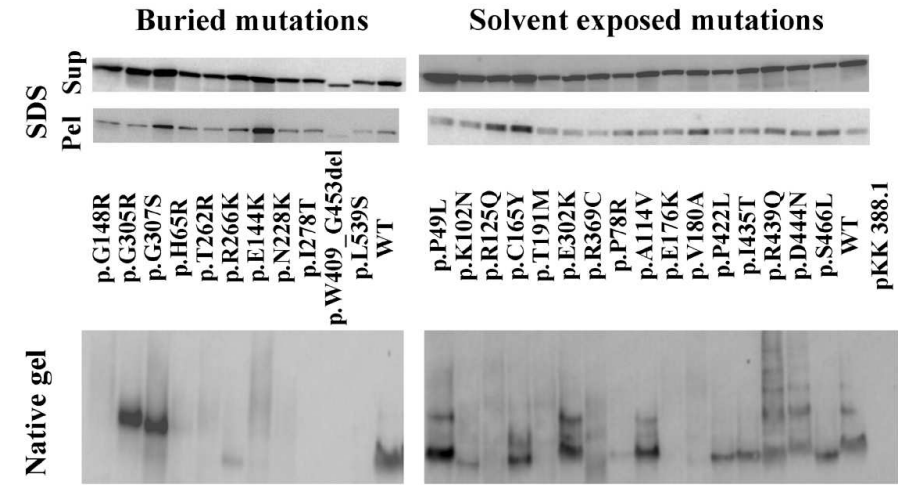
154x70mm (248 x 248 DPI)

Peer Review

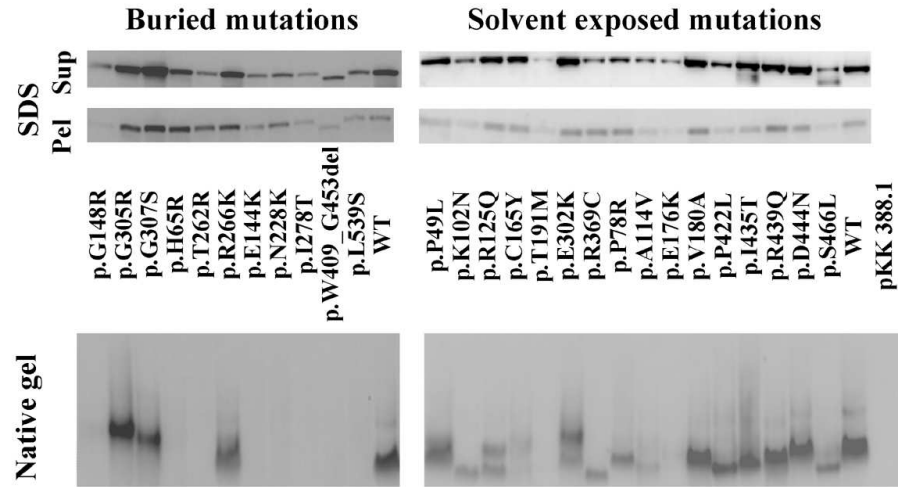


154x70mm (248 x 248 DPI)

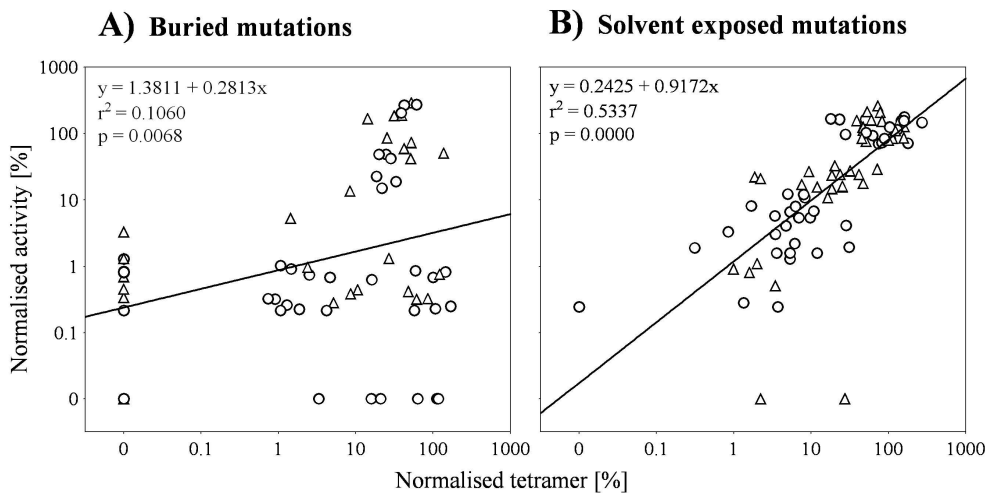
A) Expression at 37°C



B) Expression at 18°C



99x129mm (300 x 300 DPI)



179x94mm (384 x 384 DPI)

