CHARLES UNIVERSITY IN PRAGUE
FIRST FACULTY OF MEDICINE

Study program: Biochemistry and Pathobiochemistry

Ing. Aleš Hnízda

MOLECULAR MECHANISMS IN HOMOCYSTINURIA: SPATIAL ARRANGEMENT OF HUMAN CYSTATHIONINE β-SYNTHASE

MOLEKULOVÉ MECHANISMY HOMOCYSTINURIE: PROSTOROVÉ USPOŘÁDÁNÍ LIDSKÉ CYSTATHIONIN β-SYNTHASY

Ph.D. Thesis

Supervisor : doc. MUDr. Viktor Kožich, CSc.

Consultant : prof. RNDr. Milan Kodíček, CSc.

Prague, 2012

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem řádně uvedl a citoval všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu.

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

V Praze, 2.4. 2012 ALEŠ HNÍZDA

Podpis

Identifikační záznam:

HNÍZDA, Aleš. Molekulové mechanismy homocystinurie: prostorové uspořádání cystathionin beta-synthasy. [Molecular mechanisms of homocystinuria: spatial arrangement of human cystathionine beta-synthase]. Praha, 2012, 165 s., Disertační práce (Ph.D.). Univerzita Karlova v Praze, 1.lékařská fakulta, Ůstav dědičných metabolických poruch. Vedoucí práce Kožich, Viktor.

ABSTRACT

Protein misfolding is considered to be the major pathogenic mechanism in homocystinuria due to cystathionine beta-synthase (CBS) deficiency. The aim of this work was to study molecular mechanisms underlying protein misfolding of CBS mutants.

Firstly, we studied spatial arrangement of normal human CBS protein. Using data from differential covalent labeling of solvent-exposed aminoacid residues, we identified interdomain contact area between the catalytic core and the regulatory domain in human CBS, and we subsequently generated the structural model of the full-length CBS. In the next step, we studied evolutionary divergence of CBS protein structures. We performed phylogenetic analysis that revealed unique spatial arrangement of CBS enzyme in nematodes; the domain architecture of CBS in *Caenorhabditis elegans* was studied experimentally in more detail. Finally, we determined conformational properties of a representative set of human CBS mutants that exhibited in various extent affected formation of tetramers and decreased catalytic activity. Using thermolysin-based proteolytic techniques for analysis of nine mutants expressed in *E.coli*, we found that an unfolded structure is a common intermediate occurring in CBS misfolding. The importance of protein unfolding for pathogenesis of CBS deficiency was further shown by analysis of additional nine purified mutants that were properly assembled into tetramers and possessed normal catalytic activity. These data demonstrated that the altered protein unfolding is a reliable marker of pathogenicity of CBS mutants and proteolysis with thermolysin under native conditions may be an important tool for biochemical assessment of pathogenic variants.

This study advances our understanding of molecular pathology in CBS deficiency and provides knowledge that forms a base for development of chaperone therapy and future improvement in patient care.

Keywords: cystathionine beta-synthase, protein misfolding, surface mapping, protein unfolding, thermolysin, conformational analysis, enzymopathy, homocystinuria

ABSTRAKT

Chybné sbalování proteinů je považováno za hlavní patogenetický mechanismus homocystinurie z deficitu cystathionin beta-synthasy (CBS). Cílem této práce bylo studium molekulových mechanismů, které vedou k chybnému sbalování mutantních forem CBS.

V první části práce jsme studovali prostorové uspořádání normální lidské CBS. Pomocí diferenčního kovalentního značení povrchově dostupných aminokyselinových zbytků jsme identifikovali kontaktní plochu mezi katalytickým jádrem a regulační doménou v lidské CBS a následně jsme navrhli strukturní model plnodélkového enzymu. V další části práce jsme studovali evoluční divergenci proteinových struktur CBS. Provedli jsme fylogenetickou analýzu, která odhalila unikátní uspořádání pro CBS z třídy nematod; doménová architektura CBS z *Caenorhabditis elegans* byla podrobně studována experimentálně. Nakonec jsme studovali konformační vlastnosti vybraných mutantních forem lidské CBS, které měly do různé míry narušenou tvorbu tetrameru a sníženou enzymovou aktivitu. Pomocí proteolytických technik s využitím thermolysinu jsme analyzovali devět mutantních forem, které byly exprimovány v *E.coli*. Zjistili jsme, že rozbalení struktury je běžným jevem při chybném sbalování mutantních CBS. Důležitost rozbalení proteinů pro patogenesi deficitu CBS byla dále prokázána pomocí analýzy dalších devíti purifikovaných mutantních variant, které disponovaly nenarušenou tetramerní strukturou a normální enzymovou aktivitou. Tato data ukázala, že odlišná míra rozbalení proteinu je spolehlivým ukazatalem patogenity mutantních CBS a proteolýza thermolysinem za nativních podmínek může být důležitým nástrojem pro biochemické vyhodnocení patogenních variant.

Tato práce zvyšuje porozumění patogenních mechanismů deficitu CBS a poskytuje znalosti, které mohou být využity pro vývoj chaperonové terapie a následně kvalitnější péči o pacienty.

Klíčová slova: cystathionin beta-synthasa, chybné sbalování proteinů, povrchové mapování, rozbalení proteinu, thermolysin, konformační analýza, enzymopatie, homocystinurie

Dedicated to prof. Milan Elleder
(† 25.9. 2011, RIP)

Věnováno prof. Milanu Ellederovi
(† 25.9. 2011, Odpočívejte v pokoji)

ACKNOWLEDGEMENTS

I would like to thank my entire family and especially my wife Michaela who always supported me at my work; without this support I would have given this study up.

I also want to express gratitude to my colleagues from the lab. My special thanks belong to MSc. Roman Vozdek for fruitful discussions, for his friendship and for his neverending optimistic mood, to Ms. Kateřina Raková for her excelent technical assistance and to Dr. Jakub Krijt for his useful advises. I would like to thank Dr. Miroslav Janošík and MSc. Jana Kopecká for their help during my first steps in the lab. Last but not least, I thank my supervisor Assoc. Prof. Viktor Kožich for his supervision, support and revision of the publications including this thesis.

In addition, my thanks belong to my co-workers from other institutions, namely prof. Milan Kodíček, Assoc. Prof. Vojtěch Spiwok and MSc. Vojtěch Jurga from Institute of Chemical Technology in Prague, prof. Jan P. Kraus and Dr. Tomáš Majtan from University of Colorado at Denver. They helped me to conduct and successfully finish our projects.

This study was supported by Welcome Trust International Senior Research Fellowship in Biomedical Science in Central Europe (reg. No 070255/Z/03/Z), by the Research Project from Ministry of Education of the Czech Republic No. MSM0021620806, by the research program of the Charles University in Prague PRVOUK-P24/LF1/3, by the grants of Grant Agency of Charles University in Prague No. 7310 and No. 2170, by the project 303/03/H065 of Czech Science Foundantion and by the grants SVV262502 and SVV260501 of Charles University in Prague. This financial support is kindly acknowledged.

CONTENT

1. INTRODUCTION.....13

1.1 Metabolism of sulfur amino acids……..……...……..13

1.2 CBS deficiency…….....……...15

1.3 Properties of CBS protein……...……. 16

1.4 Location of pathogenic mutations in the CBS protein……...21

1.5 Protein misfolding and genetic diseases…….…… 24

1.6 Rescue of misfolded CBS mutants by chaperone treatment…….……...…... 25

2. AIMS OF THE STUDY……...……...…….. 27

3. RESULTS AND DISCUSSION……...…….…28

3.1 Surface mapping of CBS…….……....…….…28

3.2 Characterization of CBS in *Caenorhabditis elegans*…….…..…….…..….33

3.3 Structural characterization of CBS mutants……..…....…....35

3.4 Analysis of CBS protein in human plasma…….....38

4. CONCLUSIONS……...……....…..…39

5. AUTHOR´S PUBLICATIONS, PRESENTATIONS AND GRANTS……...…...….41

6. ABBREVIATIONS……..……...46

7. REFERENCES…….....……..…47

SUPPLEMENT……...……....……...……...……55

Publication 5.1.1: Hnízda et al.: Journal of Biochemical and Biophysical Methods 2008…..57

Publication 5.1.2: Hnízda et al.: Biochemistry 2010…….……...…....65

Publication 5.1.3: Vozdek et al.: Biochemical Journal 2012…….……...……91

Publication 5.1.4: Hnízda et al: Journal of Inherited Metabolic Disease (On-line article)….111

Manuscript 5.1.5: Hnízda et al.: Unpublished manuscript……...…...…....131

Publication 5.1.6: Krijt et al.: Journal of Inherited Metabolic Disease 2011……..….....157

1. INTRODUCTION

1.1 Metabolism of sulfur amino acids

Sulfur amino acids and related metabolites play an important role in many biological processes, such as protein synthesis, methylation reactions, signal transduction and protection against oxidative stress. In biological systems, there are two genetically coded amino acids containing sulfur, namely methionine and cysteine.

1.1.1 Methionine cycle

Methionine is an essential amino acid that is converted in three steps to homocysteine (Fig.1). In the first step, methionine is activated by condensation with ATP that is catalyzed by ATP:L-methionine S-adenosyltransferase to form S-adenosylmethionine (AdoMet) (*1*). AdoMet serves as a donor of methyl group in many transmethylation reactions involving AdoMet-dependent methyltransferases to yield methylated products and S-adenosyl-homocysteine (AdoHCy); AdoHCy is further hydrolyzed by S-adenosylhomocysteine hydrolase to form homocysteine and adenosine. Homocysteine represents a branchpoint intermediate in the methionine metabolism; this metabolite may be remethylated back to methionine or may be converted via the transsulfuration pathway to cysteine.

The remethylation is mediated by two different mechanisms. The first enzyme system employs methionine synthase (*2*). The donor of methyl group is methylcobalamin which is regenerated by 5-methyltetrahydrofolate, formed from 5,10-methylentetrahydrofolate by methylentetrahydrofolate reductase. The second remethylation mechanism, catalysed by betaine:homocysteine methyltransferase uses another compound, betaine, as a methyl donor (*3*).

1.1.2 Transsulfuration pathway

Transsulfuration pathway is an irreversible sequence of reactions which leads to formation of cysteine from homocysteine. The first step of the transsulfuration pathway, i.e. condensation of homocysteine with serine to form cystathionine, is catalysed by cystathionine beta-synthase (CBS). Cystathionine is subsequently cleaved by cystathionine gamma-lyase to yield cysteine and α-ketobutyrate. Cysteine may be incorporated into proteins or glutathionine and may be also converted to taurine or inorganic sulfur.

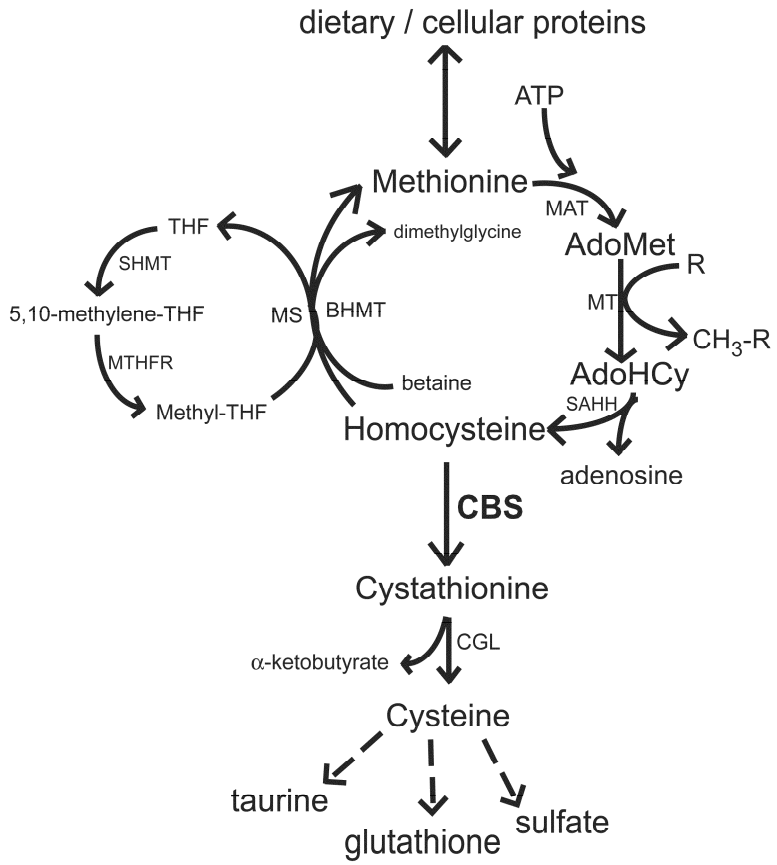


Figure 1 Metabolism of methionine: enzymes catalyzing the reactions are depicted by arrows: MAT – ATP:L-methionine S-adenosyltransferase; MT – methyltransferase; SAHH – S-adenosylhomocysteine hydrolase; BHMT – betaine:homocysteine methyltransferase; MS – methionine synthase; SHMT – serine:hydroxymethyl transferase; MTHFR – methyltetrahydrofolate reductase; CBS – cystathionine beta-synthase; CGL – cystathionine gamma-lyase. Metabolites are abbreviated as follows: ATP – adenosine-5´-triphosphate; THF – tetrahydrofolate; AdoMet – S-adenosylmethionine; AdoHcy – S-adenosylhomocysteine.

1.1.1 Regulation of methionine metabolism

The majority of tissues regulate the level of intermediates of methionine metabolism by folate-dependent remethylation and by export of the excess of homocysteine from cells (*4*). The liver represents the main organ for maintaining methionine homeostasis since hepatocytes possess the enzymes of the two remethylation systems and of the transsulfuration pathway.

Methionine metabolism is regulated by the distribution of homocysteine between remethylation and transsulfuration (*5*). Intrinsic properties of the enzymes, such as their kinetic parameters and allosteric responses to the effectors, constitute a regulatory mechanism

that reacts rapidly to perturbations in methionine metabolism. Based on different K_m values of the metabolic enzymes, it has been suggested that remethylation pathway is favored under lower concentrations of homocysteine while a more intensive flux through the transsulfuration pathway is prefered when homocysteine concentration is increased (*5*). The catalytic activity of enzymes involved in methionine metabolism is further modulated by AdoMet. Increased concentrations of AdoMet stimulate catalytic activity of CBS and, on the contrary, inhibit both remethylation systems (*6*).

Methionine metabolism may be also regulated at the level of protein expression of the respective enzymes. It was reported that intake of dietary protein increases expression of CBS, cystathionine gamma-lyase, S-adenosylhomocysteine hydrolase and betaine:homocysteine methyltransferase and decreases the level of methionine synthase (*7*). The amounts of enzymes involved in methionine metabolism may be also modulated by endocrine activity, age and other factors.

1.2 CBS deficiency

For each of the above described enzymatic steps deficiencies are known in humans. Inherited diseases were reported in the conversion of methionine to homocysteine (deficiency of ATP:L-methionine S-adenosyltransferase, glycine N-methyltransferase and S-adenosylhomocysteine hydrolase), of the remethylation system (methylenetetrahydrofolate reductase deficiency or functional deficiency of methionine synthase) and of the transsulfuration pathway (CBS or cystathionine gamma-lyase deficiency). The most common inborn error of methionine metabolism is homocystinuria due to CBS deficiency, an autosomal recessive disease with the worldwide prevalence of 1:344,000 (*8*). However, the true frequency may be underestimated since several molecular epidemiological studies suggested that the incidence may be around 1:10,000 (*9-11*).

CBS deficiency is a multisystemic disease that manifests by lens dislocation, early tromboembolic events, mental retardation and skeletal abnormalities such as osteoporosis and marfanoid features (*8*). Biochemical symptoms in CBS deficient patients include elevated concentration of total homocysteine, decreased concentrations of cysteine and cystathionine, and varying elevation of methionine (*12*). Diagnosis may be confirmed by assaying CBS activity in patient-derived fibroblast cell lines or by DNA analysis (*13*).

About half of patients respond to administration of pharmacological doses of pyridoxine (vitamin B_6), with significant biochemical and clinical improvement. Partially B_6 responsive and unresponsive patients are further treated by methionine restriction and cysteine supplementation. Some patients may also need a treatment with high doses of betaine that enhances the remethylation of homocysteine (*8, 13*).

More than 150 mutant alleles have been described in CBS-deficient patients (annotated in CBS Mutation Database; http://cbs.lf1.cuni.cz/cbsdata/cbsmain.htm) many of the mutations are private. Nevertheless, several variants are more frequent, often confined to specific ethnic groups (*14*). Missense mutations represent 87 % of all analyzed alleles indicating that structural alternations due to aminoacid substitutions in the CBS protein play an important role in the pathogenesis of CBS deficiency.

1.3 Properties of CBS protein

CBS is a pyridoxal-5'-phosphate (PLP) - dependent enzyme which catalyses the first step in the transsulfuration pathway. Human CBS is a homotetrameric protein (*15*) that in its purified form is prone to formation of higher-order oligomers and aggregates (*16, 17*). CBS performs its function mainly in cytosol while a minor portion of the protein pool may be modified by small ubiquitin-like modifier-1 protein and subsequently transported to the nucleus (*18*).

1.3.1 Domain architecture of CBS

Each subunit of this modular protein consists of 551 amino acids (61 kDa); the sequence comprises three regions: the N-terminal heme-binding domain (residues 1-69), an evolutionary conserved catalytic core (residues 70-413) and the C-terminal regulatory domain (residues 414-551) (*19*).

1.3.1.1 Heme-binding domain

The heme moiety is bound to the enzyme via C52 and H65 residues coordinating the iron atom (*20*) and via H67 and R266 residues that form second coordination sphere of the heme-binding pocket (*21*). The function of heme in CBS is still unclear. Banerjee´s group proposed that heme serves as a redox sensor that modulates catalytic activity of CBS to enhance production of glutathione under oxidative conditions (*22, 23*). This suggestion was recently

supported by a study showing that reversible reduction of heme to ferrous state and subsequent inhibition of CBS activity may be mediated by physiologically relevant reducing partner such as flavin-dependent methionine synthase reductase (*24*). On the contrary, a study by Majtan et al. revealed that heme may play a structural role in CBS protein since the heme moiety can be replaced by several metal-substituted porfyrins unresponsive to changes in redox potential (*25*).

1.3.1.2 CBS catalytic core

The highly conserved catalytic core belongs to the β-family of PLP-dependent enzymes such as O-acetylserine sulfhydrylase or tryptofan synthase. PLP, a cofactor that is necessary for catalysis, is bound to the enzyme via K119 (*26*). The mechanism of reaction catalyzed by CBS was described in details for the yeast enzyme using stopped-flow kinetics (*27, 28*). The yeast CBS was feasible for this study since it lacks heme that otherwise interferes in spectral analyses with the PLP-bound intermediates. The reaction proceeds by ping-pong mechanism. Initially, serine is bound to PLP forming an external aldimine that is converted to an aminoacrylate intermediate. Subsequently, homocysteine is attached to form an external aldimine of cystathionine. The last action, i.e. the release of cystathionine from the enzyme, is the slowest process during catalysis and thus the rate-limiting step (Fig.2).

The catalytic activity of CBS is confined to β-replacement reactions, i.e. the replacement of the electronegative substituent at the β-position in α-aminoacids, and the CBS enzyme is promiscuous towards various substrates. CBS can catalyse the following reactions:

i, serine with homocysteine to produce cystathionine and water;
ii, cysteine with homocysteine to form cystathionine and hydrogen sulfide,
iii, cysteine with water to yield serine and hydrogen sulfide;
iv, O-acetyl serine with hydrogen sulfide forming cysteine and acetate;
v, serine with hydrogen sulfide to produce cysteine and water

Enzyme kinetics for these reactions catalysed by human CBS were reported in details elsewhere (*29, 30*).

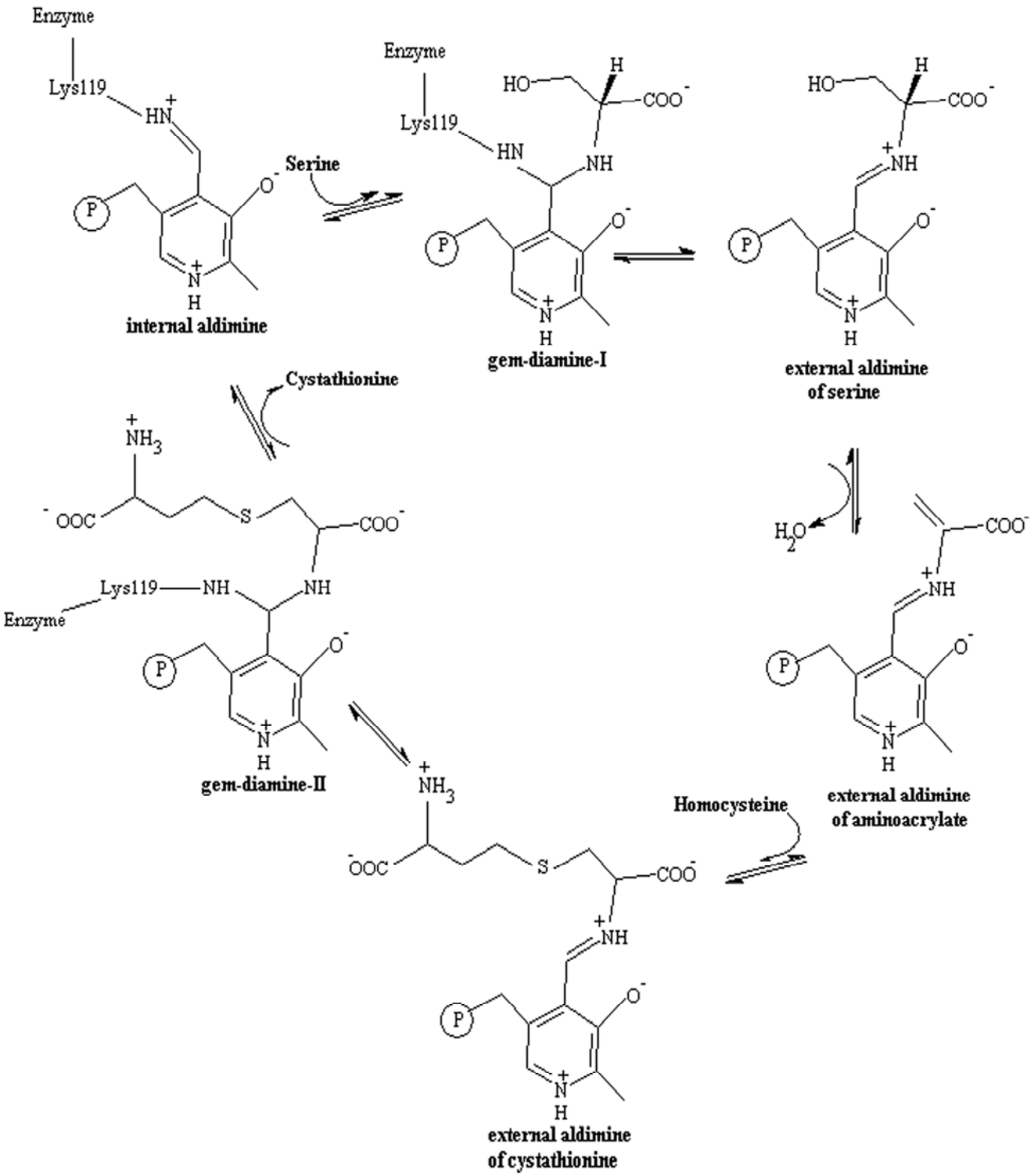


Figure 2 Catalytic mechanism for condensation of serine with homocysteine catalyzed by CBS; the figure was adopted from (*27*).

1.3.1.3 C-terminal regulatory domain

The C-terminal region contains a pair of CBS domains which together form the Bateman domain (*31*). Bateman domains are conserved in many otherwise unrelated proteins (such as chloride channel, adenosyl-monophosphate activated protein kinase or inosine-5´-monophosphate dehydrogenase) and act as sensing and regulatory elements in these proteins (*32*). In human CBS, Bateman domain functions as an autoinhibitory module that contains the binding site for AdoMet, an allosteric activator of the enzyme (*33*). In addition, the C-terminal domain is also responsible for tetramerization of CBS protein (*34*).

1.3.1.4 Evolutionary divergence of CBS domain architecture

The above described domain architecture of CBS is well conserved in mammals but not across the phylla. In addition to the human enzyme, CBS proteins from *Saccharomyces cerevisiae (35), Trypanosoma cruzi* (*36*) and *Drosophila melanogaster (37)* have been studied by biochemical and biophysical approaches. In contrast to its presence in drosophila, the N-terminal heme-binding module is absent in yeast and protozoa. The catalytic domain is conserved in all characterized CBS enzymes whereas the C-terminal part is highly variable. The yeast CBS is not activated by AdoMet; nevertheless, the C-terminal module retains autoinhibitory features and its proteolytic removal leads to increase of the CBS activity. Furthermore, the yeast C-terminal domain is responsible for formation of tetramers and octamers of the CBS protein. The C-terminal region of protozoan CBS is shortened and regulatory function is missing but it is still able to form tetramers. On the contrary, the drosophila CBS does not bind AdoMet and forms only dimers. The variable domain architecture of CBS indicates that modulation of activity of these enzymes has evolved differently in evolutionary distant organisms.

1.3.2 Allosteric regulation of catalytic activity of human CBS

Results of different studies indicate that cross-talk between the catalytic core and the C-terminal autoinhibitory domain plays an important role in modulation of human CBS activity. The activity may be stimulated *in vitro* by several processes: by binding of AdoMet to the Bateman domain (*33*), by proteolytic cleavage yielding the C-terminally truncated dimer (*34*), or by partial heat denaturation (*16, 33*).

It is suggested that allosteric activation of CBS activity upon AdoMet binding plays an important role in maintaining homeostasis of methionine metabolism (*6*). This notion was further supported by studies using cultured cell lines together with mathematical modeling of

methionine metabolism (*38, 39*). In addition, biological relevance of proteolytic activation of CBS was supported by observations in rat liver extracts (*15*) and in HepG cells (*40*).

The spatial arrangement of human CBS was determined by X-ray crystalography for the 45kDa truncated enzyme that lacks the C-terminal regulatory domain (45CBS, amino acids 1-413) and that forms dimers (*20, 23*). These studies have revealed important structural features of CBS proteins at an atomic resolution, such as the active site containing PLP, the heme-binding pocket and dimer interface of the catalytic core. However, the 3-D structure of the full-length CBS has not yet been solved and therefore the structural mechanisms of the cross-talk between the catalytic core and the Bateman domain are still unknown. While extensive hydrophobicity of the C-terminal domain and putative flexible interdomain motions prevented successful crystalization of the full-length CBS, alternative techniques can provide at least partial information about the allostery in the CBS molecule. The Banerjee´s group performed conformational study using H/D exchange demonstrating that the region 356-385 may be involved in the interdomain communication (*41*). These data were used for a protein-protein docking exercise and the first structural model of the full-length CBS was proposed (Fig. 3). However, additional experiments using complementary structural techniques are required to support and/or refine the proposed structure.

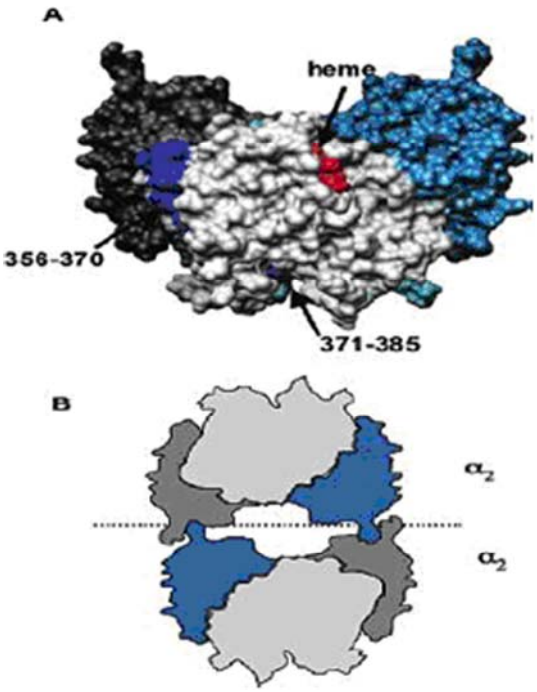


Figure 3 The proposed structural model of the full-length CBS based on data from H/D exchange. A, Peptides exhibiting different kinetics of hydrogen exchange upon C-terminal truncation are pointed by arrows on the structure of 45CBS. B, The scheme represents possible assembly of the full-length CBS into tetramer. Figures are adopted from Sen S. et al. *(41)*.

1.4 Location of pathogenic mutations in the CBS molecule

Disease-causing mutations have been observed in each functional domain of human CBS. The location of aminoacid substitutions in respect to the known structural arrangement of the wild-type CBS enzyme has been reviewed in details previously (*42, 43*). In the next section, knowledge about structure-function relationships of selected mutant proteins will be discussed.

1.4.1 Mutations in the heme-binding domain

Several mutations that directly affect the coordination of the heme moiety in CBS protein have been observed in CBS-deficient patients, namely the H65R, R266K and R266G.

Although mutation H65R directly affects the proximal ligand group coordinating the heme moiety (Fig.4A), the mutant exhibited 40 % heme saturation compared to the wild-type enzyme. This suggests that other aminoacid residues, such as H67, may partially substitute the missing imidazole ring of the H65 (*44*). Despite the partial heme saturation, the mutant enzyme retained only around 20 % of the wild-type activity indicating that the mutation in the heme-binding site disrupts proper catalytic function. The effect of mutation R266K - an aminoacid substitution located in the secondary coordination sphere of heme (Fig.4A) - was also studied in more detail. This mutant exhibited normal (*21*) or slightly lower (*45, 46*) catalytic activity; mild decrease in enzymatic activity was associated with slightly reduced thermal stability of the R266K mutant compared to the wild-type enzyme (*46*). On the other hand, the R266G mutant could not be purified into homogeneity due to substantial destabilization of protein structure that is accompanied with a complete loss of CBS activity (*21*). It suggests that a loss of positive charge in the position 266 affects binding of heme that is important for correct folding of the CBS molecule (*26*).

1.4.2 The most frequent mutations are located in the catalytic core

The T191M mutation is prevalent among patients originating from the Iberian Peninsula (*47*). The mutated residue T191 is solvent-exposed and located at the periphery of the catalytic core (Fig.4B) (*48*). The T191M mutant was found to be virtually inactive and to form aggregates on native electrophoresis (*47*). Based on similarity with O-acetylserine sulfhydrylase from *Salmonella typhyrium*, it was suggested that the region 186-222 is

involved in conformational changes upon PLP binding and that the mutation T191M impairs binding of PLP to the CBS protein (*47*).

The most common pathogenic CBS variant is the I278T mutation (*14, 49*). This aminoacid substitution is buried in the catalytic core and located at a distance from the active site or the heme-binding pocket (Fig.4B) (*42, 48*). The I278T mutant forms high molecular weight aggregates devoid of heme and exhibits very low residual catalytic activity (*50*). Interestingly, the effect of this mutation has been suppressed by the point mutations in the regulatory domain or by complete deletion of the C-terminal region (*51, 52*). The rescue of the mutant protein by C-terminal removal indicates that the residue I278 may be involved in the regulatory motions between the active site and the autoinhibitory domain.

The active site mutation G307S represents another variant frequently found in CBS-deficient patients being highly prevalent in Northern Europe (*53*). The residue G307 is located at the entry to the catatalytic cleft (Fig.4C). The incorporation of the larger serine residue to this position probably causes overpacking since the side-chains of P282, S285 and Y301 are located in the proximity of the G307 in the crystal structure of the C-terminally truncated human CBS protein (*20*) (Fig.4C). It has been suggested that homocysteine binds in the vicinity of G307 and that the aminoacid substitution would thus decrease binding of homocysteine (*42*). The mutant protein G307S is catalytically inactive, with residual activity lower than 1 % of wild-type CBS enzyme activity (*54*) despite unimpaired formation of tetramers (*48*).

1.4.3. Mutations in the C-terminal regulatory domain

Mutations in the C-terminal domain represent an interesting group of disease-causing variants. Patients carrying these mutations suffered from mild CBS deficiency (*55*). Interestingly, the purified C-terminal mutants I435T, S466L (*33*) and D444N (*56*) exhibited increased catalytic activity and failed to be further stimulated by AdoMet since they were locked in a superactive conformation (*33, 41*). It was proposed that the altered response to AdoMet is responsible for the pathogenicity of the C-terminal mutations (*38, 48*). Hovewer, the expression of S466L in a transgenic mouse led to lower steady-state protein levels and to decreased catalytic efficiency *in vivo* (*57*). Decreased amounts of the C-terminal mutants were also observed in cell lysates from the patient-derived fibroblast cell lines harboring the genotype I435T/del ex8 (*55*) and D444N/D444N (*56*). These data indicate that the altered conformation of these “superactive mutants“ may lead to more rapid degradation with resulting decreased amounts of mutant proteins *in vivo*.

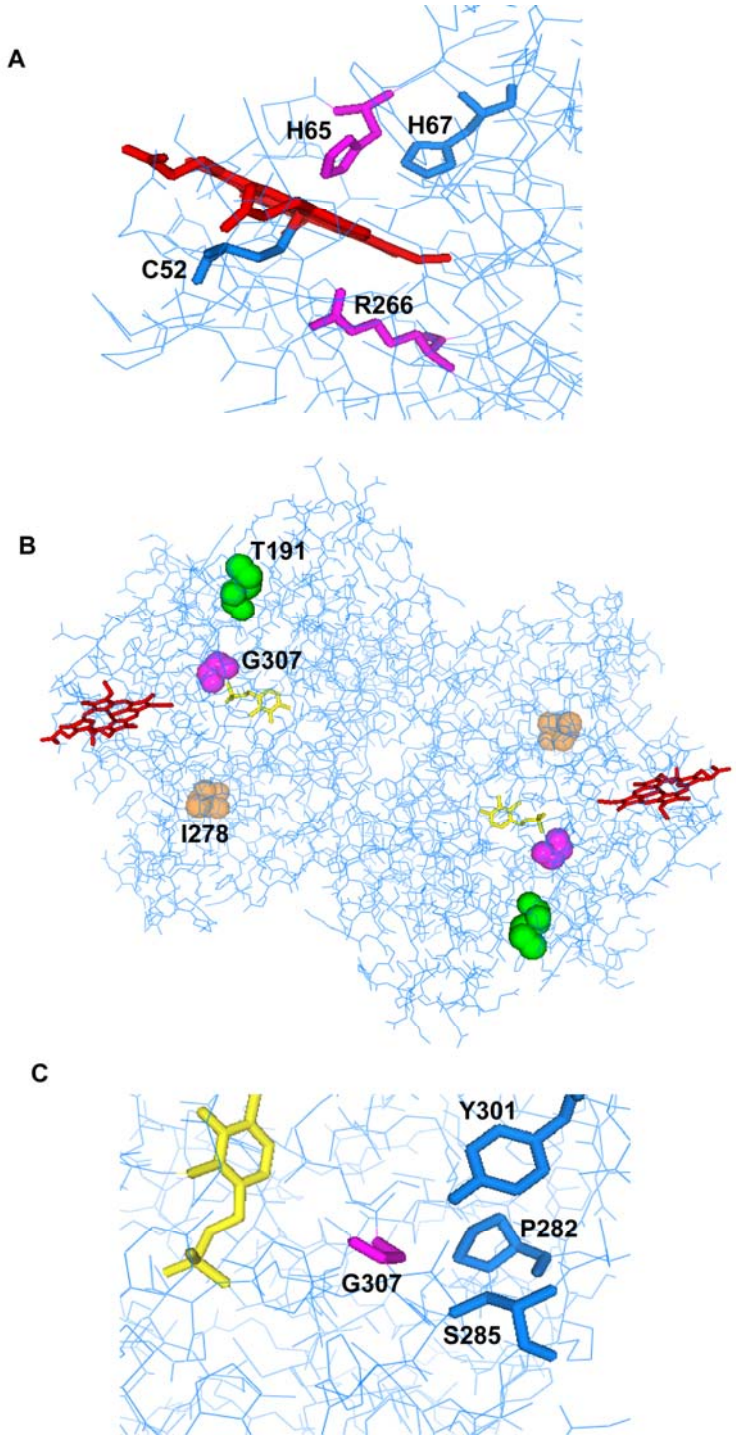


Figure 4 Location of the described mutations in the 3-D structure of the 45CBS (Protein Data Bank (PDB) ID 1JBQ). **A. The heme binding pocket.** The heme moiety is depicted in red, mutated residues (H65 and R266) in magenta, and the heme ligands C52 and H67 are shown as blue sticks.

B. Location of the most frequent mutations in the 45CBS structure (T191M, I278T and G307S). The moieties of heme and PLP are depicted in red and yellow, respectively. The mutated residues are depicted as colored balls. **C. Detail of the active site with PLP and G307** (yellow and magenta sticks, respectively). The residues tightly packed against G307 are marked as blue sticks.

1.5 Protein misfolding and genetic diseases

In genetic diseases, most aminoacid substitutions do not affect key functional residues in respective proteins, but rather result in misfolded proteins with impaired biological function (*58*). This suggestion was strongly supported by *in silico* analysis of 731 proteins implicated in monogenic diseases. This analysis showed that a loss of protein structure stability is the major cause of pathogenicity of missense mutations (*59*).