Peer Review

Location	Mutation		Solvent accesibility			Number of alleles	Properties of mutant enzymes
	Protein change	cDNA change	Absolute [A ²]	Relative [%]	Class		
Active site	p.G148R	c.442 G>A	1.31	1.80	B	3	ECE: tetramer, 0-5.82% WT activity (Katsushima, et al., 2006; Orendae, et al., 2004; Urreizti, et al., 2006)
	p.G305R	c.913 G>A	3.58	4.90	B	1	
	p.G307S	c.919 G>A	0.16	0.50	B	88	ECE: undetectable activity (Hu, et al., 1993) SCE: <1.7% WT activity (Kruger, et al., 2003)
Heme binding pocket	p.H65R	c.194 A>G	22.01	30.30	S	3	PE: 11.5% WT activity, 75% PLP saturation, 40% heme saturation (Ojha, et al., 2002) ECE: undetectable activity (Chen, et al., 1999)
	p.T262R	c.785 C>G	0.14	0.20	B	1	
	p.R266K	c.797 G>A	1.76	2.90	B	10	PE: Vmax 28-114% of WT, K _m [PLP]: 3x of WT; SCE: tetramer, 49% WT activity, K _m [PLP]: 10x of WT (Chen, et al., 2006)
Other locations in active core	p.E144K	c.430 G>A	0.10	0.20	B	12	ECE: <1% WT activity (Dawson, et al., 1997; Janosik, et al., 2001b)
	p.C165Y	c.494 G>A	1.85	3.00	B	8	ECE: 0-3% WT activity (Gordon, et al., 1998; Kluijtmans, et al., 1998; Kluijtmans, et al., 1999)
	p.N228K	c.684 C>G/A	0.06	0.10	B	4	ECE: aggregates, 0% WT activity (Gaustadnes, et al., 2002)

	p.I278T	c.833 T>C	0.00	0.00	B	146	PE: V_{\max} 0.8-1.4% of WT, K_m [PLP]: 1.4x of WT (Chen, et al., 2006); ECE: aggregates, 0-2% WT activity (Janosik, et al., 2001b; Kluijtmans, et al., 1998; Kluijtmans, et al., 1999; Kozich and Kraus, 1992) SCE: WT activity, tetramer, <2%-5% WT activity, K_m [PLP]: 0.7x of WT (Chen, et al., 2006; Kruger, et al., 2003; Shan, et al., 2001; Singh, et al., 2007); MLE: ~2% WT activity (Wang, et al., 2005)
	p.P49L	c.146 C>T	35.01	61.90	S	3	
	p.K102N	c.306 G>C	17.74	43.40	S	1	ECE: 21-34% WT activity (de Franchis, et al., 1994)
	p.R125Q	c.374 G>A	8.57	16.60	S	13	ECE: 0-1.9% WT activity (Marble, et al., 1994; Sebastio, et al., 1995)
	p.T191M	c.572 C>T	6.93	11.20	S	137	ECE: 1% WT activity (Kluijtmans, et al., 1998)
	p.E302K	c.904 G>A	15.03	24.50	S	6	ECE: shorter product -partially degraded, 4.9% WT activity (de Franchis, et al., 1999)
	p.R369C	c.1105 C>T	4.55	10.90	S	4	SCE: does not affect CBS function (Kim, et al., 1997) ECE: 34% WT activity (Janosik, et al., 2009) CHO: 54% WT activity, V_{\max} 29% of WT, K_m [serin]: 0.8x of WT, K_m [homocystein]: 1.3x of WT (Janosik, et al., 2009)

Dimer - dimer interface	p.P78R	c.233C>G	47.79	65.80	S	2	ECE: 13-61% WT activity (de Franchis, et al., 1994)
	p.A114V	c.341 C>T	8.94	18.07	S	10	ECE: tetramer + aggregates (Janosik, et al., 2001b), 46.3-50% WT activity (de Franchis, et al., 1999; Sebastio, et al., 1995)
	p.E176K	c.526 G>A	16.35	26.60	S	2	ECE: aggregates, 0% WT activity (Janosik, et al., 2001b)
	p.V180A	c.539 T>C	5.15	15.50	S	1	ECE: 3-8% WT activity (Kluijtmans, et al., 1998; Kluijtmans, et al., 1999)
Connection loop between active core and regulatory domain	p.W409_G453del	c.1224-2 A>C (r.1224_1358del)			B	19	ECE: aggregates (Janosik, et al., 2001b), inactive (Kozich and Kraus, 1992)
Regulatory domain-first CBS domain	p.P422L	c.1265 C>T	4.89	8.70	B	1	ECE: ~160% WT activity, 0.16x activation by AdoMet compared to WT; CHFE: ~150% WT activity (Maclean, et al., 2002)
	p.I435T	c.1304 T>C	16.11	38.50	S	1	ECE: tetramer (Janosik, et al., 2001a), ~290% WT activity, 0.03x activation by AdoMet compared to WT (Maclean, et al., 2002); CHFE: ~100% WT activity (Maclean, et al., 2002)
	p.R439Q	c.1316 G>A	36.79	71.20	S	3	ECE: 30% WT activity (Dawson, et al., 1997)
	p.D444N	c.1330 G>A	18.73	45.80	S	13	PE: 1.5x of WT activity, $K_{act[AdoMet]}$ 62x of WT (Evande, et al., 2002); ECE: 96-100% WT activity (Kluijtmans, et al., 1999; Kluijtmans, et al., 1996)

	p.S466L	c.1397 C>T	0.23	0.40	B	1	ECE: ~310% WT activity, 0.05x activation by AdoMet compared to WT (Maclean, et al., 2002); CHFE: ~300% WT activity (Maclean, et al., 2002); MLE: tetramer, ~400% WT activity (Gupta, et al., 2008)
Regulatory domain-second CBS domain	p.L539S	c.1616 T>C	1.93	5.70	B	2	

Table 1. Panel of studied mutations. Mutations selected for this study are grouped according to their location in different domains of the CBS 3-dimensional structure (Meier, et al., 2001; Meier, et al., 2003). Nucleotide numbering reflects the cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference GenBank sequence NM_000071.2, according to the Journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon number 1. Solvent accessibility has been calculated using the program Modeller, mutations are classified based on predicted solvent accessibility as solvent exposed (S, accessible surface area larger than 40\AA^2 and/or relative accessible surface area larger than 9%) or buried (B). Number of alleles carrying particular mutation has been taken from the CBS mutation database (<http://www.uchsc.edu/cbs/cbsdata/cbsmain.htm>) as of April 7, 2010. Data on properties of mutant enzymes has been extracted from literature, expression systems are abbreviated as follows: ECE, *E.coli* crude extracts; SCE, *S.cerevisiae* crude extracts; CHFE, Chinese hamster fibroblasts crude extracts; CHO, Chinese hamster ovary cells crude extract; MLE, mouse liver extracts; PE, purified enzymes.

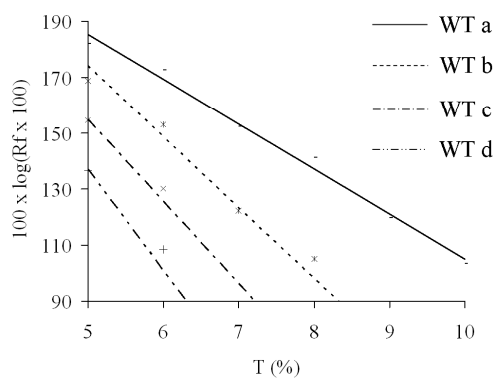
	Solvent exposure	Mutation	Activity [nmol cystathionine/mg protein/hour]					
			Expression at 37°C			Expression at 18°C		
			no addition	AdoMet (0.5mM)	AdoHcy (0.5mM)	no addition	AdoMet (0.5mM)	AdoHcy (0.5mM)
Active site	B	p.G148R	0.2 ± 0.1	0.2 ± 0.2	0.3 ± 0.3	0.4 ± 0.1	N/A	N/A
	B	p.G305R	n.d.	n.d.	n.d.	0.4 ± 0.02	N/A	N/A
	B	p.G307S	0.3 ± 0.02	0.2 ± 0.1	0.4 ± 0.01	0.5 ± 0.2	N/A	N/A
Heme binding pocket	S	p.H65R	5.6 ± 1.9	4.4 ± 1.6	1.9 ± 0.7	0.8 ± 0.2	N/A	N/A
	B	p.T262R	n.d.	n.d.	n.d.	0.2 ± 0.2	N/A	N/A
	B	p.R266K	25.3 ± 4.2	82.3 ± 23.0	29.4 ± 0.03	47.8 ± 24.7	325.3 ± 144.7	N/A
Other locations in active core	B	p.E144K	0.6 ± 0.4	0.9 ± 0.6	0.5 ± 0.5	3.3 ± 2.8	N/A	N/A
	B	p.C165Y	1.1 ± 0.1	0.6 ± 0.04	0.9 ± 0.01	0.5 ± 0.05	N/A	N/A
	B	p.N228K	n.d.	n.d.	n.d.	0.4 ± 0.005	N/A	N/A
	B	p.I278T	0.4 ± 0.04	0.6 ± 0.1	0.4 ± 0.1	0.9 ± 0.2	N/A	N/A
	S	p.P49L	147.9 ± 38.9	512.8 ± 147.3	170.9 ± 34.1	103.4 ± 43.2	469.4 ± 146.0	89.8 ± 20.7
	S	p.K102N	8.8 ± 1.7	25.7 ± 4.3	9.2 ± 0.9	13.6 ± 10.2	58.2 ± 44.5	12.0 ± 5.6
	S	p.R125Q	2.3 ± 0.3	8.1 ± 2.5	2.7 ± 0.3	19.4 ± 8.8	141.8 ± 78.0	N/A
	S	p.T191M	0.4 ± 0.02	0.2 ± 0.2	n.d.	0.6 ± 0.03	N/A	N/A
	S	p.E302K	136.2 ± 19.3	147.5 ± 10.1	112.3 ± 4.3	81.5 ± 28.9	95.3 ± 29.2	107.7 ± 68.1
Dimer-dimer interface	S	p.R369C	2.5 ± 0.4	6.4 ± 0.3	2.4 ± 0.1	13.1 ± 0.4	57.5 ± 8.1	N/A
	S	p.P78R	13.8 ± 5.0	43.4 ± 12.2	13.2 ± 4.9	26.4 ± 9.2	149.7 ± 36.1	31.0 ± 8.4
	S	p.A114V	109.7 ± 10.5	303.3 ± 36.7	117.0 ± 19.0	13.0 ± 1.2	28.0 ± 9.4	11.5 ± 2.2
	S	p.E176K	4.9 ± 0.6	2.7 ± 0.3	4.4 ± 0.6	11.5 ± 8.3	7.9 ± 5.6	N/A
	S	p.V180A	13.1 ± 3.5	24.5 ± 5.3	21.7 ± 0.05	127.3 ± 15.2	457.5 ± 65.2	N/A