Misfolded proteins are recognized by the protein-quality control systems and in many cases are rapidly degraded; consequently, the amounts of mutant proteins are decreased exemplifying the loss-of-function pathogenesis. The pathogenic mechanisms underlying protein misfolding in the field of inborn errors of metabolism have been thoroughly described for the most frequent enzymopathy - phenylketonuria due to phenylalanine hydroxylase deficiency. Protein misfolding as a major pathogenic mechanism of this metabolic disease was demonstrated by meta-analysis of experimental data in connection with computational modeling of 318 mutants; it was postulated that the energetics of protein folding is a more reliable determinant of the pathogenicity of the mutant proteins than their *in vitro* residual activity (*60*). Loss-of-function pathogenesis as a consequence of misfolding of phenylalanine hydroxylase was shown by observing more rapid degradation of seven pathogenic variants in rabbit reticulocyte extracts (*61*) and human embryonic kidney cell lines (*62*).

In another group of diseases, an aberrant protein folding may result in a production of biomolecules that are highly resistant against proteolysis and that form aggregates. In these cases, pathogenesis of protein misfolding is not only due to loss-of-function but also due to gain-of-function (*63*). This pathogenesis is very common in prion diseases and for neurodegenerative diseases such as Alzheimer, Parkinson and Huntington disease (*64*). Clinical manifestation of these disorders is caused by the presence of the neuronal deposits of misfolded protein aggregates.

1.5.1 Protein misfolding as the leading pathogenic mechanism of CBS deficiency

The previous study by Janošík et al. proposed that protein misfolding may be an important pathogenic mechanism also in CBS deficiency since the majority of tested CBS mutants, i.e. A114V, A155T, E176K and I278T, formed large inactive aggregates devoid of heme (*50*). This proposal was strongly supported by a more extensive study of 27 mutant variants representing 70 % of known allelles observed in the CBS-deficient patients (*48*). Protein

misfolding was demonstrated for the majority of the studied mutants by observing abnormal migration on native electrophoresis, an increased abundance of mutants in the insoluble fraction and by their rescue after the expression at a lower temperature facilitating correct folding. It was also shown that buried mutations affect correct folding more severely than solvent-exposed aminoacid substitutions. However, both studies have used the bacterial expression system that permits evaluation of intrinsic structural propensity of the mutant proteins while other consequences such as proteolytic degradation specific to eukaryotic cells cannot be determined by this approach (*62*). Indeed, the functioning of the misfolded CBS mutant I278T was also studied in a eukaryotic expression system, namely in *Saccharomyces cerevisiae*. The mutant I278T was rapidly degraded by the proteasome-dependent proteolysis and the protein turnover was altered by manipulation with the levels of heat shock proteins (*65*). Taken together, studies using prokaryotic and eukaryotic expression systems have shown that CBS deficiency is a conformational disorder with loss-of-function pathogenesis.

1.6 Rescue of misfolded CBS mutants by chaperone treatment

Misfolding of mutants may be corrected by the presence of chaperones during protein expression. The rescue of misfolded proteins is accomplished by treating cells with two types of compounds, i.e. chemical or pharmacological chaperones (*66*). Chemical chaperones refer to small molecular weight compounds, usually osmolytes such as glycerol and trimethylamine-N-oxide, that nonspecifically promote correct folding of proteins. However, nonspecific action and high effective doses of chemical chaperones decrease their therapeutical potential. On the other hand, pharmacological chaperones are usually derived from ligand(s) of a target protein (e.g. cofactor or inhibitor of an enzyme) and act thus highly specifically at low concentrations used (*66*). Generally, a correction of aberrant folding by chaperones represents an interesting alternative for successful therapy of conformational disorders (*67*). Clinical relevance of this strategy in the field of inherited metabolic diseases was demonstrated by the implementation of therapy for phenylketonuria using sapropterin (*68*) which possibly acts as a pharmacological chaperone (*69-71*).

The CBS mutants were successfully rescued by the treatment with chemical chaperones, ligands of CBS enzyme and inhibitors of proteasome (*72-75*) suggesting that this strategy may become a novel therapeutical option in CBS deficiency. Nevertheless, chaperones useful for therapy of CBS deficiency have not yet been found. The efficacy and the specificity of

screening for pharmacological chaperones may be significantly increased when the structural mechanisms underlying misfolding are known.

However, there are two major limitations for more efficient structure-based drug design. Firstly, the 3-D structure of normal full-length CBS that is necessary for understanding of proper protein folding has not yet been determined (as discussed in Section 1.3.2). Secondly, only a few CBS mutants (the C-terminal mutants discussed in Section 1.4.3, a double linked mutant P78R/K102N (*17*) and the R266K mutant (*46*)) have been purified into homogeneity in amounts sufficient for detailed conformational studies to assess molecular mechanisms of their pathogenicity. Six additional mutants (P49L, P78R, A114V, R125Q, E176K and P422L) were purified when the culture media contained chemical chaperones (*74*). Since chaperones facilitate correct folding, these purified mutants did not exhibit gross abnormalities in heme saturation, catalytic activity and tetramer assembly and the structural cause of the pathogenicity of these mutations was not found. Taken together, knowledge about the mechanisms underlying misfolding of the mutant CBS proteins clearly needs to be expanded in order to advance our understanding of CBS deficiency and to improve patient care.

2. AIMS OF THE STUDY

The goal of my Ph.D. project was to study structure-function relationships of various normal and mutant CBS enzymes.

The specific aims were as follows:

Spatial arrangement of the wild-type CBS

Surface mapping of the full-length CBS

1. To develop a methodology feasible for surface mapping of CBS
2. To compare the solvent accessibility of aminoacid residues between the 45CBS and the full-length CBS, to identify regulatory interface between the catalytic core and the regulatory domain, and to build the structural model of the human full-length CBS

Conservation of CBS domain architecture in other species

1. To study the conservation of CBS structure by *in silico* analysis
2. To characterize structural and enzymatic properties of unique CBS from the nematode *Caenorhabditis elegans*

Structural properties of CBS mutant proteins

Conformational analysis of mutant CBS proteins

1. To develop a methodology feasible for conformational study of the mutant CBS proteins directly in crude cell extracts without the need for purification
2. Using this approach, to study structural properties of a set of representative and clinically relevant CBS mutants that are not amenable to purification due to their instability
3. To study conformation of another set of purified CBS mutants and to explore the molecular mechanisms of their pathogenicity

Analysis of mutant CBS proteins in plasma

1. To explore whether CBS enzyme is present in plasma of healthy individuals by the available proteomic, immunological and enzymological approaches
2. To study the properties of CBS enzyme in plasma of CBS deficient patients and to test the utility of this analysis for diagnosis of the CBS deficiency

3. RESULTS AND DISCUSSION

3.1 Surface mapping of CBS

Surface mapping is a structural technique that uses specific covalent modification of the solvent-exposed aminoacid residues in native proteins. The modified residues are identified by peptide mass fingerprint using mass spectrometric detection (Fig.5).

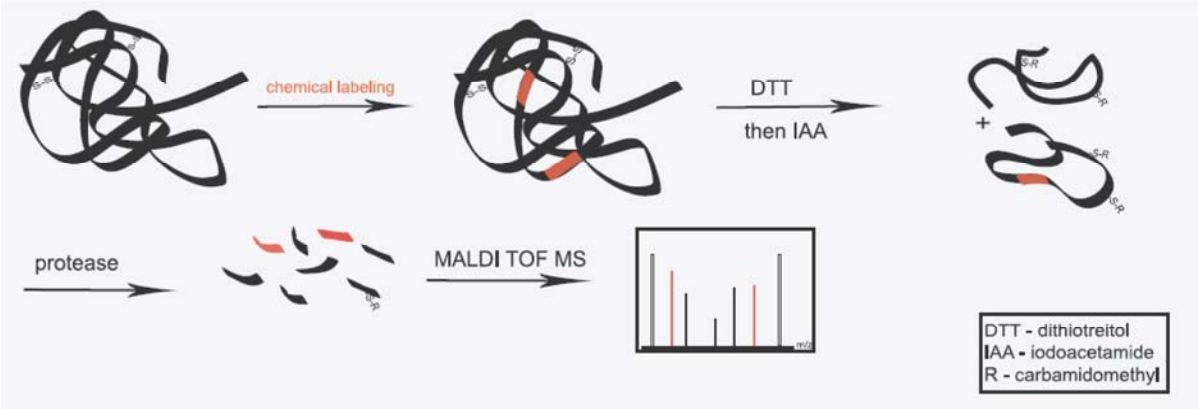


Figure 5: The scheme showing the workflow of protein surface mapping used in our study.

3.1.1 Development of the methodology for surface mapping

We found the appropriate experimental conditions for the modification reactions with nine commonly used modifiers using commercially available proteins such as chicken lysozyme, horse cytochrome c and human serum albumin. Special emphasis was put on modification of proteins with diethylpyrocarbonate (DEP) that is commonly used for modification of histidine residues but its possible reactivity towards other aminoacid residues has been neglected in previous studies. We found that DEP labeled not only the surface accessible histidine but also lysine residues in a set of model proteins. Moreover, our study revealed that the surface accessibility of the residues is a necessary, but not sufficient condition for their reactivity. Other study by Mendozza et al. (*76*) showed that DEP can react also with solvent-exposed residues of tyrosine, threonine and serine. These data show that DEP does not label only histidines and thus the identity of the modified residues should be confirmed by tandem mass spectrometry (MS/MS), i.e. by analysing the fragmentation of the modified peptides.

Feasibility of the modifiers for structural analysis of CBS was tested using the 45CBS. Since an excessive modification of protein may disrupt its native structure, we determined the

lowest concentration of the modifier that enabled sufficient detection of the labeled residues while maintaining the integrity of the modified CBS proteins on native electrophoresis and retaining the catalytic activity. Six modification agents were feasible for further study (Tab.1) since the modified CBS retained high residual activity and formed dimers on native electrophoresis. Three other agents, namely tetranitromethane and iodine (tyrosine modifiers), and 2-hydroxy-5-nitrobenzyl bromide (reactive towards tryptophans) could not be applied for the structural study of CBS since they caused formation of smears on native electrophoresis along with complete inactivation of CBS enzyme indicating severe structural impairment of modified CBS. These observations demonstrate that the assessment of the integrity of modified proteins is an important procedure for acquisition of structurally relevant data using protein surface mapping.

Table 1 Modification agents that were used for surface mapping of CBS.

Modifier	**Reactive amino acid residues**	**Mass shift in modified peptides**
diethylpyrocarbonate	histidine (lysine, cysteine, N-terminal amino acid)	+72; +15
N-bromsuccinimide	tryptophan (methionine)	+16; +32
N-ethylmaleinimide	cysteine	+125
N-acetylimidazole	tyrosine (lysine, N-terminal amino acid)	+42
sulfo-N-hydroxysuccinimido acetate	lysine (N-terminal amino acid)	+42
4-hydroxyphenylglyoxal	arginine	+132

The most reactive amino acids are placed in the first position, other reactive amino acids are shown in parentheses.

For details see papers 5.1.1 and 5.1.2 in the Supplement

3.1.2 The comparison of reactivity of residues in 45CBS and full-length CBS

Using the methodology described above, we compared reactivity of residues of the 45CBS and of the full-length CBS. In the 45CBS, we found 50 labeled residues of which only four sites were not detected in the full-length CBS. Three of these differentially modified residues (K172 and/or K177, R336 and K384) are located in the same region of the 45CBS 3-D structure (Fig. 6) suggesting that area containing these amino acids may form a regulatory

interface between the catalytic core and the C-terminal domain. This notion is also supported by the presence of hydrophobic residues in this region (*42*) and by catalytic properties of several disease-causing mutations, namely V173M (*77*), E176K (*74*) and E302K (*48*) which are located at this putative interface and which exhibit catalytic activity similar to the wild-type but fail to be allosterically stimulated by AdoMet.

Differential reactivity of the four modification sites could be explained by interdomain sterical hindrance that is independent of the allosteric motions, or by conformational changes which are responsible for regulation of activity. Therefore, we tested whether the labeling of differentially reactive residues can be restored by a stimulation of the full-length CBS catalytic activity. Since the covalent labeling could not have been performed in the presence of AdoMet due to its reactivity towards the modifiers used, we analysed thermally activated CBS as a surrogate. The restoration of reactivity was observed only for the residues K172 and/or K177 indicating their possible involvement in the regulatory motions. Three other differentially reactive residues were not labeled in the thermally activated CBS and they probably form sterically fixed interdomain interface.

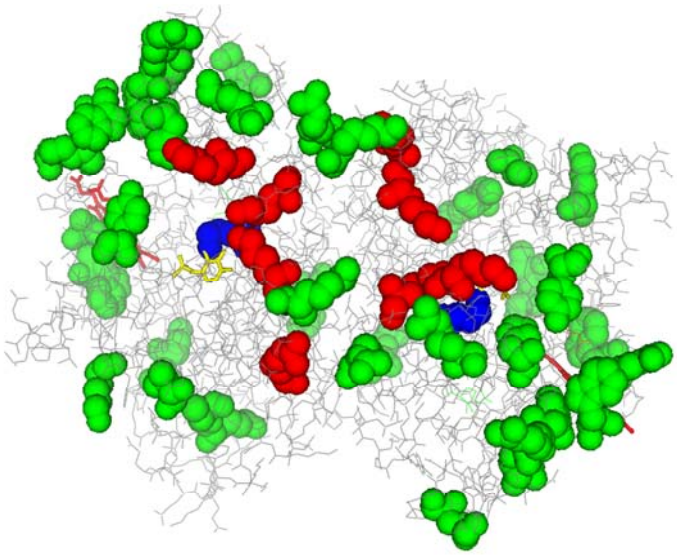


Figure 6 The modification sites shown in the 45CBS crystal structure (PDB ID 1JBQ). The identically modifed residues, i.e. the residues reactive in both the 45CBS and the full-length CBS, are shown as green balls, the differentially modifed sites as red balls. The moieties of heme and PLP are indicated as red and yellow sticks, respectively; the active site Lys119 is shown as blue balls.

Surface mapping showed that the autoinhibition of CBS by the C-terminal regulatory domain is associated with subtle changes at the protein surface. Our observations are consistent with the results from the H/D exchange study by Sen et al. (*41*) suggesting that the allostery in the CBS is not driven by conformational motions but rather by changes in structural flexibility, and/or by population shifts in the structural ensemble.

For details see paper 5.1.2 in the Supplement

3.1.3 Generation of 3-D structural model of the full-length CBS and its comparison with other models

The data from covalent labeling were further used for modeling the 3-D structure of the full-length CBS. Initially, a homology model of the C-terminal regulatory domain was built using recently reported archeal CBS domain binding AdoMet (*78*) as a template. In the next step, the modeled regulatory domain was docked onto the available structure of 45CBS. Our experimental data perfectly fitted to the resulting modeled structure, i.e. the differentially reactive residues formed an interdomain interface while the identically modified sites were located at the surface of the full-length enzyme.

However, the previously published study employing H/D exchange (*41*) revealed interdomain interface in a different region than our study. Although the changes in the surface accessibility of K384 were observed by both H/D exchange and the covalent labeling, other differentially solvent-exposed regions were identified using only one technique. We found differential surface accessibility of the residues K172/K177, R336 and W408/409/410 but they were not reported by Sen et al. On the other hand, the H/D exchange revealed differentially accessible residues in the region 359-370 that were not observed by us; in this segment we identified three identically reactive sites, namely K359, R369 and C370.

The discrepancies between results from these two studies are not clear. The differences may reflect methodological limitations of each technique or may have arisen from different conditions and procedures during preparation of CBS proteins for the structural analyses. It should be also noted that the both proposed models were built as dimers and thus do not provide structural basis of protein tetramerization. It is tempting to speculate that one differentially accessible region may form regulatory interface and the other would be responsible for tetramerization of the full-length CBS.

The knowledge of the 3-D structure of CBS proteins has been recently expanded as the crystal structure of the full-length CBS from *Drosophila melanogaster* has been solved at an atomic resolution (*37*). Although the fruitfly CBS does not bind AdoMet and forms dimers (not tetramers as the human CBS), these data permit inferences on the human ortholog. Comparison of our model with the solved structure of the drosophila CBS indicates that the interface between the catalytic core and the C-terminal domain is formed by virtually identical regions corresponding to residues K172/K177, R336, K384 and W408/W409/W410 in human CBS (Fig.7). These findings suggest that the interdomain interface may be conserved in Bateman domain-containing CBS orthologs including mammalian enzymes.

In addition, it is tempting to speculate that the contact area proposed by the H/D exchange study is not involved in the interdomain communication but appears to be responsible for tetramerization of the full-length human CBS (Fig.7). Nevertheless, these suggestions should be confirmed by advanced molecular dynamic study together with relevant experimental confinements.

In conclusion, our work using surface mapping contributed significantly to expanding the knowledge on the 3-D structure of human CBS.

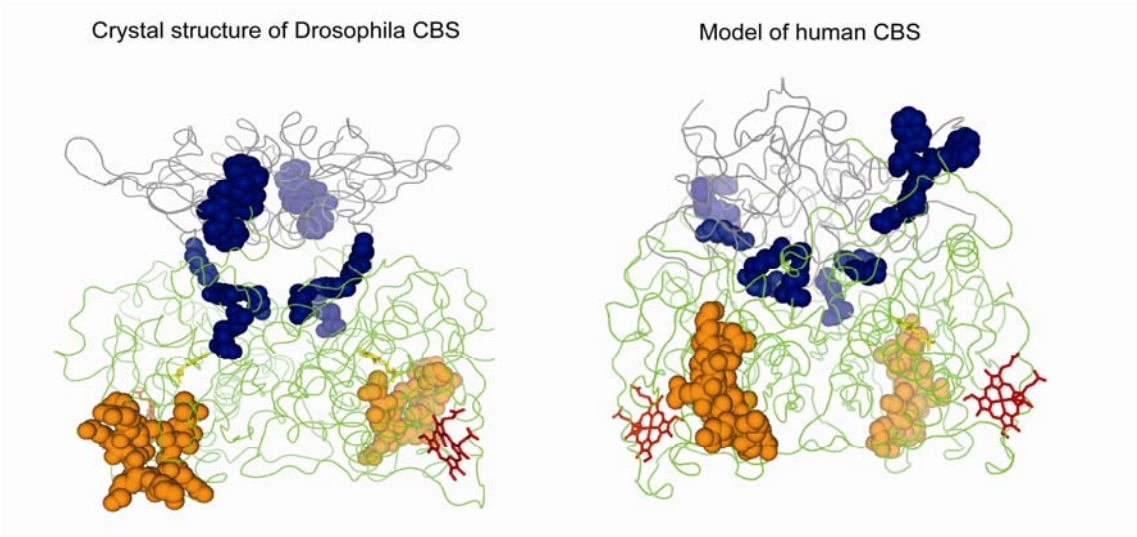


Figure 7 Comparison of the solved 3-D structure from *Drosophila melanogaster* (*37*) (PDB ID 3PC2) with the model of dimeric human CBS based on our data from the surface mapping. The catalytic core is colored in green, the C-terminal domain in grey. The differentially reactive residues from our study are depicted as dark blue balls, the differentially accessible area revealed from the H/D exchange as orange balls. The moieties of heme and PLP are shown as red and yellow sticks, respectively.

For details see paper 5.1.2 in the Supplement

3.2 Characterization of CBS in *Caenorhabditis elegans*

Caenorhabditis elegans represents a well established animal model that is used to study pathogenetic mechanisms of human diseases. However, metabolism of sulfur amino acids in this organism has not yet been elucidated. Thus, we aimed at identification and characterization of CBS in *C.elegans*.

Using *in silico* searches in Blastp,Worm Base and Peptide Atlas together with the experimental procedures such as reverse transcription PCR and the expression analysis of GFP-tagged proteins, we found that the ZC373.1 is the only transcriptionally active gene encoding cystathionine beta-synthase in *C.elegans* (CBS-1).

In further study we demonstrated that CBS-1 has a unique domain architecture that had not been reported for CBS from other species. In contrast to other described metazoan CBS proteins the nematode enzyme lacks both the heme-binding pocket and the Bateman domain, and it also does not assemble into oligomers. Moreover, CBS-1 contains two conserved tandemly arranged PLP-binding domains in one polypeptide (Fig. 8). Interestingly, only the C-terminal catalytic domain binds PLP and catalyses β-replacement reactions as was demonstrated by analysis of truncated (Δ377-704) and active site (E62K and K421A) CBS-1 mutants. Function of the N-terminal noncatalytic conserved domain is unclear; our study showed that the N-terminal domain may be important for proper folding and stabilization of the full-length protein but additional roles such as allosteric regulation of catalytic activity or modulation of protein-protein interactions need to be addressed in future studies.

Interestingly, the above described differences between CBS-1 and mammalian orthologs are also associated with changes in protein stability. Using thermal-based shift assay and pulse proteolysis in a urea gradient, we found that CBS-1 is more sensitive towards denaturation than the human 45CBS; the melting point of CBS-1 was 10 °C lower than that of the human 45CBS, and the resistance towards CBS-1 to urea-induced unfolding was lower by ~ 2.8 M of denaturant compared to the human 45CBS. On the other hand, analysis of enzyme kinetics revealed that the nematode CBS-1 subunit is approximately 4-fold more active - expressed by the turnover number - compared to human 45CBS.

Phylogenetic analysis revealed that the unique structural arrangement is specific only for nematodes. As the CBS-1 protein forms a monomer, it is probable that its N-terminal and C-terminal modules interact to form a structure similar to that of the human 45CBS dimer.

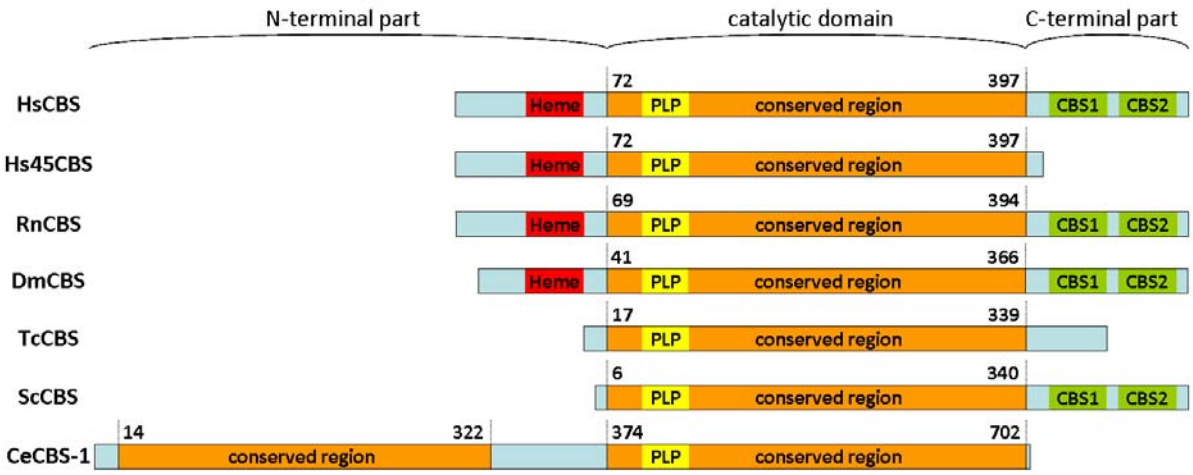


Figure 8 Domain architecture of various CBS proteins. Heme-binding site is colored in red, PLP-binding site in yellow, PLP-dependent conserved regions in orange, flexible loop in light blue and Bateman domain composed of two CBS domains (CBS1 and CBS2) in green. The structure is shown for CBS proteins from different species as follows: HsCBS – *Homo sapiens*; Hs45CBS – C-terminally truncated CBS from *Homo sapiens*; RnCBS – *Rattus norvegicus*; DmCBS – *Drosophila melanogaster*; TcCBS – *Trypanosoma cruzi*; ScCBS – *Saccharomyces cerevisiae*; CeCBS-1 – *Caenorhabditis elegans*.

However, the proposed interdomain interaction cannot be sufficiently supported by the computational modeling procedures using previously solved crystal structures of CBS proteins, and thus it requires further study to determine spatial arrangement of the CBS-1 at an atomic resolution.

In conclusion, our study provides novel insight into the evolutionary divergence of CBS enzymes and demonstrates that the previously described 3-D structure of human and fruitfly CBS protein is not canonical in all metazoan species.

For details see paper 5.1.3 in the Supplement

3.3 Structural characterization of CBS mutants

3.3.1 Conformational analysis of CBS mutant proteins in crude cell extracts

Since the majority of CBS mutants could not be purified in sufficient yields due to an excessive aggregation, we developed an approach for conformational analysis of the mutant variants without the need of purification. This approach is based on two techniques using thermolysin, namely on proteolysis under native conditions, and on pulse proteolysis in a urea gradient. The rate of proteolysis under native conditions provides a knowledge about the extent of unfolding since proteases can cleave only flexible regions and/or unfolded structures (*79*). On the other hand, pulse proteolysis is a technique that monitors urea-induced unfolding of proteins; after a short proteolytic pulse only the unfolded species are digested whereas compact proteins remain uncleaved. Using this assumption in samples with varying concentration of urea, the c_m value as a measure of thermodynamic stability can be determined (*80*).

Initially, we tested feasibility of thermolysin-based approach for analysis of CBS proteins using purified 45CBS and the full-length CBS in the presence and absence of AdoMet. Interestingly, the levels of catalytic activity of CBS proteins are directly proportional to the conformational stability expressed as c_m values. It suggests that autoinhibitory motions in the full-length CBS are accompanied by changes in protein stability. These data further support a notion that allostery in CBS is driven by changes in protein energetics but not by extensive conformational movements.

In the next step, we used proteolytic techniques directly in cell lysates for the analysis of nine pathogenic CBS variants expressed in *E.coli*. For this study, we selected three most common mutants together with representative aminoacid substitutions in each functional domain of the CBS protein. We found that proteolysis under native conditions is a robust technique being feasible even for highly unstable CBS mutants, whereas pulse proteolysis in a urea gradient has limited value for the majority of studied proteins due to their instability and rapid cleavage.

All mutants in the active core (H65R, A114V, T191M, I278T, E302K, G307S, R369C) exhibited increased unfolding. These data show that the unfolded structure is a common intermediate occurring in CBS misfolding. In contrast, mutants in the C-terminal regulatory domain (R439Q and D444N) exhibited decreased proteolytic susceptibility indicating a loss of structural flexibility. We speculate that the rigidified conformation of regulatory domain

mutants is recognized by the protein-quality control machinery and leads to accelerated proteolytic turnover as observed previously *in vivo* (*55-57*).

The extent of unfolding of the studied mutants inversely correlated with the previously determined *(48)* degree of tetrameric assembly (R^2 = 0.6646; p = 0.0074) and with the catalytic activity (R^2 = 0.7566; p = 0.0015). These data may imply that an optimal degree of flexibility is a necessary condition for correct formation of tetramers which further determines the catalytic activity of CBS proteins.

Based on data from this study, the mutants may be divided into three subgroups (Fig. 9):

I, The mutants with affected formation of tetramers, low catalytic activity and extensive unfolding (H65R, T191M, I278T and R369C)

II, Correctly assembled variants with slightly increased unfolding (A114V, E302K and the inactive G307S possessing aminoacid substitution in the proximity of the active site)

III, Correctly assembled hyperactive mutants showing higher stability and lower flexibility (R439Q and D444N)

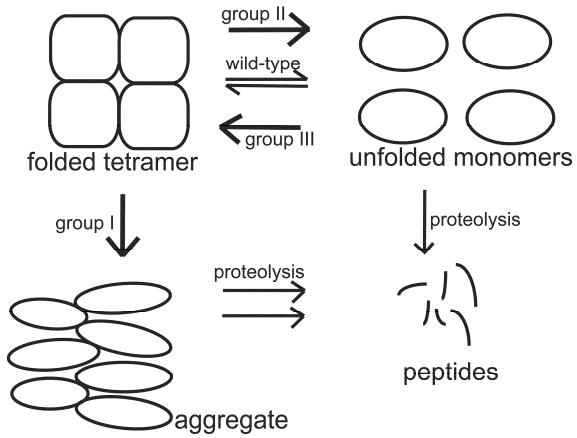


Figure 9 Schematic representation of misfolding of CBS mutants. Model for folded tetramer - unfolded monomers equilibrium was adopted and modified from *(81)*. Lines represent behavior of different mutants: group I, the extensively unfolded mutants forming aggregates (H65R, T191M, I278T, R369C); group II, the proteins with slightly increased unfolding (A114V, E302K, G307S); group III, more tightly folded proteins (R439Q, D444N).

For details see papers 5.1.2 and 5.1.4 in the Supplement

3.3.2 Conformational study of purified CBS mutants

In the previous study by Majtan et al. (*74*), it has been shown that several CBS mutants could be purified into homogeneity if the culture media contained chemical chaperones. However, these purified mutants exhibited unaltered catalytic activity, heme saturation and tetramer assembly making their pathogenicity obscure. In our recent study, we analysed conformation of nine purified mutant proteins from the Majtan´s studies (*46, 74*), namely P49L, P78R, A114V, R125Q, E176K, R266K, P422L, I435T and S466L. Using far-UV circular dichroism, fluorescence and second-derivative UV spectroscopy, we found that global structure of the CBS mutants is similar to the wild-type protein but the conformational properties of the studied mutants differ in the microenviroment of chromophores. Proteolysis with thermolysin under native conditions revealed that the majority of CBS mutants is significantly more susceptible towards cleavage than the wild-type CBS (Fig. 10) indicating propensity of mutant proteins to unfolding. Pulse proteolysis in a urea gradient showed impaired cooperativity of tertiary structure of extensively unfolded mutants R125Q and E176K and, contradictory, unaffected global protein stability of CBS mutants with slightly increased or normal degree of unfolding. This study further supports the hypothesis that increased protein unfolding has important functional consequences even in otherwise unimpaired CBS mutants and proteolysis with thermolysin under native conditions may be a useful tool for the assessment of biochemical penalty of the pathogenic variants.

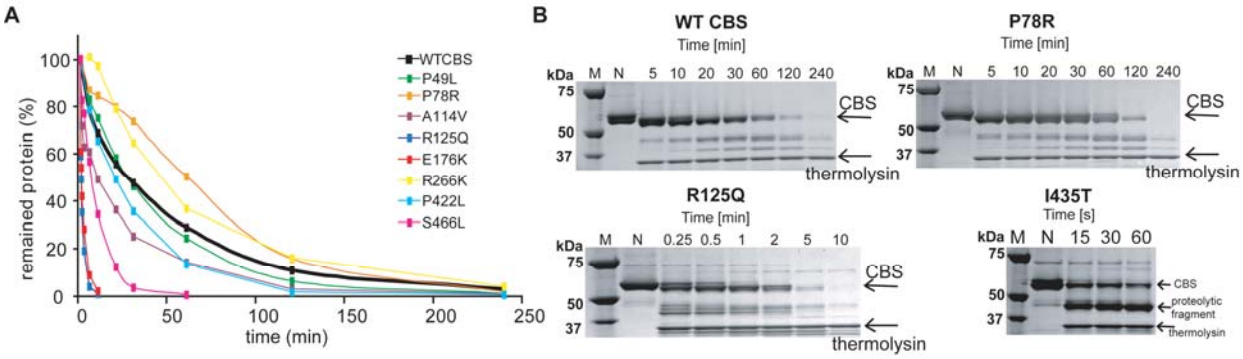


Figure 10 Proteolysis with thermolysin under native conditions of purified CBS mutants. A. Comparison of wild-type CBS with mutant proteins. Each point represents a mean from two independent experiments. B. Representative gels depicting proteolytic cleavage of selected mutants. P78R represents the proteolytically resistant mutants, R125Q belongs to the more rapidly cleaved proteins. The mutant I435T is rapidly cleaved with the formation of the major fragment of molecular weight ≈ 40 kDa. "M" refers to molecular weight marker, "N" refers to the uncleaved control sample.

For details see manuscript 5.1.5 in the Supplement

3.4 Analysis of CBS protein in human plasma

Confirmation of diagnosis in CBS-deficient patients is usually performed by determining the catalytic activity in patient-derived fibroblast cell lines. In this study, we explored whether the CBS protein is present in human plasma and we tested a utility of the analyses in plasma for improved diagnostic procedure.

The CBS activity in human plasma was determined using deuterium-labeled substrate 2,3,3-^{2}H serine together with sensitive detection of labeled product cystathionine by LC-MS/MS. The median CBS activity in healthy controls was 404 nmol/h/L (range 66-1,066, n=57) which, regarding to the activity of the purified enzyme, corresponds to the quantity of CBS protein approximately in the range of nanograms per millilitre. However, these amounts do not permit successful detection by the available immunological techniques, such as Western Blotting or ELISA. The presence of CBS protein in plasma was further evidenced by the stimulation of CBS activity by addition of AdoMet and/or PLP to the reaction mixture. Moreover, detection of CBS-related peptides was reported in a study using shotgun sequencing that was annotated in the Peptide Atlas database (www.peptideatlas.org). Therefore, determination of CBS activity using LC-MS/MS is the only available tool for the analysis of the CBS protein in human plasma. In the next step, we tested utility of this approach for diagnostic purposes. The plasma CBS activity in pyridoxine-nonresponsive CBS-deficient patients was significantly decreased, with the median catalytic activity of 0 nmol/h/L (range 0-9, n=26), whereas the activity in pyridoxine-responsive patients (median 16 nmol/h/L; range 0-358, n=28) overlapped with activities of control subjects.

These data indicate that determination of CBS activity in human plasma represents a new technique for diagnosis of pyridoxine-unresponsive patients obviating the need to set up cell cultures of skin fibroblasts.

For details see paper 5.1.6 in the Supplement

4. CONCLUSIONS

1) Using surface mapping for comparison of the 45CBS with the full-length CBS, we identified differentially reactive residues (K172 and/or K177, R336, K384 and W408 and/or W409 and/or W410) that form a contact area between the catalytic core and the C-terminal regulatory domain. These data were used to generate the structural model of the full-length CBS. The modeled structure is consistent with recently solved 3-D spatial arrangement of the CBS from *Drosophila melanogaster* indicating that the regulatory interface in human CBS is conserved in Bateman domain-containing orthologs. Our study significantly expanded a knowledge about spatial arrangement of human CBS.

2) We identified and characterized the CBS-1 enzyme in *Caenorhabditis elegans*. We found that this CBS protein possesses a unique tandem repeat of conserved catalytic domains in one subunit that does not assemble into oligomers. This study provides novel data on evolutionary divergence of the CBS structure; it shows that the previously described structural features of human and fruitfly CBS proteins are not conserved in all metazoan species.

3A) Using proteolysis with thermolysin under native conditions for analysis of nine CBS mutants in bacterial lysates, we found that the unfolded structure is a common intermediate occurring in protein misfolding of mutants in the heme-binding domain (H65R) and the catalytic core (A114V, T191M, I278T, E302K, G307S, R369C). In contrast, the C-terminal mutants (R439Q and D444N) exhibited lower flexibility and higher structural stability indicating their rigidification. The extent of unfolding of the studied CBS mutants correlated inversely with catalytic activity and with degree of tetrameric assembly.

3B) Furthermore, the majority of purified CBS mutants (A114V, R125Q, E176K, P422L, I435T and S466L) were more susceptible to proteolysis despite exhibiting normal catalytic activity, heme saturation and tetrameric assembly. Both thermolysin-based studies show that the increased proteolytic sensitivity of the CBS mutants may represent an important marker of pathogenicity of mutations in the CBS biomolecule.

4) Our study demonstrated that CBS protein is present in human plasma in low amounts, i.e. approximately in the range of nanograms per milliliter that are insufficient for detection by

immunological techniques. The only feasible approach to detect CBS in the plasma is a determination of its catalytic activity by LC-MS/MS; we showed that this technique may be a useful tool for diagnosing pyridoxine-unresponsive CBS deficiency.