Connection loop between active core and regulatory domain	B	p.W409_G453del	1.2 ± 0.2	1.2 ± 0.2	1.4 ± 0.1	0.8 ± 0.3	N/A	N/A
Regulatory domain - first CBS domain	B	p.P422L	65.5 ± 12.6	165.8 ± 38.1	67.0 ± 16.6	113.2 ± 77.9	228.4 ± 80.3	119.3 ± 92.0
	S	p.I435T	196.5 ± 48.2	268.8 ± 27.9	186.6 ± 52.8	215.4 ± 72.4	575.2 ± 206.5	215.5 ± 50.7
	S	p.R439Q	167.3 ± 28.4	774.0 ± 56.4	179.8 ± 0.04	76.4 ± 14.8	705.3 ± 257.0	N/A
	S	p.D444N	233.7 ± 39.9	817.9 ± 101.5	313.6 ± 22.6	117.0 ± 27.4	675.0 ± 174.2	140.4 ± 3.8
	B	p.S466L	346.2 ± 69.1	325.2 ± 63.1	353.7 ± 59.9	208.7 ± 98.0	254.4 ± 116.9	217.2 ± 88.3
Regulatory domain - second CBS domain	B	p.L539S	1.2 ± 0.1	1.2 ± 0.1	1.6 ± 0.2	4.1 ± 0.2	N/A	
Wild type CBS			142.7 ± 27.2	530.5 ± 70.0	191.4 ± 4.5	85.5 ± 27.0	498.2 ± 204.4	N/A
Median of all mutations (n=27)			4.9	6.4	2.7	13		
Median of buried mutations (n=13)			0.6	0.6	0.5	0.8		
Median of solvent exposed mutations (n=14)			13.5	34.6	17.5	22.9		

Table 2. Activities of mutant enzymes in non-particulate water-soluble fractions of bacterial extracts. Each mutant was expressed at 37°C and 18°C in three and two independent experiments, respectively. The numbers in the table represent means and standard deviations of activities determined by LC-MS/MS method with the exception of three bottom lines, which show medians of groups. The effect of AdoMet and AdoHcy on activities of mutants expressed at 37°C was determined in at least 2 experiments, effect on mutants expressed at 18°C was examined only in selected cases.

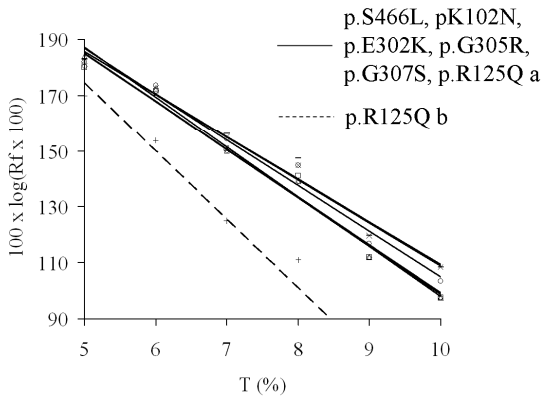
n.d., not detected (activity lower than the limit of quantifications, i.e. 0.3 nmol cystathionine/mg protein/hour); B, buried mutations; S, solvent exposed mutations; N/A, not analyzed