5. AUTHOR´S PUBLICATIONS, PRESENTATIONS AND GRANTS

5.1 PUBLICATIONS CITED IN PHD. THESIS

5.1.1 **Hnízda A**., Šantrůček J., Šanda M., Strohalm M. and Kodíček M.: Reactivity of histidine and lysine side-chains with diethylpyrocarbonate – a method to identify surface exposed residues in proteins. J. Biochem. Biophys. Methods 70, 1091-7 (2008). **IF 1.994**

5.1.2 **Hnízda A**., Spiwok V., Jurga V., Kožich V., Kodíček M., Kraus J.P.: Cross-talk between the catalytic core and the regulatory domain in cystathionine β-synthase: study by differential covalent labeling and computational modeling. Biochemistry 49, 10526-34 (2010). **IF 3.226**

5.1.3 Vozdek R., **Hnízda A.**, Krijt J., Kostrouchová M., Kožich V.: Novel structural arrangement of nematode cystathionine beta-synthases: characterization of Caenorhabditis elegans CBS-1. Biochem. J. 443, 535-547 (2012). **IF 5.016**

5.1.4 **Hnízda A**., Jurga V., Raková K., Kožich V.: Cystathionine beta-synthase mutants exhibit changes in protein unfolding: conformational analysis of misfolded variants in crude cell extracts. J. Inherit. Metab. Dis., DOI 10.1007/s10545-010-9178-3. **IF 3.808**

5.1.5 **Hnízda A.**, Majtan T., Li L., Pey. A.L., Carpenter J., Kodíček M., Kožich V., Kraus J.P.: Conformational analysis of nine purified cystathionine beta-synthase mutants, Unpublished manuscript

5.1.6 Krijt J., Kopecká J., **Hnízda A.,** Moat S., Kluijtmans L.A., Mayne P., Kožich V.: Determination of cystathionine beta-synthase activity in human plasma by LC-MS/MS: potential use in diagnosis of CBS deficiency. J. Inher. Met. Dis. 34 (1), 49-55 (2011). **IF 3.808**

5.2 PUBLICATIONS UNRELATED TO PHD. THESIS

5.2.1 Kučková S., Crhová M., Vaňková L., **Hnízda A**., Hynek R. and Kodíček M.: Towards proteomic analysis of milk proteins in historical building materials. Int. J. Mass. Spec. 284, 42-46 (2009). **IF 2.117**

5.2.2 Zikánová M., Škopová V., **Hnízda A**., Krijt J. and Kmoch S.: Biochemical and structural analysis of 14 mutant adsl enzyme complexes and correlation to phenotypic heterogeneity of adenylosuccinate lyase deficiency. Hum. Mutat. 31(4), 445-55 (2010). **IF 5.956**

5.2.3 Vliet L.K., Wilkinson T.G. 2nd, Duval N., Vacano G., Graham C., Zikánová M, Škopová V., Baresova V., **Hnízda A.** , Kmoch S., Patterson D.: Molecular characterization of the AdeI mutant of Chinese hamster ovary cells: a cellular model of adenylosuccinate lyase deficiency. Mol. Genet. Metab. 102(1), 61-8 (2011). **IF 3.539**

5.3 PRESENTATIONS

5.3.1 **Hnízda A.**, Janošík M., Kodíček M., Kožich V. Surface mapping of cystathionine β-synthase using MALDI-TOF mass spectrometry. Poster presentation. *22nd workshop - Inherited Metabolic Diseases*, Prague, Czech Republic, 16.-18.5. 2007

5.3.2 **Hnízda A.**, Janošík M., Kodíček M., Kožich V. Surface mapping of cystathionine β-synthase using MALDI-TOF mass spectrometry. Poster presentation. *8th Student Scientific Conference, First Faculty of Medicine, Charles University in Prague*, Prague, Czech Republic, 22.5.2007

5.3.3 Kopecká J., **Hnízda A.**, Kožich V. Effect of heme and S-adenosyl-L-methionine on folding and activity of cystathionine beta-synthase mutants. Poster presentation. *6th Conference on Homocysteine Metabolism: World Congress on hyperhomocysteinemia*. Saarbruecken, Germany, 5.-9.6. 2007

5.3.4 **Hnízda A.**, Kožich V. Surface mapping of human cystathionine β-synthase by chemical modification of amino acids. Oral and poster presentation. *23rd workshop - Inherited Metabolic Diseases*, Senec, Slovakia, 14.-16.5. 2008

5.3.5 **Hnízda A.**, Kožich V. Surface mapping of human cystathionine β-synthase by chemical modification of amino acids. Oral presentation. *9th Student Scientific Conference, First Faculty of Medicine, Charles University in Prague*, Prague, Czech Republic, 21.5.2008

5.3.6 **Hnízda A.**, Kožich V. Surface mapping of human cystathionine β-synthase by chemical modification of amino acids. Poster presentation. *7th Human Proteome Organization Conference*, Amsterdam, Netherlands, 15.- 21.8. 2008

5.3.7 **Hnízda A.**, Kraus J.P., Kožich V. Surface mapping of cystathionine beta synthase: insight into enzyme autoinhibition using mass spectrometry. Oral and poster presentation. *24th workshop - Inherited Metabolic Diseases,* Lázně Jeseník, Czech Republic, 13.-15.5. 2009

5.3.8 **Hnízda A.**, Kraus J.P., Kožich V. Surface mapping of cystathionine beta synthase: insight into enzyme autoinhibition using mass spectrometry. Oral presentation. *10th Student Scientific Conference, First Faculty of Medicine, Charles University in Prague*, Prague, Czech Republic, 27.5.2009

5.3.9 Vozdek R., **Hnízda A.**, Krijt J., Kodíček M., Kožich V. Characterization of recombinant cysteine synthase in *Caenorhabditis elegans*. Poster presentation. *7th Conference on Homocysteine Metabolism*, Prague, Czech Republic, 21.-25.6. 2009

5.3.10 **Hnízda A.**, Kraus J.P., Kožich V. Surface mapping of cystathionine beta synthase: insight into enzyme autoinhibition using mass spectrometry. Poster presentation. *7th Conference on Homocysteine Metabolism*, Prague, Czech Republic, 21.-25.6. 2009

5.3.11 Vozdek R., **Hnízda A.**, Krijt J., Kodíček M., Kožich V. Characterization of recombinant cysteine synthase in *Caenorhabditis elegans*. Poster presentation. *34th FEBS Congress*, Prague, Czech Republic, 2.-9.7. 2009, Abstract in FEBS J. (2009) Supplement 1, pp. 351

5.3.12 **Hnízda A.**, Kraus J.P., Kožich V. Surface mapping of cystathionine beta synthase: insight into enzyme autoinhibition using mass spectrometry. Poster presentation. *Young Scientists Forum and 34th FEBS Congress*, Prague, Czech Republic, 3.-9.7. 2009, Abstract in FEBS J. (2009) Supplement 1, pp. 369

5.3.13 **Hnízda A.**, Kraus J.P., Kožich V. Surface mapping of cystathionine beta synthase: insight into enzyme autoinhibition using mass spectrometry. Poster presentation. *18th Intermational Mass Spectrometry Conference*, Bremen, Germany, 29.8.-3.9.2009

5.3.14 **Hnízda A.**, Jurga V., Raková K., Kožich V. Conformational properties of cystathionine β-synthase mutants: analysis in crude cell extracts by proteolytic aproaches. Oral presentation. *11th Student Scientific Conference, First Faculty of Medicine, Charles University in Prague*, Prague, Czech Republic, 19.5.2010

5.3.15 **Hnízda A.**, Jurga V., Raková K., Kožich V. Conformational properties of cystathionine β-synthase mutants: analysis in crude cell extracts by proteolytic approaches. Poster presentation. *Gordon Research Conferences: Enzymes, coenzymes and metabolic pathways*, Waterville Valey, New Hampshire, USA, 18.-23.7. 2010

5.3.16 **Hnízda A.**, Jurga V., Raková K., Kožich V. Conformational properties of cystathionine β-synthase mutants: analysis in crude cell extracts by proteolytic approaches. Poster presentation. *Intristically proteins in biomedicin*e, Barcelona, Spain, 3.-6.10. 2010

5.3.17 Vozdek V., **Hnízda A.**, Kožich V. Unusual structural arrangement of nematode cystathionine beta-synthases suggests evolutionary divergent regulation of their catalytic activity. Poster presentation. 8*th Conference on Homocysteine Metabolism*, Lisboa, Portugal, 19.-22.6.2011

5.3.18 **Hnízda A.**, Jurga V., Raková K., Kožich V. Conformational properties of cystathionine β-synthase mutants: analysis in crude cell extracts by proteolytic approaches. Poster presentation. 8*th Conference on Homocysteine Metabolism*, Lisboa, Portugal, 19.-22.6.2011

5.3.19 Vozdek V., **Hnízda A.**, Kožich V. Unusual structural arrangement of nematode cystathionine beta-synthases suggests evolutionary divergent regulation of their catalytic activity. Poster presentation. *25th Symposium of the Protein Society*, Boston, USA, 23.-27.7. 2011

5.3.20 **Hnízda A.**, Jurga V., Raková K., Kožich V. Conformational properties of cystathionine β-synthase mutants: analysis in crude cell extracts by proteolytic approaches. Poster presentation. *25th Symposium of the Protein Society*, Boston, USA, 23.-27.7. 2011

5.4 GRANTS

2010-2011 Grant of Grant Agency of Charles University in Prague (No. 7310) Conformational analysis of cystathionine beta-synthase mutants

5.5 EDUCATIONAL ACTIVITIES

2009 Consultant of Master Thesis: Jurga V. (Institute of Chemical Technology, Prague) Conformational stability of cystathionine-β-synthase and its mutants

2011 Tutor of IMFSA project: Kamo H. (St. Marianna University, Kanagawa, Japan) Conformational studies of cystathionine beta-synthase mutants

5.6 AWARDS

2010 2nd Price – The Best Publication of Ph.D. students at the First Medical Faculty; Endowment Fund Scientia

6. ABBREVIATIONS

45CBS	C-terminally truncated cystathionine beta-synthase
AdoHcy	S-adenosylhomocysteine
AdoMet	S-adenosylmethionine
ATP	adenosine-5´-triphosphate
BHMT	betaine:homocysteine methyltranferase
CBS	cystathionine beta-synthase
CGL	cystathionine gamma-lyase
DEP	diethylpyrocarbonate
GFP	green fluorescent protein
LC-MS/MS	liquid chromatography-tandem mass spectrometry
MAT	ATP:L-methionine S-adenosyltransferase
MS	methionine synthase
MT	methyltransferase
MTHFR	methylenetetrahydrofolate reductase
PCR	polymerase chain reaction
PDB	Protein Data Bank
PLP	pyridoxal-5´-phosphate
SHMT	serine:hydroxymethyltransferase
THF	tetrahydrofolate

7. REFERENCES

1. Mudd, S. H., and Cantoni, G. L. (1958) Activation of methionine for transmethylation. III. The methionine-activating enzyme of Bakers' yeast, *J Biol Chem 231*, 481-492.
2. Li, Y. N., Gulati, S., Baker, P. J., Brody, L. C., Banerjee, R., and Kruger, W. D. (1996) Cloning, mapping and RNA analysis of the human methionine synthase gene, *Hum Mol Genet 5*, 1851-1858.
3. McKeever, M. P., Weir, D. G., Molloy, A. and Scott, J. M. (1991) Betaine-homocysteine methyltransferase: organ distribution in man, pig and rat and subcellular distribution in the rat, *Clin Sci 81*, 551-556.
4. Blom, H. J., and Smulders, Y. (2011) Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects, *J Inherit Metab Dis 34*, 75-81.
5. Finkelstein, J. D. (1998) The metabolism of homocysteine: pathways and regulation, *Eur J Pediatr 157 Suppl 2*, S40-44.
6. Finkelstein, J. D. (2007) Metabolic regulatory properties of S-adenosylmethionine and S-adenosylhomocysteine, *Clin Chem Lab Med 45*, 1694-1699.
7. Finkelstein, J.D. (2001) Regulation of homocysteine metabolism. In: Carmel R. and Jacobsen D.W. (eds), Homocysteine in health and disease, Cambridge University Press. 92-99.
8. Mudd, S.H., Levy H.L., and Kraus J.P. (2001) Disorders of transsulfuration. In: Scriver, C.R., Beaudet A.L., Sly, W.S. and Valle D. (eds), The metabolic and molecular bases of inherited disease, 8th edn. New York: McGraw-Hill. 2007-2056.
9. Gaustadnes, M., Ingerslev, J., and Rutiger, N. (1999) Prevalence of congenital homocystinuria in Denmark, *N Engl J Med 340*, 1513.
10. Refsum, H., Fredriksen, A., Meyer, K., Ueland, P. M., and Kase, B. F. (2004) Birth prevalence of homocystinuria, *J Pediatr 144*, 830-832.
11. Janosik, M., Sokolova, J., Janosikova, B., Krijt, J., Klatovska, V., and 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*, 431-437.

12. Stabler, S. P., Lindenbaum, J., Savage, D. G., and Allen, R. H. (1993) Elevation of serum cystathionine levels in patients with cobalamin and folate deficiency, *Blood 81*, 3404-3413.
13. Kraus, J.P., and Kozich, V. (2001) Cystathionine-β-synthase and its deficiency. In: Carmel R. and Jacobsen D.W. (eds), Homocysteine in health and disease, Cambridge University Press. 223-243.
14. Mudd, S. H. (2011) Hypermethioninemias of genetic and non-genetic origin: A review, *Am J Med Genet C Semin Med Genet 157*, 3-32.
15. Skovby, F., Kraus, J. P., and Rosenberg, L. E. (1984) Biosynthesis and proteolytic activation of cystathionine beta-synthase in rat liver, *J Biol Chem 259*, 588-593.
16. Frank, N., Kery, V., Maclean, K. N., and Kraus, J. P. (2006) Solvent-accessible cysteines in human cystathionine beta-synthase: crucial role of cysteine 431 in S-adenosyl-L-methionine binding, *Biochemistry 45*, 11021-11029.
17. Sen, S., and Banerjee, R. (2007) A pathogenic linked mutation in the catalytic core of human cystathionine beta-synthase disrupts allosteric regulation and allows kinetic characterization of a full-length dimer, *Biochemistry 46*, 4110-4116.
18. Kabil, O., Zhou, Y., and Banerjee, R. (2006) Human cystathionine beta-synthase is a target for sumoylation, *Biochemistry 45*, 13528-13536.
19. Oliveriusova, J., Kery, V., Maclean, K. N., and Kraus, J. P. (2002) Deletion mutagenesis of human cystathionine beta-synthase. Impact on activity, oligomeric status, and S-adenosylmethionine regulation, *J Biol Chem 277*, 48386-48394.
20. Meier, M., Janosik, M., Kery, V., Kraus, J. P., and Burkhard, P. (2001) Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein, *Embo J 20*, 3910-3916.
21. Singh, S., Madzelan, P., Stasser, J., Weeks, C. L., Becker, D., Spiro, T. G., Penner-Hahn, J., and Banerjee, R. (2009) Modulation of the heme electronic structure and cystathionine beta-synthase activity by second coordination sphere ligands: The role of heme ligand switching in redox regulation, *J Inorg Biochem 103*, 689-697.
22. Taoka, S., Ohja, S., Shan, X., Kruger, W. D., and Banerjee, R. (1998) Evidence for heme-mediated redox regulation of human cystathionine beta-synthase activity, *J Biol Chem 273*, 25179-25184.
23. Taoka, S., Lepore, B. W., Kabil, O., Ojha, S., Ringe, D., and Banerjee, R. (2002) Human cystathionine beta-synthase is a heme sensor protein. Evidence that the redox

sensor is heme and not the vicinal cysteines in the CXXC motif seen in the crystal structure of the truncated enzyme, *Biochemistry 41*, 10454-10461.

24. Kabil, O., Weeks, C. L., Carballal, S., Gherasim, C., Alvarez, B., Spiro, T. G., and Banerjee, R. (2011) Reversible Heme-Dependent Regulation of Human Cystathionine beta-Synthase by a Flavoprotein Oxidoreductase, *Biochemistry 50*, 8261-8263.
25. Majtan, T., Singh, L. R., Wang, L., Kruger, W. D., and Kraus, J. P. (2008) Active cystathionine beta-synthase can be expressed in heme-free systems in the presence of metal-substituted porphyrins or a chemical chaperone, *J Biol Chem 283*, 34588-34595.
26. Kery, V., Poneleit, L., Meyer, J. D., Manning, M. C., and Kraus, J. P. (1999) Binding of pyridoxal 5'-phosphate to the heme protein human cystathionine beta-synthase, *Biochemistry 38*, 2716-2724.
27. Jhee, K. H., Niks, D., McPhie, P., Dunn, M. F., and Miles, E. W. (2001) The reaction of yeast cystathionine beta-synthase is rate-limited by the conversion of aminoacrylate to cystathionine, *Biochemistry 40*, 10873-10880.
28. Taoka, S., and Banerjee, R. (2002) Stopped-flow kinetic analysis of the reaction catalyzed by the full-length yeast cystathionine beta-synthase, *J Biol Chem 277*, 22421-22425.
29. Frank, N., Kent, J. O., Meier, M., and Kraus, J. P. (2008) Purification and characterization of the wild type and truncated human cystathionine beta-synthase enzymes expressed in E. coli, *Arch Biochem Biophys 470*, 64-72.
30. Singh, S., Padovani, D., Leslie, R. A., Chiku, T., and Banerjee, R. (2009) Relative contributions of cystathionine beta-synthase and gamma-cystathionase to H2S biogenesis via alternative trans-sulfuration reactions, *J Biol Chem 284*, 22457-22466.
31. Bateman, A. (1997) The structure of a domain common to archaebacteria and the homocystinuria disease protein, *Trends Biochem Sci 22*, 12-13.
32. Baykov, A. A., Tuominen, H. K., and Lahti, R. (2011) The CBS Domain: A Protein Module with an Emerging Prominent Role in Regulation, *ACS Chem Biol*. 6, 1156-1163.
33. Janosik, M., Kery, V., Gaustadnes, M., Maclean, K. N., and Kraus, J. P. (2001) 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*, 10625-10633.

34. Kery, V., Poneleit, L., and Kraus, J. P. (1998) Trypsin cleavage of human cystathionine beta-synthase into an evolutionarily conserved active core: structural and functional consequences, *Arch Biochem Biophys 355*, 222-232.
35. Jhee, K. H., McPhie, P., and Miles, E. W. (2000) Domain architecture of the heme-independent yeast cystathionine beta-synthase provides insights into mechanisms of catalysis and regulation, *Biochemistry 39*, 10548-10556.
36. Nozaki, T., Shigeta, Y., Saito-Nakano, Y., Imada, M., and Kruger, W. D. (2001) Characterization of transsulfuration and cysteine biosynthetic pathways in the protozoan hemoflagellate, Trypanosoma cruzi. Isolation and molecular characterization of cystathionine beta-synthase and serine acetyltransferase from Trypanosoma, *J Biol Chem 276*, 6516-6523.
37. Koutmos, M., Kabil, O., Smith, J. L., and Banerjee, R. (2010) Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine {beta}-synthase, *Proc Natl Acad Sci U S A 107*, 20958-20963.
38. Prudova, A., Martinov, M. V., Vitvitsky, V. M., Ataullakhanov, F. I., and Banerjee, R. (2005) Analysis of pathological defects in methionine metabolism using a simple mathematical model, *Biochim Biophys Acta 1741*, 331-338.
39. Martinov, M. V., Vitvitsky, V. M., Banerjee, R., and Ataullakhanov, F. I. (2011) The logic of the hepatic methionine metabolic cycle, *Biochim Biophys Acta 1804*, 89-96.
40. Zou, C. G., and Banerjee, R. (2003) Tumor necrosis factor-alpha-induced targeted proteolysis of cystathionine beta-synthase modulates redox homeostasis, *J Biol Chem 278*, 16802-16808.
41. Sen, S., Yu, J., Yamanishi, M., Schellhorn, D., and Banerjee, R. (2005) Mapping peptides correlated with transmission of intrasteric inhibition and allosteric activation in human cystathionine beta-synthase, *Biochemistry 44*, 14210-14216.
42. Meier, M., Oliveriusova, J., Kraus, J. P., and Burkhard, P. (2003) Structural insights into mutations of cystathionine beta-synthase, *Biochim Biophys Acta 1647*, 206-213.
43. Yamanishi, M., Kabil, O., Sen, S., and Banerjee, R. (2006) Structural insights into pathogenic mutations in heme-dependent cystathionine-beta-synthase, *J Inorg Biochem 100*, 1988-1995.
44. Ojha, S., Wu, J., LoBrutto, R., and Banerjee, R. (2002) Effects of heme ligand mutations including a pathogenic variant, H65R, on the properties of human cystathionine beta-synthase, *Biochemistry 41*, 4649-4654.