A Analysis of wild type CBS



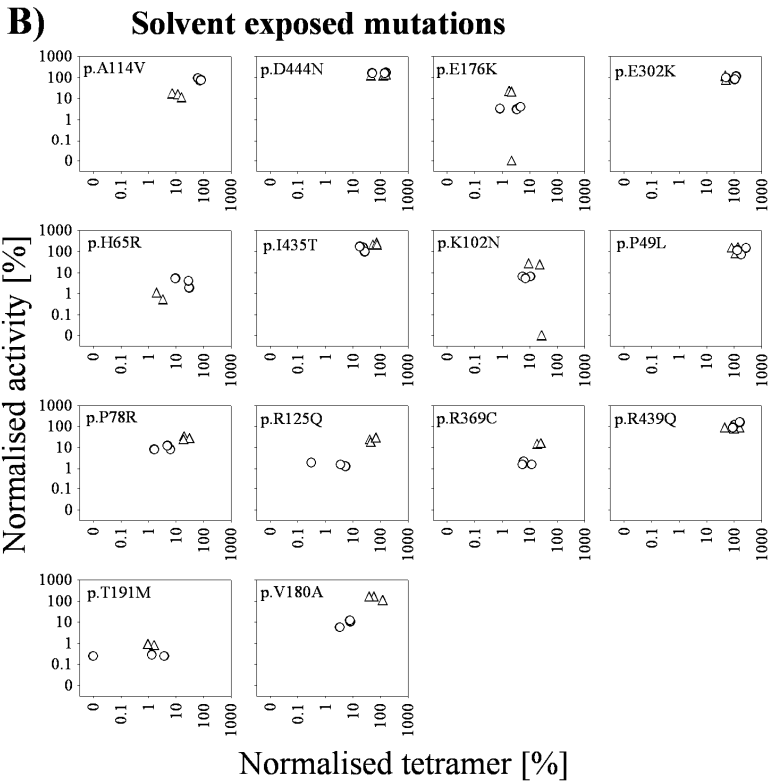
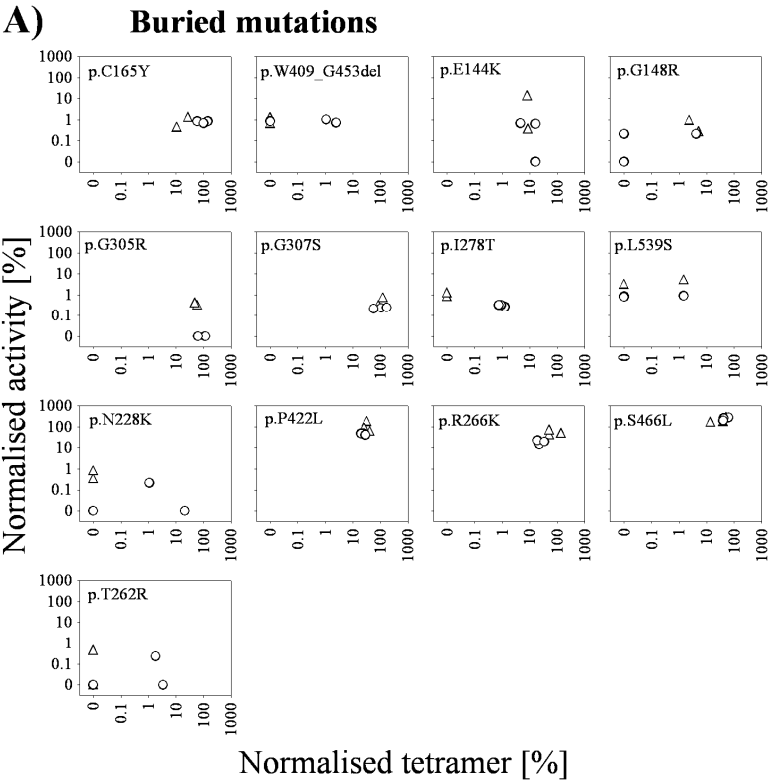
WT fractions	Negative slope	Mr	Oligomeric structure
WT a (fastest)	16.08	177	tetramer
WT b	25.39	398	octamer
WT c	29.36	512	dodecamer
WT d (slowest)	36.03	724	hexadecamer

B Analysis of mutants



Mutant fractions	Negative slope	Mr	Deduced oligomeric structure
p.S466L (fast)	15.383	166	tetramer
p.K102N (fast)	15.249	162	tetramer
p.E302K (slow)	17.171	199	tetramer
p.G305R (slow)	17.263	199	tetramer
p.G307S (slow)	17.849	213	tetramer
p.R125Q a (fast)	16.377	182	tetramer
p.R125Q b (slow)	24.509	372	octamer

Supplementary Figure S1. Ferguson plot. The slopes of regression lines obtained by plotting the relative mobilities (R_f) of markers against the gel concentrations in logarithmic scale were used to calculate the molecular weight of CBS fractions. The tables on the bottom show the values of calculated molecular weights and the deduced oligomeric structure of the four fractions of the wild type CBS enzyme (Panel A) and of six selected mutants (Panel B).



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary Figure S2. Relation of activity and amounts of tetramer. For each mutant data on activity in relation to the amount of tetramers/oligomers from five independent expression experiments are plotted; O and Δ indicate expression at 37°C and 18°C, respectively. All values are normalized to the amounts of tetramers and activities of the wild type CBS, data are shown on a logarithmic scale. Panel A, buried mutations; panel B, solvent exposed mutations.