45. Chen, X., Wang, L., Fazlieva, R., and Kruger, W. D. (2006) Contrasting behaviors of mutant cystathionine beta-synthase enzymes associated with pyridoxine response, *Hum Mutat 27*, 474-482.
46. Majtan, T., and Kraus, J. P. (2012) Folding and activity of mutant cystathionine beta-synthase depends on the position and the nature of the purification tag: Characterization of the R266K CBS mutant, *Protein Expr Purif 82*, 317-324.
47. Urreizti, R., Balcells, S., Rodes, M., Vilarinho, L., Baldellou, A., Couce, M. L., Munoz, C., Campistol, J., Pinto, X., Vilaseca, M. A., and Grinberg, D. (2003) Spectrum of CBS mutations in 16 homocystinuric patients from the Iberian Peninsula: high prevalence of T191M and absence of I278T or G307S, *Hum Mutat 22*, 103.
48. Kozich, V., Sokolova, J., Klatovska, V., Krijt, J., Janosik, M., Jelinek, K., and Kraus, J. P. (2010) Cystathionine beta-synthase mutations: effect of mutation topology on folding and activity, *Hum Mutat 31*, 809-819.
49. Moat, S. J., Bao, L., Fowler, B., Bonham, J. R., Walter, J. H., and Kraus, J. P. (2004) The molecular basis of cystathionine beta-synthase (CBS) deficiency in UK and US patients with homocystinuria, *Hum Mutat 23*, 206.
50. Janosik, M., Oliveriusova, J., Janosikova, B., Sokolova, J., Kraus, E., Kraus, J. P., and Kozich, V. (2001) Impaired heme binding and aggregation of mutant cystathionine beta-synthase subunits in homocystinuria, *Am J Hum Genet 68*, 1506-1513.
51. Shan, X., and Kruger, W. D. (1998) Correction of disease-causing CBS mutations in yeast, *Nat Genet 19*, 91-93.
52. Shan, X., Dunbrack, R. L., Jr., Christopher, S. A., and Kruger, W. D. (2001) Mutations in the regulatory domain of cystathionine beta synthase can functionally suppress patient-derived mutations in cis, *Hum Mol Genet 10*, 635-643.
53. Gallagher, P. M., Ward, P., Tan, S., Naughten, E., Kraus, J. P., Sellar, G. C., McConnell, D. J., Graham, I., and Whitehead, A. S. (1995) High frequency (71%) of cystathionine beta-synthase mutation G307S in Irish homocystinuria patients, *Hum Mutat 6*, 177-180.
54. Hu, F. L., Gu, Z., Kozich, V., Kraus, J. P., Ramesh, V., and Shih, V. E. (1993) Molecular basis of cystathionine beta-synthase deficiency in pyridoxine responsive and nonresponsive homocystinuria, *Hum Mol Genet 2*, 1857-1860.
55. Maclean, K. N., Gaustadnes, M., Oliveriusova, J., Janosik, M., Kraus, E., Kozich, V., Kery, V., Skovby, F., Rudiger, N., Ingerslev, J., Stabler, S. P., Allen, R. H., and Kraus, J. P. (2002) High homocysteine and thrombosis without connective tissue disorders

are associated with a novel class of cystathionine beta-synthase (CBS) mutations, *Hum Mutat 19*, 641-655.

56. Evande, R., Blom, H., Boers, G. H., and Banerjee, R. (2002) Alleviation of intrasteric inhibition by the pathogenic activation domain mutation, D444N, in human cystathionine beta-synthase, *Biochemistry 41*, 11832-11837.
57. Gupta, S., Wang, L., Hua, X., Krijt, J., Kozich, V., and Kruger, W. D. (2008) Cystathionine beta-synthase p.S466L mutation causes hyperhomocysteinemia in mice, *Hum Mutat 29*, 1048-1054.
58. Bross, P., Corydon, T. J., Andresen, B. S., Jorgensen, M. M., Bolund, L., and Gregersen, N. (1999) Protein misfolding and degradation in genetic diseases, *Hum Mutat 14*, 186-198.
59. Yue, P., Li, Z., and Moult, J. (2005) Loss of protein structure stability as a major causative factor in monogenic disease, *J Mol Biol 353*, 459-473.
60. Pey, A. L., Stricher, F., Serrano, L., and Martinez, A. (2007) Predicted effects of missense mutations on native-state stability account for phenotypic outcome in phenylketonuria, a paradigm of misfolding diseases, *Am J Hum Genet 81*, 1006-1024.
61. Waters, P. J., Parniak, M. A., Akerman, B. R., Jones, A. O., and Scriver, C. R. (1999) Missense mutations in the phenylalanine hydroxylase gene (PAH) can cause accelerated proteolytic turnover of PAH enzyme: a mechanism underlying phenylketonuria, *J Inherit Metab Dis 22*, 208-212.
62. Eiken, H. G., Knappskog, P. M., Apold, J., and Flatmark, T. (1996) PKU mutation G46S is associated with increased aggregation and degradation of the phenylalanine hydroxylase enzyme, *Hum Mutat 7*, 228-238.
63. Gregersen, N. (2006) Protein misfolding disorders: pathogenesis and intervention, *J Inherit Metab Dis 29*, 456-470.
64. Soto, C., and Estrada, L. D. (2008) Protein misfolding and neurodegeneration, *Arch Neurol 65*, 184-189.
65. Singh, L. R., and Kruger, W. D. (2009) Functional rescue of mutant human cystathionine beta-synthase by manipulation of Hsp26 and Hsp70 levels in Saccharomyces cerevisiae, *J Biol Chem 284*, 4238-4245.
66. Loo, T.W., and Clarke, D.M. (2007) Chemical and pharmacological chaperones as new therapeutic agents, *Expert Rev Mol Med 9*, 1-18.
67. Leandro, P., and Gomes, C. M. (2008) Protein misfolding in conformational disorders: rescue of folding defects and chemical chaperoning, *Mini Rev Med Chem 8*, 901-911.

68. Muntau, A. C., Roschinger, W., Habich, M., Demmelmair, H., Hoffmann, B., Sommerhoff, C. P., and Roscher, A. A. (2002) Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria, *N Engl J Med 347*, 2122-2132.

69. Erlandsen, H., Pey, A. L., Gamez, A., Perez, B., Desviat, L. R., Aguado, C., Koch, R., Surendran, S., Tyring, S., Matalon, R., Scriver, C. R., Ugarte, M., Martinez, A., and Stevens, R. C. (2004) Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations, *Proc Natl Acad Sci U S A 101*, 16903-16908.

70. Pey, A. L., Perez, B., Desviat, L. R., Martinez, M. A., Aguado, C., Erlandsen, H., Gamez, A., Stevens, R. C., Thorolfsson, M., Ugarte, M., and Martinez, A. (2004) Mechanisms underlying responsiveness to tetrahydrobiopterin in mild phenylketonuria mutations, *Hum Mutat 24*, 388-399.

71 Perez, B., Desviat, L. R., Gomez-Puertas, P., Martinez, A., Stevens, R. C., and Ugarte, M. (2005) Kinetic and stability analysis of PKU mutations identified in BH4-responsive patients, *Mol Genet Metab 86 Suppl 1*, S11-16.

72. Singh, L. R., Chen, X., Kozich, V., and Kruger, W. D. (2007) Chemical chaperone rescue of mutant human cystathionine beta-synthase, *Mol Genet Metab 91*, 335-342.

73. Singh, L. R., Gupta, S., Honig, N. H., Kraus, J. P., and Kruger, W. D. (2010) Activation of mutant enzyme function in vivo by proteasome inhibitors and treatments that induce Hsp70, *PLoS Genet 6*, e1000807.

74. Majtan, T., Liu, L., Carpenter, J. F., and Kraus, J. P. (2010) Rescue of cystathionine beta-synthase (CBS) mutants with chemical chaperones: purification and characterization of eight CBS mutant enzymes, *J Biol Chem 285*, 15866-15873.

75. Kopecka, J., Krijt, J., Rakova, K., and Kozich, V. (2011) Restoring assembly and activity of cystathionine beta-synthase mutants by ligands and chemical chaperones, *J Inherit Metab Dis 34*, 39-48.

76. Mendoza, V. L., and Vachet, R. W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection, *Anal Chem 80*, 2895-2904.

77. Urreizti, R., Asteggiano, C., Cozar, M., Frank, N., Vilaseca, M. A., Grinberg, D., and Balcells, S. (2006) Functional assays testing pathogenicity of 14 cystathionine-beta synthase mutations, *Hum Mutat 27*, 211.

78. Lucas, M., Encinar, J. A., Arribas, E. A., Oyenarte, I., Garcia, I. G., Kortazar, D., Fernandez, J. A., Mato, J. M., Martinez-Chantar, M. L., and Martinez-Cruz, L. A. (2010) Binding of S-methyl-5'-thioadenosine and S-adenosyl-L-methionine to protein

MJ0100 triggers an open-to-closed conformational change in its CBS motif pair, *J Mol Biol 396*, 800-820.

79. Young, T. A., Skordalakes, E., and Marqusee, S. (2007) Comparison of proteolytic susceptibility in phosphoglycerate kinases from yeast and E. coli: modulation of conformational ensembles without altering structure or stability, *J Mol Biol 368*, 1438-1447.

80. Park, C., and Marqusee, S. (2005) Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding, *Nat Methods 2*, 207-212.

81. Park, C., and Marqusee, S. (2004) Analysis of the stability of multimeric proteins by effective DeltaG and effective m-values, *Protein Sci 13*, 2553-2558.

SUPPLEMENT

Supplement contains reprints of publications and manuscripts 5.1.1 – 5.1.6.

Publication 5.1.1

A. Hnízda et al.

Reactivity of histidine and lysine side-chains with diethylpyrocarbonate – A method to identify surface exposed residues in proteins

Journal of Biochemical and Biophysical Methods, 2008

Available online at www.sciencedirect.com
ScienceDirect
J. Biochem. Biophys. Methods 70 (2008) 1091 – 1097
JOURNAL OF biochemical and biophysical methods
www.elsevier.com/locate/jbbm

Reactivity of histidine and lysine side-chains with diethylpyrocarbonate — A method to identify surface exposed residues in proteins

Aleš Hnízda [a], Jiří Šantrůček [b], Miloslav Šanda [c], Martin Strohalm [b], Milan Kodíček [b,*]

[a] Institute of Inherited Metabolic Disorders, 1st School of Medicine, Charles University, 128 00 Praha 2, Czech Republic
[b] Department of Biochemistry and Microbiology, Institute of Chemical Technology in Prague, 166 28 Praha 6, Czech Republic
[c] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 166 10 Praha 6, Czech Republic

Received 7 June 2007; received in revised form 13 July 2007; accepted 15 July 2007

Abstract

The chemical modification of amino acid side-chains followed by mass spectrometric detection can reveal at least partial information about the 3-D structure of proteins. In this work we tested diethylpyrocarbonate, as a common histidyl modification agent, for this purpose. Appropriate conditions for the reaction and detection of modified amino acids were developed using angiotensin II as a model peptide. We studied the modification of several model proteins with a known spatial arrangement (insulin, cytochrome c, lysozyme and human serum albumin). Our results revealed that the surface accessibility of residues is a necessary, although in itself insufficient, condition for their reactivity; the microenvironment of side-chains and the dynamics of protein structure also affect the ability of residues to react. However the detection of modified residues can be taken as proof of their surface accessibility, and of direct contact with solvent molecules.

Keywords: Diethylpyrocarbonate; Histidine modification; Mass spectrometry; Protein modification; Surface accessibility; Surface mapping

1. Introduction

Knowledge of the spatial arrangement of proteins is crucial for understanding of their function in biological systems. X-ray crystallography and nuclear magnetic resonance (NMR) are considered to be the most powerful techniques for providing extended information about the 3-D structure of proteins. Both techniques, however, have limitations. High-quality protein crystals are necessary for successful diffraction analysis, but it can be both laborious and difficult to harvest them. On the other hand, NMR spectroscopy enables the measurement of solutions, but it can only be used for smaller soluble proteins [1]. Consequently there is a great need for other methods that can provide at least partial structural information, while also being less labor and time-intensive.

The chemical modification of proteins for microenvironment probing has been employed for the past 30 years [2]. Great advances in the development of mass spectrometric techniques in the last decade have led to greater sophistication in the study of this area. The approach of side-chain-specific modification, for example, has been applied to the monitoring of conformational changes, especially during ligand binding [3,4] and protein–protein interaction [5]. This method provides information about protein surface topology because in some cases the surface accessibility of residues can be directly correlated with their reactivity. This relationship has been established for lysine modification by succinic anhydride [6], as well as for tryptophan modification by 2-hydroxy-5-nitrobenzyl bromide (Koshland's reagent) [7]. However the relationship between reactivity and surface accessibility can be ambiguous, e.g. tyrosine modification by iodine or tetranitromethane [8,9]. This approach has been further developed by usage of bifunctional

Abbreviations: BSA, bovine serum albumin; DEP, diethylpyrocarbonate; DHB, 2,5-dihydroxybenzoic acid; DTT, dithiothreitol; HSA, human serum albumin; IAA, iodoacetamide; LC-ESI MS/MS, liquid chromatography–electrospray ionization tandem mass spectrometry; MALDI-TOF MS, matrix assisted laser desorption/ionization time of flight mass spectrometry; NMR, nuclear magnetic resonance; PSD, post source decay; Q-TOF, quadrupole-time of flight; SDS PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis.

* Corresponding author. Institute of chemical technology in Prague, Technická 5, 166 28 Praha 6, Czech Republic. Fax: +420 220 448 5167.
E-mail address: Milan.Kodicek@vscht.cz (M. Kodíček).

0165-022X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbbm.2007.07.004

agents that might be used for protein cross-linking, with the distance between cross-linked residues being estimated according to the length of the spacer [10].

More recently the technique of the chemical oxidation of proteins has been developed. Hydrogen peroxide can be used as a source of hydroxyl radicals, the generation of which is catalyzed by chelated iron [11], by UV-radiation [12] or by laser flash [13]. Other protein oxidation methods include applying an electrical discharge (in electrospray ionization mass spectrometry experiments) [14], or the radiolysis of water using synchrotron [15]. The surface mapping of proteins using protein oxidation is a very promising area, in particular because the residues primarily modified are solvent-accessible. The only exceptions are the spontaneously oxidizing buried methionine side-chains [11,12].

The main advantages of these methods for the determination of protein 3-D structure are their low material demands, and the fact that they provide an opportunity to study the desired protein in its native conditions, which is especially valuable for membrane proteins and lipoprotein particles [9,16]. In addition, data collected from the chemical modification of proteins might also be used to predict protein structure using computational modeling [17,18].

In this work we focused on the reactivity of amino acids with diethylpyrocarbonate (DEP), the agent most frequently used for histidine modification [2]. In proteins, the ethoxyformyl group is bound to the nitrogen atom of imidazole (Fig. 1). At higher DEP concentration levels, other products are formed, in which the imidazole ring is disrupted [19].

It should be noted that DEP is a quite non-specific modifier, and that other residues, namely lysine and tyrosine, can react [2]. Formerly, absorption spectrophotometry and the hydroxylamine treatment of modified proteins were the only tools available for discriminating among the reacting groups [20]. The reaction between DEP and histidine is accompanied by an increase in absorbance in the wavelength range of 230–250 nm. The extent of histidine modification can be quantified using a molar extinction coefficient, the most frequently used value being $\varepsilon_{240\ nm}=3200\ dm^3\ mol^{-1}\ cm^{-1}$ [21]. Tyrosine modification causes a decrease in absorbance at the wavelength of 278 nm. The hydroxylamine treatment of modified proteins confirms lysine modification, which, in contrast to other modifications, is not removed by this procedure [2,20].

Modification of proteins by DEP has been studied in several papers. Using NMR spectroscopy it has been found that the extent of histidine modification in cytochrome b_5 is influenced by imidazole surface accessibility, pK_A and hydrogen-binding involvement [22]. In the study using mass spectrometry (MS) for detection of DEP modification, angiotensin II and insulin were used as model compounds, and histidine reactivity was found to correlate with the surface accessibility of imidazole nitrogen atoms [23]; however this same study neglected tyrosine and lysine reactivity.

We conducted a detailed analysis of protein modification, including lysine and tyrosine modification. We chose model proteins of different sizes and structural characteristics (insulin, cytochrome *c*, lysozyme and human serum albumin), and systematically studied the reactivity of their side-chains in an attempt to find a correlation between modification ability and protein surface topology.

2. Materials and methods

Human angiotensin II (hereinafter angiotensin), cytochrome *c* from horse heart (cytochrome), lysozyme from hen egg white (lysozyme), human serum albumin (HSA), α-chymotrypsin (chymotrypsin), endoprotease Glu-C (Glu-C), iodoacetamide (IAA), dithiothreitol (DTT), trifluoroacetic acid and DEP were all purchased from SIGMA. Porcine insulin was supplied by

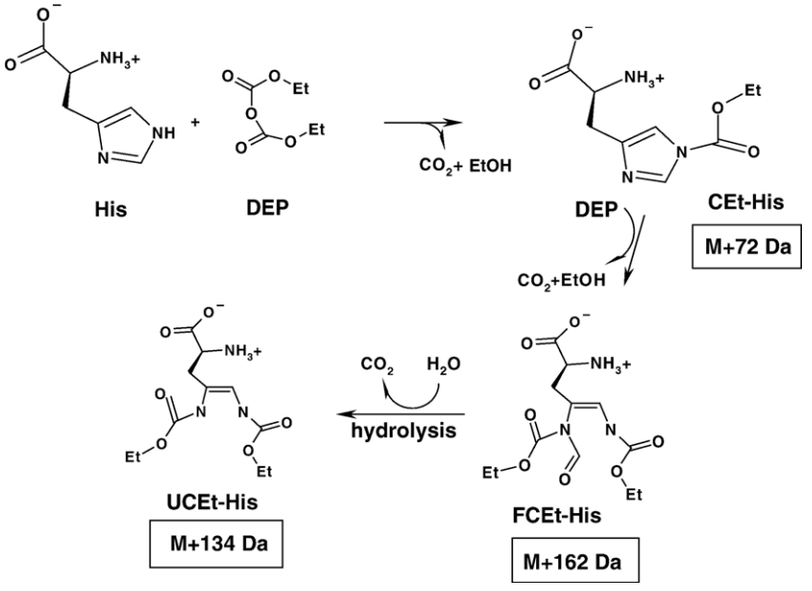


Fig. 1. Modification of histidine residue by DEP (CEt-His — carbethoxyhistidine, FCEt-His-formyl-biscarbethoxyhistidine, UCEt-His — urethane-carbethoxyhistidine) [19].