For Peer Review

Mutation	Primer	
p.H49L	S	5'- gAT CCg gCC CgA TgC TCT gAg CAg gTg CAC - 3'
	AS	5'- gAg CAT Cgg gCC ggA TCC ACA ggg gCT CCT - 3'
p.H65R	S	5'- TCA CCA CAC TgC CCC ggC AAA ATC TCC AAA - 3'
	AS	5'- gCC ggg gCA gTg Tgg TgA CgT ggg gAC TCg gA - 3'
p.P78R	S	5'- AAT CTC CAA AAA TCT TgC gAg ATA TTC TgA - 3'
	AS	5'- gCA AgA TTT TTg gAg ATT TTg CCg ggg CAg - 3'
p.K102N	S	5'- gAA gAA gTT Cgg CCT gAA CTg TgA gCT CTT - 3'
	AS	5'- TTC Agg CCg AAC TTC TTC CCA ATC TTg TTg - 3'
p.R125Q	S	5'- gAT gAT TgA ggA TgC TgA gCg CgA Cgg gAC - 3'
	AS	5'- TCA gCA TCC TCA ATC ATC TgC Agg CTg ATg - 3'
p.E144K	S	5'- gCC Cgg ggA CAC gAT TAT CAA gCC gAC ATC Cg - 3'
	AS	5'- gAT AAT CgT gTC CCC ggg CTT CAg CgT CCC - 3'
p.C165Y	S	5'- AgT gAg ggg CTA TCg CTA CAT CAT CgT gAT - 3'
	AS	5'- AgC gAT AgC CCC TCA CTg CCg CAg CCA ggg - 3'
p.V180A	S	5'- CTC CgA gAA ggT ggA CgC gCT gCg ggC ACT - 3'
	AS	5'- CgT CCA CCT TCT Cgg AgC TCA TCT TCT CTg - 3'
p.T191M	S	5'- gCC CAC CAA TgC CAg gTT CgA CTC CCC ggA - 3'
	AS	5'- CgA ACC Tgg CAT Tgg Tgg gCA TCC TCA CAA TC - 3'
p.N228K	S	5'- TAC CgC AAC gCC AgC AAg CCC CTg gCT CAC - 3'
	AS	5'- TTg CTg gCg TTg Cgg TAC Tgg TCT Agg ATg - 3'
p.T262R	S	5'- CAC ggg Cgg CAC CAT CAg ggg CAT TgC CAg - 3'
	AS	5'- TgA Tgg TgC CgC CCg TgC CCA CTg AAg CCA - 3'
p.R266K	S	5'- gAA gCT gAA ggA gAA gTg TCC Tgg ATg CAg - 3'
	AS	5'- ACT TCT CCT TCA gCT TCT Tgg CAA TgC CCg - 3'
p.I278T	S	5'-TgT CCT ggA TgC Agg ATC ACT ggg gTg gAT - 3'
	AS	5'-TgA TCC TgC ATC CAg gAC ACT TCT CCT TCA - 3'
p.E302K	S	5'- Cgg AgC AgA CAA CCT ACA Agg Tgg AAg ggA - 3'
	AS	5'- gTA ggT TgT CTg CTC CgT CTg gTT CAg CTC - 3'
p.G305R	S	5'- CAA CCT ACg Agg Tgg AAA ggA TCg gCT ACg - 3'
	AS	5'- TTC CAC CTC gTA ggT TgT CTg CTC CgT CTg - 3'
p.G307S	S	5'- CTA CgA ggT ggA Agg gAT CAg CTA CgA CTT CA - 3'
	AS	5'- gAT CCC TTC CAC CTC gTA ggT TgT CTg CTC Cg - 3'
p.R369C	S	5'- gCT gCg Tgg TCA TTC TgC CCg ACT CAg TgC - 3'
	AS	5'- gCA gAA TgA CCA CgC AgC ACT ggC CCT CCT - 3'
p.R439Q	S	5'- ACA CCA TCg AgA TCC TCC Agg AgA Agg gCT TC - 3'
	AS	5'- ggA ggA TCT CgA Tgg TgT gCC CAC Agg Tg - 3'
p.D444N	S	5'- TCC ggg AgA Agg gCT TCA ACC Agg CgC CCg - 3'
	AS	5'- gAA gCC CTT CTC CCg gAg gAT CTC gAT ggT - 3'
p.L539S	S	5'- ggT CAC CgC CAT TgA CTC gCT gAA CTT CgT - 3'
	AS	5'- AgT CAA Tgg Cgg TgA CCA CCC CgA ACA CCA - 3'

Supplementary Table S1. Mutagenic primers. Sequences of mutagenic primers used to introduce the appropriate mutation into the plasmid pHCS3 which carries the wild type CBS sequence.