NovoNordisk, trypsin (sequencing grade) by Promega, 2,5-dihydroxybenzoic acid (DHB) by Bruker Daltonics, and hydroxylamine hydrochloride by Fluka.

2.1. Protein modification by DEP

Model peptides and proteins (1 mg/ml) were dissolved in 50 mM phosphate buffer (pH 6.8). Insulin, which is insoluble under these conditions, was dissolved in 50 mM $NaHCO_3$ (pH 9), with pH then adjusted with concentrated HCl to 6.8. Freshly prepared DEP solution in methanol was added to the protein to obtain a protein:DEP molar ratio in the range of 1:1 to 1:100. Incubation took 1 h at room temperature. Excess DEP was removed by gel filtration using a PD-10 column (Amersham Biosciences) and 50 mM NH_4HCO_3 as an elution buffer [23,24].

2.2. Reaction of modified proteins with hydroxylamine

After protein incubation with DEP, freshly prepared hydroxylamine solution (2 M; pH value adjusted with KOH to 6.8) was added, the concentration of hydroxylamine in the reaction mixture being in the range of 0.1–1 M. Incubation took 1 h at room temperature. Hydroxylamine was separated from protein by gel filtration (see above).

2.3. In-solution digestion

Model proteins were reduced by DTT (30 min, 50 °C, final concentration 5 mM). Thiol groups were modified by IAA (30 min in dark, room temperature, final concentration 25 mM). The proteins were digested by trypsin, chymotrypsin or Glu-C in 50 mM NH_4HCO_3 (1 h, 37 °C). Cytochrome was digested for only 10 min to achieve partial proteolysis. The protein:enzyme ratio was always 20:1 (w/w). The protein digest was mixed with matrix solution, and the mixture directly analyzed by MALDI-TOF MS (see below).

2.4. In-gel digestion

Modified cytochrome was used to test the application of in-gel digestion to ethoxyformylation analysis. After SDS-PAGE, the gel was stained with Coomasie Brilliant Blue. Gel slices were excised and washed with distilled water. After destaining with a solution of acetonitrile:50 mM NH_4HCO_3 (1:1 v/v), reduction by 10 mM DTT occurred for 45 min at 55 °C. Reduced thiols were modified by 55 mM IAA for 30 min in the dark at room temperature. Protein was digested by trypsin (12 ng/μl in NH_4HCO_3) overnight, at 37 °C. After cleavage, peptides were extracted from the gel with 0.1% trifluoroacetic acid in an ultrasonic bath.

2.5. MALDI-TOF MS

All samples were measured using a Biflex IV (Bruker Daltonic) equipped with a nitrogen laser (337 nm) in positive reflector mode. A DHB matrix was prepared as a solution of 10 mg/ml in a mixture of acetonitril:0.1% trifluoroacetic acid (1:2 v/v). The sample was mixed with DHB in a matrix:sample ratio of 3.5:1, and 1 μl of this mixture was deposited on the target plate (dried droplet method). The spectrometer was calibrated with a BSA tryptic digestion mixture before each analysis. Peptide fragmentation was analyzed by MALDI-PSD. Data were collected and compiled using Bruker XTof and mMass software [25].

2.6. LC-ESI MS/MS

The peptides produced by the protein digestion process were separated by a 2D CapLC (Waters), using the first trapping column (Symetry300 with stationary phase C_{18}) with mobile phase A (2% acetonitril + 0.1% formic acid), and the second separating column (Atlantis with stationary phase dC_{18}), with mobile phase A and mobile phase B (90% acetonitril, 0.1% formic acid). Initial conditions, maintained for 3 min, consisted of 90% A. The peptides were eluted (flow rate was 200 nl/min) using a gradient program (1st min 90% A, 17th min 60% A, 19th min 20% A, 22nd min 20% A, 23rd min 90% A, 37th min 90% A). Mass spectrometric detection was performed using Q-TOF (Waters Micromass). Electrospray voltage was 3500 V, cone voltage 40 V. The modified peptides were measured in positive mode, energy of 35 eV being used to fragment double-charged precursor ions.

2.7. Surface accessibility determination

Surface accessibility values were collected from PDB files of the analyzed proteins (insulin — 4INS, cytochrome — 1AKK, lysozyme — 1IO5 and 2CDS, HSA — 1AO6 and 1UOR) using MBG (free) software, which functions according to an algorithm designed by B. Lee and F.M. Richards [26].

3. Results

3.1. Analysis of angiotensin as a model peptide

Octapeptide angiotensin (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe) contains three sites that can react with DEP: α-amine, p-hydroxyphenyl, and imidazol groups [2]. The concentration of DEP in the reaction mixture was chosen to obtain peptide:DEP molar ratios of 1:2 and 1:20 (0.92 mM and 9.22 mM respectively). Mass increases in multiplies of 72 Da were detected even for the lower concentration of DEP, indicating that several groups reacted simultaneously (Fig. 2). For the higher concentration of DEP, mass shifts of +134 and +162 Da were also detected in some cases, indicating histidyl "overmodification" (Fig. 1).

Modified histidine was identified using MALDI-PSD; in the spectra the intensive peak of its immonium ion (182 Da) [23] was detected. The amount of modified histidine and tyrosine residues can be estimated using difference spectrophotometry [21], which, in our study, showed that 22% of histidine residues and 8% of tyrosine residues reacted at the peptide:DEP ratio of 1:2. For the higher concentration of DEP, the molar extinction coefficient could not be applied because "overmodification" caused undefined spectral changes.

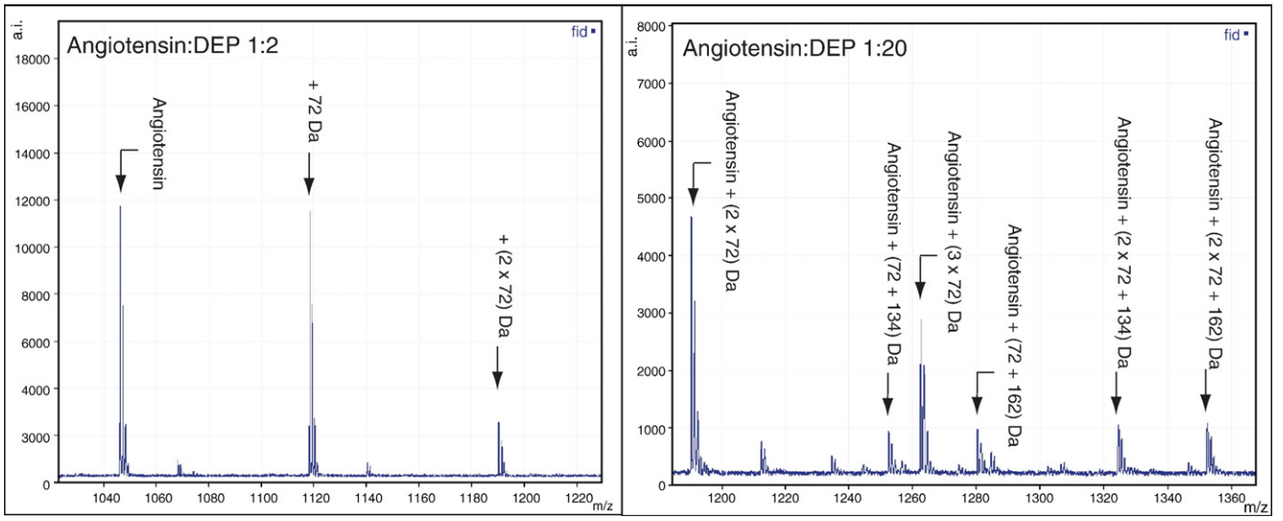


Fig. 2. Spectra of modified angiotensin. Peptide: DEP ratios were 1:2 (left) and 1:20 (right). Mass shift +72 Da corresponds to ethoxyformyl binding, the shifts +134 Da and +162 Da correspond to histidine "overmodification".

The hydroxylamine treatment of modified peptides can serve as a useful tool for the identification of reacting groups. The ethoxyformylated amino group is the only resistant one, the other groups losing their modifications easily under the procedure [21]. In the case of angiotensin, the difference absorbance decreased markedly after hydroxylamine treatment, indicating histidine and tyrosine modification. In the mass spectra, only one (primary amine) modification was found after hydroxylamine treatment, which is in strong agreement with data obtained from spectrophotometry (Fig. 3).

3.2. Analysis of model proteins

3.2.1. General remarks about the method

Spectrophotometry showed that tyrosine residues in the studied proteins are non-reactive with DEP at the concentrations used. To differentiate between modified lysine and histidine residues, we primarily used hydroxylamine treatment. As MALDI-PSD MS is not very effective for more complicated peptide mixtures, it was necessary to use more sophisticated MS techniques for peptide sequencing, which in the case of insulin involved LC-ESI MS/MS. However, the low stability of histidine modification during liquid chromatography in acidic pH (the most common solution being trifluoroacetic acid) limits the use of LC-ESI MS/MS. This is particularly the case for proteins of a higher molecular weight because their digests contain several modified peptides at relatively lower concentrations. These problems were observed during analysis of HSA, but also of cytochrome.

3.2.2. Insulin

Insulin modification by DEP was studied for insulin:DEP ratios ranging from 1:1 to 1:100 (the DEP concentration in the reaction mixture being in the range of 0.15–15 mM). For detailed study we chose an insulin:DEP ratio of 1:10 (1.5 mM

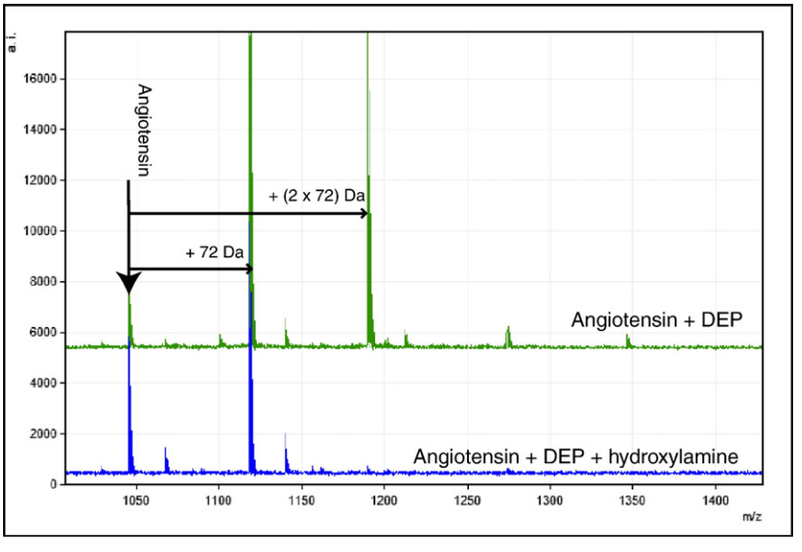


Fig. 3. The spectra of DEP modified angiotensin before and after hydroxylamine treatment (peptide:DEP ratio 1:2).

Table 1
The outline of modified side-chains in insulin

Protease	Peptide	Detected mass [Da]	Change of MW [Da]	Modified residue	Confirmed by method
Glu-C	FVNQHLCGSHLVE	1611.74	+72	Phe1	MALDI-PSD MS; ESI MS/MS
Glu-C	FVNQHLCGSHLVE	1698.67	+(2×72+15)	Phe1, His5, His10	ESI MS/MS
Glu-C	RGFFYTPKA	1158.56	+72	Lys26	xxx [a]
Chymotrypsin	FVNQHL	829.4	+72	Phe1	MALDI-PSD MS
Chymotrypsin	CGSHLVEALY	1220.61	+72	His10	MALDI-PSD MS

[a] The confirmation of Lys26 modification was not necessary in this case.

DEP); the peaks of modified peptides were relatively intensive, and irreversible overmodification of the imidazole ring was not being detected.

We discovered that histidine modification disables entire cystein acetamidation prior to proteolytic cleavage; we detected an increase in peptide mass by just 15 Da, rather than an increase by 72 Da, which suggested that acetamidation, which increases peptide mass by 57 Da, didn't occur (i.e. 72–57= +15 Da). Use of the MALDI-PSD MS and LC-ESI MS/MS methods confirmed this hypothesis (Table 1). Having understood the problem, we were able to reliably identify modifications of His5, His10 and Lys26, as well as the α-amino group Phe1. Our findings show that "+15 Da peaks" should be searched for in every detailed analysis of protein modification by DEP in the case when acetamidation took place in the sample preparation. Indeed, we also found this type of mass shift in the case of HSA modification.

3.2.3. Cytochrome

Cytochrome was modified by DEP at protein:DEP ratios ranging from 1:1 to 1:100 (0.08–8 mM DEP). At the ratio of 1:5 (0.4 mM), modifications of 5 lysine residues and 1 histidine residue were identified (Table 2). In addition, His18 was determined to be non-reactive because mass shift of +72 Da of peptide 14–22 was not observed. It was more complicated to analyze the reactivity of His26. By complete trypsin digestion, dipeptide 26–27, which exhibits low molecular weight and cannot be detected by MALDI-TOF MS, was formed. However, after partial trypsin digestion (10 min), His26 was found in peptide 26–38, with one uncleaved site; this peptide also contains Lys27 and His33. The modification of His26 containing an accessible nitrogen atom was only observed for the highest concentration of DEP (ratio 1:100), resulting in a triple-modified peptide. Using a lower concentration of DEP (ratio 1:10), we were only able to prove the modification of Lys27, together with at least one of the histidines in this peptide. Because modified His33 was also found in peptide 28–38, the modification of His26 at lower DEP concentration levels seems improbable.

Cytochrome was also used as a model for testing the method of in-gel digestion for the identification of modified residues. We found that ethoxyformyl–histidine cannot be detected in this way because of its low stability during the procedure. However, because modified lysine is stable, it was possible to use this method to find all modified lysines also detectable by in-solution digestion.

Table 2
The outline of modified residues in cytochrome for the ratio protein:DEP of 1:5

Peptide	Sequence (trypsin digest)	Detected mass of modified peptide [Da]	Reactive residue	Surface accessibily of nitrogen atoms [Å^2] (PDB — 1AKK)
8–13	KIFVQK	834.53	Lys8	48.7
14–25	CAQCHTVEKGGK	1947.85	Lys22	33.8
28–38	TGPNLHGLFGR	1240.70	His33	19.4 (ND1); 1.1 (NE2)
39–53	KTGQAPGFTYTDANK	1670.90	Lys39	41.8
73–79	KYIPGTK	878.52	Lys73	21.4
80–87	MIFAGIKK	979.62	Lys86	11.5

Nomenclature of nitrogen atoms is according to PDB files.

3.2.4. Lysozyme

Using the protein:DEP ratio of 1:10 (0.7 mM DEP), 5 reactive lysine side-chains were identified (Table 3). The modification of His15, the only histidine in this protein, was only detected using the highest concentration of DEP (ratio 1:100). However, contrary to our mass spectrometry analysis, absorption spectrophotometry indicated histidine modification even for the lowest concentration used (ratio 1:1).

3.2.5. Human serum albumin

HSA was chosen to represent a "mid-sized" protein. Absorption spectrophotometry revealed that histidine residues were highly reactive in HSA. Using the highest concentration of DEP (ratio 1:100, concentration of DEP 1.5 mM) we observed a reaction of up to 10 histidine side-chains per molecule of HSA; it should be emphasized here that spectrophotometry is

Table 3
Surface accessibilities of reactive atoms in lysozyme

	Structure 1IO5		Structure 2CDS	
Residue	Surface accessibility NZ [Å^2]		Surface accessibility NZ [Å^2]	
Lys1-N_α	0.9		6	
Lys1-N_ε	2.5		10	
Lys13	14.5		23.6	
Lys33	14		26.6	
Lys97	4		16.7	
Lys116	24		16.9	
	Surface accessibility ND1 [Å^2]	Surface accessibility NE2 [Å^2]	Surface accessibility ND1 [Å^2]	Surface accessibility NE2 [Å^2]
His15	0	0.1	9.6	0.4

Table 4
Surface accessibilities of reactive atoms in HSA

	Stucture 1AO6		Structure 1UOR	
Residue	Surface accessibility NZ [Å^2]		Surface accessibility NZ [Å^2]	
Lys137	38; 32.9		4.2	
Lys205	42.7; 5.9		6	
Lys212	37.5; 23.3		11.4	
Lys262	48.4; 19.1		48	
Lys351	20.2; 20.6		11.9	
Lys414	8.4; 3.1		14.5	
Lys525	11.5;17.1		49.2	
	Surface accessibility NE2 [Å^2]	Surface accessibility NE2 [Å^2]	Surface accessibility NE2 [Å^2]	Surface accessibility NE2 [Å^2]
His67	2.5; 3.1	0.3; 0.1	16.9	0
His105	0.5; 5.8	12.6; 1.9	2.9	18
His128	9.3; 7.3	9.1; 12.6	3.4	31.4
His146	5.3; 1.4	18.1; 15.1	0.4	0
His367	9.3; 11.3	1.4; 3.6	4.3	0.5

Structure 1AO6 of HSA in dimer form; first value is for chain a, second one is for chain b.

considered to be only semi-quantitative method in this case. Using mass spectrometry it was only possible to prove with high reliability the modification of 5 histidine and 7 lysine residues (Table 4). The other potentially reactive histidines were found in the peptides containing lysine as well, but hydroxylamine treatment confirmed lysine modification only. Furthermore, the methods of peptide sequencing used to confirm histidine modification (MALDI-PSD MS, LC-ESI MS/MS) were inefficient in the case of HSA.

4. Discussion

4.1. Methodological aspects

Reaction of angiotensin with DEP revealed that the agent is non-specific even for its relatively low concentration in the mixture. Consequently, it was necessary to apply technique which allows to discriminate among reactive groups. We consider peptide sequencing by mass spectrometry and mainly hydroxylamine treatment as available and sufficient tools for this purpose. For higher concentration levels of DEP (angiotensin:DEP ratio of 1:20), disruption of imidazole ring was observed. That is why the careful choice of DEP concentration for modification experiment is fundamental. Level of agent should be chosen high enough for successful detection of modified peptides, on the other hand, exceedingly high DEP concentration can significantly affect studied 3-D structure of protein.