S, sense primer; AS, antisense primer

	Solvent exposure	Mutation	Temperature of expression 37°C			
			SDS-soluble CBS antigen in fractions		Water-soluble CBS antigen in supernatant	
			Particulate fraction of the total antigen of each mutant	Total antigen of mutants relative to wild type	Tetrameric fraction of the total mutant antigen	Tetramers of mutants relative to tetramer of wild type
Active site	B	p.G148R	8.8 ± 2.7	52.8 ± 29.5	3.0 ± 4.3	1.4 ± 2.0
	B	p.G305R	7.7 ± 3.5	65.8 ± 9.1	83.4 ± 8.3	98.0 ± 24.0
	B	p.G307S	21.2 ± 2.3	67.9 ± 14.3	66.7 ± 9.3	111.7 ± 46.1
Heme binding pocket	S	p.H65R	17.7 ± 1.0	91.2 ± 12.5	31.4 ± 16.6	23.1 ± 9.5
	B	p.T262R	16.0 ± 5.4	49.6 ± 3.8	2.4 ± 2.3	1.7 ± 1.4
	B	p.R266K	21.1 ± 8.7	85.7 ± 9.3	38.4 ± 15.6	24.7 ± 6.3
Other locations in active core	B	p.E144K	49.6 ± 3.9	255.2 ± 165.6	4.9 ± 1.4	12.3 ± 5.4
	B	p.C165Y	28.8 ± 4.7	75.4 ± 41.0	68.0 ± 9.7	101.2 ± 34.7
	B	p.N228K	9.2 ± 3.6	37.5 ± 21.9	13.6 ± 16.7	7.4 ± 9.8
	B	p.I278T	28.2 ± 2.0	74.8 ± 23.9	3.5 ± 1.1	1.0 ± 0.2
	S	p.P49L	10.8 ± 0.3	144.9 ± 75.1	88.0 ± 1.9	194.6 ± 60.1
	S	p.K102N	17.6 ± 8.1	39.2 ± 4.1	29.4 ± 5.6	7.7 ± 2.3
	S	p.R125Q	20.5 ± 6.1	55.5 ± 22.3	7.2 ± 5.4	3.1 ± 2.1
	S	p.T191M	36.5 ± 10.4	35.6 ± 4.0	2.9 ± 2.2	1.7 ± 1.5
	S	p.E302K	11.6 ± 1.3	88.3 ± 0.1	71.2 ± 17.9	90.7 ± 27.8
Dimer-dimer interface	S	p.R369C	27.7 ± 11.5	65.8 ± 12.2	15.7 ± 6.7	7.8 ± 2.9
	S	p.P78R	17.8 ± 0.6	57.6 ± 22.3	41.3 ± 41.9	4.3 ± 1.9
	S	p.A114V	12.7 ± 0.4	111.1 ± 20.2	85.2 ± 3.8	73.7 ± 8.5
	S	p.E176K	15.0 ± 7.6	54.7 ± 3.1	17.7 ± 13.0	3.0 ± 1.6
	S	p.V180A	26.1 ± 1.9	54.0 ± 3.7	20.1 ± 12.7	6.6 ± 2.3

Connection loop between active core and regulatory domain	B	p.W409-G453del	23.2 ± 4.6	43.7 ± 21.7	4.3 ± 3.2	1.2 ± 1.0
Regulatory domain - first CBS domain	B	p.P422L	26.1 ± 11.1	76.1 ± 6.0	61.9 ± 17.5	24.5 ± 3.6
	S	p.I435T	20.9 ± 11.8	64.6 ± 0.4	55.2 ± 23.0	23.2 ± 4.1
	S	p.R439Q	10.7 ± 8.1	322.1 ± 59.6	91.6 ± 7.7	117.6 ± 29.8
	S	p.D444N	5.1 ± 3.0	174.6 ± 16.5	88.7 ± 5.8	120.2 ± 49.7
	B	p.S466L	6.4 ± 6.4	100.0 ± 49.9	67.2 ± 21.0	47.8 ± 10.2
Regulatory domain - second CBS domain	B	p.L539S	5.1 ± 1.9	36.6 ± 11.6	1.0 ± 1.5	0.5 ± 0.7
Wild type CBS			12.4 ± 11.1	100.0	88.0 ± 10.3	100.0
Median of all mutations (n=27)			17.7	65.8	31.4	12.3
Median of buried mutations (n=13)			21.1	67.9	13.6	12.3
Median of solvent exposed mutations (n=14)			17.7	65.2	36.4	15.5

Supplementary Table S2. Amounts of mutant CBS antigen in fractions of bacterial extracts expressed at 37°C. The SDS-soluble antigen was determined by boiling the particulate and non-particulate fractions of bacterial extracts in 3% SDS and by SDS-PAGE followed by western blotting; the signal was quantified by chemiluminescence; the percentual proportion of antigen in the pellets is shown in the first column, total CBS antigen normalized to the wild type CBS is given in the second column. The ability of mutants to fold and assemble was examined by electrophoresis of water soluble non-particulate fractions under non-denaturing conditions and western blotting; percentual proportion of tetramers of the total signal is shown in the third column, yield of tetramers/oligomers in comparison to the amount of wild type tetramers is shown in the fourth column. All values with the exception of last 3 lines are means and standard deviations from 2-3 determinations.

	Solvent exposure	Mutation	Temperature of expression 18°C			
			SDS-soluble CBS antigen in fractions		Water-soluble CBS antigen in supernatant	
			Particulate fraction of the total antigen of each mutant	Total antigen of mutants relative to wild type	Tetrameric fraction of the total mutant antigen	Tetramers of mutants relative to tetramer of wild type
Active site	B	p.G148R	17.3 ± 14.4	23.3 ± 8.4	94.0 ± 8.4	3.2 ± 1.5
	B	p.G305R	9.4 ± 0.8	100.6 ± 31.7	68.4 ± 18.8	51.6 ± 7.2
	B	p.G307S	9.8 ± 3.9	189.1 ± 20.7	74.2 ± 9.3	97.5 ± 17.7
Heme binding pocket	S	p.H65R	38.0 ± 12.5	45.0 ± 7.0	12.7 ± 9.0	1.8 ± 1.4
	B	p.T262R	14.5 ± 8.7	28.9 ± 15.0	n.d.	n.d.
	B	p.R266K	27.0 ± 16.9	117.8 ± 47.5	90.6 ± 8.0	81.0 ± 40.6
Other locations in active core	B	p.E144K	22.5 ± 6.6	39.4 ± 28.0	41.2 ± 42.7	5.7 ± 4.0
	B	p.C165Y	6.6 ± 1.9	63.4 ± 8.0	56.6 ± 12.2	17.3 ± 7.0
	B	p.N228K	26.7 ± 7.9	59.0 ± 31.9	n.d.	n.d.
	B	p.I278T	43.6 ± 9.7	49.1 ± 47.0	n.d.	n.d.
	S	p.P49L	6.1 ± 2.3	92.0 ± 58.0	91.1 ± 2.5	112.0 ± 23.0
	S	p.K102N	23.1 ± 9.5	61.8 ± 12.9	64.1 ± 14.6	20.2 ± 7.8
	S	p.R125Q	18.0 ± 12.3	71.0 ± 1.0	84.1 ± 5.7	53.4 ± 13.0
	S	p.T191M	31.6 ± 14.3	23.9 ± 5.6	9.4 ± 4.6	1.1 ± 0.4
	S	p.E302K	15.8 ± 7.3	113.8 ± 68.5	74.0 ± 6.4	48.8 ± 1.8
Dimer-dimer interface	S	p.R369C	27.8 ± 8.4	58.5 ± 20.2	75.8 ± 8.0	23.3 ± 3.2
	S	p.P78R	4.7 ± 0.8	77.5 ± 21.6	71.6 ± 6.0	23.7 ± 5.9
	S	p.A114V	4.0 ± 1.4	56.5 ± 11.0	59.6 ± 17,3	12.0 ± 3.6
	S	p.E176K	4.9 ± 4.9	39.8 ± 7.2	32.7 ± 8.8	2.1 ± 0.2

	S	p.V180A	7.1 ± 0.1	136.2 ± 1.8	84.7 ± 8.9	75.0 ± 36.5
Connection loop between active core and regulatory domain	B	p.W409-G453del	22.8 ± 9.4	36.0 ± 8.6	n.d.	n.d.
Regulatory domain - first CBS domain	B	p.P422L	9.2 ± 8.0	78.5 ± 11.8	67.6 ± 4.0	33.3 ± 6.9
	S	p.I435T	15.1 ± 8.0	96.9 ± 18.6	82.7 ± 4.4	68.1 ± 10.8
	S	p.R439Q	14.6 ± 3.8	174.5 ± 79.1	89.0 ± 2.9	101.6 ± 45.6
	S	p.D444N	7.8 ± 3.3	281.0 ± 180.5	80.1 ± 7.0	113.3 ± 49.0
	B	p.S466L	8.3 ± 8.3	100.4 ± 61.4	74.9 ± 7.8	35.5 ± 15.9
Regulatory domain - second CBS domain	B	p.L539S	5.3 ± 3.0	31.7 ± 23.7	4.5 ± 3.4	0.9 ± 0.7
Wild type CBS			12.7 ± 16.6	100.0	93.1 ± 3.2	100.0
Median of all mutations (n=27)			14.6	63.4	68.4	23.3
Median of buried mutations (n=13)			14.5	59.0	56.6	5.7
Median of solvent exposed mutations (n=14)			14.9	74.3	74.9	36.3

Supplementary Table S3. Amounts of mutant CBS antigen in fractions of bacterial extracts expressed at 18°C. The SDS-soluble antigen was determined by boiling the particulate and non-particulate fractions of bacterial extracts in 3% and by SDS-PAGE followed by western blotting; the signal was quantified by chemiluminescence; the percentual proportion of antigen in the pellets is shown in the first column, total CBS antigen normalized to the wild type CBS is given in the second column. The ability of mutants to fold and assemble was examined by electrophoresis of water soluble non-particulate fractions under non-denaturing conditions and western blotting; percentual proportion of tetramers of the total signal is shown in the third column, yield of tetramers/oligomers in comparison to the amount of wild type tetramers is shown in the fourth column. All values with the exception of last 3 lines are means and standard deviations from 2-3 determinations.

n.d., not detected