4.2. Structural interpretation

4.2.1. Insulin

We identified four reactive residues — Phe1, His5, His10 and Lys26. The ability of these residues to react is in strong agreement with the high surface accessibility of reactive atoms (Table 5). The reaction of DEP with insulin has already been described [20], but we found some discrepancies between the published results and our results. In the study cited, insulin was dissolved in ammonium acetate (pH 6.8), but we were unable to repeat this procedure; we were only able to dissolve insulin in $NaHCO_3$ at pH 9, with the pH then being adjusted to 6.8 by HCl. The fact that the authors used an insulin:DEP ratio of 1:50 is also questionable. At this concentration of DEP we observed irreversible disruption of the imidazol ring, which can affect protein spatial arrangement. Our study of the reaction of DEP with other proteins also revealed that such high concentrations of DEP increase the range of modifications, thereby making both spectra analysis and structural interpretation of the results difficult.

4.2.2. Cytochrom

All reactive residues contained nitrogen atoms with relatively high surface accessibility (Table 2), although we found other accessible nitrogen atoms that were non-reactive (i.e. Lys 87, 48 Å^2). It was also interesting to observe that the reactive side-chain of Lys86 had the lowest surface accessibility of all lysines in cytochrome. Nevertheless, His18 was identified as non-reactive which is in good agreement with the surface inaccessibility of this residue, which stabilizes hem by interaction with ferric ion. These results imply that surface accessibility is not the only important characteristic in determining the reactivity of amino acids with DEP.

4.2.3. Lysozyme

Studying the relationship between reactivity and surface accessibility was difficult in this case because we experienced essential discrepancies among the lysozyme PDB files (Table 3). In spite of this fact out data confirmed that only surface accessible residues of protein can react.

4.2.4. Human serum albumin

Again, structural interpretation of our results proved to be difficult because in certain cases both PDB files contained different characteristics for the same atom. Nevertheless, all modified residues contained surface accessible nitrogen atoms in at least one of the files (Table 4). Furthermore, two histidines (His39, His464) were found to be inaccessible in both files and both of them were non-reactive. Therefore we conclude that

Table 5
Surface accessibilities of reactive atoms in insulin (PDB file — 4INS)

Position	Surface accessibility of ND1 atoms [Å^2]	Surface accessibility of NE2 atoms [Å^2]
His5b	18.6; 0.8	2.3; 24.2
His10b	15; 14.3	29; 29.8
Position	Surface accessibility of N atoms [Å^2]	
Phe1b	39.1; 37.1	
Lys29b	14.9; 39.3	

Insulin crystalized in dimer form; in the table there are values of accessibilities for both monomers.

surface accessibility is a necessary condition for the reactivity of HSA with DEP.

5. Simplified description of method and its applications

In this paper, we tested whether DEP can be used for surface protein mapping. The method is technically relatively very simple. First, the native protein is modified by DEP. Then, the excess of reagent is removed by gel filtration and the modified protein is digested by specific protease (usually trypsin). Finally, mass spectra of the native and modified proteins are compared and typical shifts of molecular masses are detected. The main limitations of the method are the non-specificity of DEP reaction and the relative instability of histidyl ethoxyformylation.

Only surface-accessible lysyl and histidyl nitrogen atoms in model proteins were found to be able to react with DEP. Consequently, the surface accessibility is a necessary condition (although not the only one) for their reactivity. That is why the described method enables to determine important structural characteristics of proteins with unknown spatial arrangement.

Acknowledgement

The support of this work by the Czech Ministry of Education (grant number: 604 613 7305) is gratefully acknowledged.

References

[1] Coligan JE. Structural biology — introduction. In: Coligan JE, Dunn BH, Speicher DW, Wingfield DT, editors. Current protocols in protein science. Wiley & Sons Inc.; 2004. p. 3–4. Chapter 17.
[2] Lundblad RL. Chemical reagents for protein modification. 3rd edition. CRC Press; 2004.
[3] Carven CJ, Stern LJ. Probing the ligand-induced conformational change in HLA-DR1 by selective chemical modification and mass spectrometric mapping. Biochemistry 2005;44:13625–37.
[4] Calvete JJ, Rhodes MAC, Raida M, Sanz L. Characterisation of the conformational and quaterinary structure-dependent heparin-binding of bovine seminal plasma protein PDC-109. FEBS Lett 1999;444:260–4.
[5] Fiedler W, Borchers C, Macht M, Deininger SO, Przybylski M. Molecular characterization of a conformational epitope of hen egg white lysozyme by differential chemical modification of immune complexes and mass spectrometric peptide mapping. Bioconjug Chem 1998;9:236–41.
[6] Glocker MO, Borchers C, Fiedler W, Suckau D, Przybylski M. Molecular characterization of surface topology in protein tertiary structures by amino-acylation and mass spectrometric peptide mapping. Bioconjug Chem 1994;5:583–90.
[7] Strohalm M, Šantrůček J, Hynek R, Kodíček M. Analysis of tryptophan surface accessibility in proteins by MALDI-TOF mass spectrometry. Biochem Biophys Res Commun 2004;323:1134–8.
[8] Šantrůček J, Strohalm M, Kadlcík V, Hynek R, Kodícek M. Tyrosine residues modification studied by MALDI-TOF mass spectrometry. Biochem Biophys Res Commun 2004;323:1151–6.
[9] Scherrer P, Stoeckenius W. Selective nitration of tyrosines-26 and -64 in bacteriorhodopsin with tetranitromethane. Biochemistry 1984;23:6195–202.
[10] Sinz A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. J Mass Spectrom 2003;38:1225–37.
[11] Sharp JS, Becker JM, Hettich RL. Protein surface mapping by chemical oxidation: structural analysis by mass spectrometry. Anal Biochem 2003;313:216–35.
[12] Sharp JS, Becker JM, Hettich RL. Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectrometry. Anal Chem 2004;76:672–83.
[13] Hambly DM, Gross ML. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. J Am Soc Mass Spectrom 2005;16:2057–63.
[14] Maleknia SD, Chance MR, Downard KM. Electrospray-assisted modification of proteins: a radical probe of protein structure. Rapid Commun Mass Spectrom 1999;13:2352–8.
[15] Guan JQ, Vorobiev S, Almo SC, Chance MR. Mapping the G-actin binding surface of cofilin using synchrotron protein footprinting. Biochemistry 2002;41:5765–75.
[16] Schnaible V, Michels J, Zeth K, Freigang J, Weite W, Bühler S, et al. Approaches to the characterization of membrane channel proteins (porins) by UV MALDI-MS. Int J Mass Spectrom Ion Process 1997;169:165–77.
[17] Back JW, de Jong L, Muijsers AO, de Koster CG. Chemical cross-linking and mass spectrometry for protein structural modeling. J Mol Biol 2003;331:303–13.
[18] Sharp JS, Guo J, Uchiki T, Xu Y, Dealwis C, Hettich RL. Photochemical surface mapping of C14S-Sm1p for constrained computational modeling of protein structure. Anal Biochem 2005;340:201–12.
[19] Foti S, Marletta D, Saletti R. Fast-atom bombardment mass spectrometry of peptide derivates with diethylpyrocarbonate. Rapid Commun Mass Spectrom 1991;5:336–9.
[20] Kumagai H, Utagawa T, Yamada H. Studies on tyrosine phenol lyase. Modification of essentials histidyl residues by diethylpyrocarbonate. J Biol Chem 1975;250:1661–7.
[21] Miles W. Modification of histidyl residues in proteins by diethylpyrocarbonate. Methods Enzymol 1977;47:431–43.
[22] Altman J, Lipka JJ, Kuntz I, Waskeil L. Identification by proton nuclear magnetic resonance of the histidines in cytochrome b_5 modified by diethyl pyrocarbonate. Biochemistry 1989;28:7516–23.
[23] Kalkum M, Przybylski M, Glocker MO. Structure characterization of functional histidyl residues and carbethoxylated derivates in peptides and proteins by mass spectrometry. Bioconjug Chem 1998;9:226–35.
[24] Dage LJ, Sun H, Halsall HB. Determination of diethylpyrocarbonate-modified amino acid residues in α1-acid glycoprotein by high-performance liquid chromatography electrospray ionization-mass spectrometry and matrix-assisted laser desorption/ionization time-of-flight-mass spectrometry. Anal Biochem 1998;257:176–85.
[25] Internet, http://mmass.biographics.cz/, 21.3. 2007.
[26] Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol 1971;55:379–400.

Publication 5.1.2

A. Hnízda et al.

Cross-Talk between the Catalytic Core and the Regulatory Domain in Cystathionine β-Synthase: Study by Differential Covalent Labeling and Computational Modeling

Biochemistry, 2010

10526 *Biochemistry* **2010,** *49,* 10526–10534
DOI: 10.1021/bi101384m

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Cross-Talk between the Catalytic Core and the Regulatory Domain in Cystathionine β-Synthase: Study by Differential Covalent Labeling and Computational Modeling[†]

Aleš Hnízda,[‡] Vojtěch Spiwok,*[,§] Vojtěch Jurga,[§] Viktor Kožich,[‡] Milan Kodíček,[§] and Jan P. Kraus[∥]

[‡]*Institute of Inherited Metabolic Disorders, First Medical Faculty, Charles University in Prague and General University Hospital in Prague, Ke Karlovu 2, Prague 2, 128 00 Czech Republic,* [§]*Department of Biochemistry and Microbiology, Institute of Chemical Technology in Prague, Technická 5, Prague 6, 166 28 Czech Republic, and* [∥]*Department of Pediatrics, University of Colorado School of Medicine, Aurora, Colorado 80045, United States*

Received August 27, 2010; Revised Manuscript Received November 9, 2010

ABSTRACT: Cystathionine β-synthase (CBS) is a modular enzyme which catalyzes condensation of serine with homocysteine. Cross-talk between the catalytic core and the C-terminal regulatory domain modulates the enzyme activity. The regulatory domain imposes an autoinhibition action that is alleviated by *S*-adenosyl-L-methionine (AdoMet) binding, by deletion of the C-terminal regulatory module, or by thermal activation. The atomic mechanisms of the CBS allostery have not yet been sufficiently explained. Using pulse proteolysis in urea gradient and proteolytic kinetics with thermolysin under native conditions, we demonstrated that autoinhibition is associated with changes in conformational stability and with sterical hindrance of the catalytic core. To determine the contact area between the catalytic core and the autoinhibitory module of the CBS protein, we compared side-chain reactivity of the truncated CBS lacking the regulatory domain (45CBS) and of the full-length enzyme (wtCBS) using covalent labeling by six different modification agents and subsequent mass spectrometry. Fifty modification sites were identified in 45CBS, and four of them were not labeled in wtCBS. One differentially reactive site (cluster W408/W409/W410) is a part of the linker between the domains. The other three residues (K172 and/or K177, R336, and K384) are located in the same region of the 45CBS crystal structure; computational modeling showed that these amino acid side chains potentially form a regulatory interface in CBS protein. Subtle differences at CBS surface indicate that enzyme activity is not regulated by conformational conversions but more likely by different allosteric mechanisms.

Cystathionine β-synthase (CBS,[1] EC 4.2.1.22) is a pyridoxal 5′-phosphate (PLP) dependent enzyme which catalyzes the first step of the transsulfuration pathway, namely, the condensation of serine with homocysteine to cystathionine (*1*). Its deficiency due to missense mutations in the *CBS* gene is the most common cause of inherited homocystinuria, a treatable multisystemic disease affecting to various extent vasculature, connective tissues, and central nervous system (http://www.ncbi.nlm.nih.gov/omim/236200). More than 100 different pathogenic amino acid substitutions in the CBS protein were described, and the missense mutations represent 86% of all analyzed patient alleles (http://www.uchs.edu/cbs/cbsdata/cbsmain.htm).

Human CBS is a homotetrameric protein, and each subunit (61 kDa) consists of 551 amino acids. The protein sequence comprises three regions: the N-terminal heme-binding domain (1−69), a highly conserved catalytic core (70−413), and the C-terminal regulatory domain (414−551) (*2*), an autoinhibitory module with binding site for the allosteric activator, AdoMet (*3*).

CBS activity can be stimulated *in vitro* by several processes: by allosteric binding of *S*-adenosyl-L-methionine (AdoMet) (*3*), by proteolytic cleavage yielding the C-terminally truncated dimer containing identical subunits with molecular mass of 45 kDa (*4*), or by heat activation (*3*, *5*). Proteolytic activation of CBS was observed also *in vivo* in rat liver extract (*6*) and in HepG cell lines (*7*).

The spatial arrangement of CBS molecule was solved by X-ray crystalography for the truncated 45 kDa enzyme lacking the C-terminal regulatory domain (amino acids 1−413, 45CBS) only (*8*, *9*); the 3-D structure belongs to the β-family of PLP enzymes such as *O*-acetylserine sulfhydrylase or tryptophan synthase. However, the 3-D structure of the full-length CBS (wtCBS) has not yet been determined, and therefore the atomic basis of the enzyme regulation is still unclear. While hydrophobicity of the C-terminal module and putative interdomain motions prevented successful crystallization of wtCBS, alternative techniques can yield at least partial information about the allosteric communication in the wtCBS protein. Using H/D exchange, Sen et al. showed that the region 356−385 exhibited significantly slower rate of deuterium incorporation for wtCBS compared to 45CBS (*10*). The data were used for evaluation of a protein−protein docking exercise, and a structural model of the full-length CBS was proposed. However, this model has not yet been supported and/or refined by other structural techniques.

[†]This work was supported by the Welcome Trust International Senior Research Fellowship in Biomedical Science in Central Europe (Reg. No. 070255/Z/03/Z) and by grants from Grant Agency of Charles University in Prague (No. 0073/2010 and No. 260501). Institutional support was provided by the Research Projects from Ministry of Education of the Czech Republic (No. MSM0021620806 and No. MSM6046137305). J.P.K. was supported by NIH Grant HL065217, by American Heart Association Grant-in-Aid 09GRNT2110159, and by a grant from the Jerome Lejeune Foundation.

*Corresponding author. E-mail: spiwokv@vscht.cz. Telephone: +420 220 44 30 28. Fax: +420 220 44 51 67.

[1]Abbreviations: CBS, cystathionine β-synthase; PLP, pyridoxal 5′-phosphate; AdoMet, *S*-adenosyl-L-methionine; 45CBS, C-terminally truncated CBS; wtCBS, full-length CBS; DEP, diethyl pyrocarbonate; NAI, *N*-acetylimidazole; NEM, *N*-ethylmaleimide; NHS, sulfo-*N*-hydroxysuccinimido acetate; NBS, *N*-bromosuccinimide; HPG, 4-hydroxyphenylglyoxal.

 Published on Web 11/09/2010

In this study, we developed a procedure for covalent labeling of solvent-accessible amino acid residues (*11*) in purified CBS. Using this technique, we compared reactivity of the side chains in 45CBS and wtCBS with six modifiers. These commonly used compounds specifically react with histidines (diethyl pyrocarbonate; DEP), tyrosines (*N*-acetylimidazole; NAI), cysteines (*N*-ethylmaleimide; NEM), lysines (sulfo-*N*-hydroxysuccinimido acetate; NHS), tryptophans (*N*-bromosuccinimide; NBS), and arginines (4-hydroxyphenylglyoxal; HPG) (*12*). Surface mapping provided data which faciliated development of the refined model for wtCBS spatial arrangement and enabled insight into the structural basis of the enzyme allosteric regulation.

EXPERIMENTAL PROCEDURES

Materials. If not specified otherwise, all chemicals were purchased from Sigma-Aldrich.

Preparation of 45CBS and wtCBS. The 45CBS and the wtCBS were expressed in *Escherichia coli* and purified to homogenity as previously described (*13*, *14*).

Pulse Proteolysis. Pulse proteolysis was performed as described previously (*15*, *16*) with some modifications. Purified 45CBS or wtCBS (0.5 mg/mL) was equilibrated overnight at 4 °C in 20 mM Tris-HCl (pH 8.0) containing 10 mM $CaCl_2$ and urea (0−7 M) and then digested by thermolysin from *Bacillus thermoproteolyticus* (0.1 mg/mL). To carry out pulse proteolysis of wtCBS in the presence of AdoMet, wtCBS was incubated with 300 μM AdoMet at room temperature for 10 min prior to equilibration in urea. The proteolytic pulse (1 min) was quenched in 20 mM EDTA. Protein samples (7.5 μg) were analyzed by SDS−PAGE using Tris−acetate SDS running buffer with 3−8% gradient Tris−acetate precast gels (Invitrogen) and visualized by Coomassie blue solution. Experiments were repeated three times. Band intensities were quantified using GeneTools software (Syngene) and were fitted into the sigmoidal equation:

$$f_{\text{fold}} = \frac{1}{1 + e^{p(c_m - c)}}$$

using Origin 8.0 (Originlab); f_{fold} represents a fraction of folded proteins remaining intact after proteolytic pulse, c_m urea concentration at which f_{fold} is 0.5, and c urea concentration. Value of p is a slope of curve at c_m, and it reflects unfolding cooperativity.

Proteolytic Kinetics under Native Conditions. Purified proteins (0.5 mg/mL) were diluted in 20 mM Tris-HCl (pH 8.0) containing 10 mM $CaCl_2$ and digested by thermolysin (0.1 mg/mL). At the chosen time point, proteolysis was quenched in 20 mM EDTA. SDS−PAGE and band quantification were performed as described for pulse proteolysis. First-order kinetic constant of proteolysis (k_p) for each protein was determined by nonliner curve fitting (*17*).

Preparation of Modified Protein Samples. CBS proteins (1 mg/mL) were diluted in modification buffer and covalently labeled. Each labeling procedure (*18*−*23*) (Table S1 in the Supporting Information) was repeated three times. Reaction was quenched by buffer exchange using Zeba Desalt spin columns (ThermoFischer Scientific) with elution by 50 mM NH_4HCO_3.

Analysis of Modified Proteins. (*i*) *Native Electrophoresis*. Labeled proteins (3 μg) were separated using Laemmli buffer system without sodium dodecyl sulfate and with 3−8% gradient Tris−acetate precast gel at 4 °C and visualized by silver staining kit (Promega) according to manufacturer's manual.

(*ii*) *CBS Activity Assay*. Enzyme activity of the proteins was determined in the absence or presence of 1 mM AdoMet by radiometric assay using [^{14}C]-L-serine (Moravek Biochemicals); the previously described method (*24*) was slightly modified. The reactants and products were separated by thin-layer chromatography using cellulose−HPTLC sheets (Merck) and subsequently visualized using PhosphorImager system (Molecular Dynamics); amount of radioactive cystathionine as the reaction product was determined by ImageQuant 5.0 software (Molecular Dynamics).

(*iii*) *In-Solution Proteins Digestion and Mass Spectrometric Analysis*. Labeled proteins were reduced in 5 mM dithiothreitol at 50 °C for 30 min; reduced cysteines were acetamidated in 25 mM iodoacetamide in the dark at room temperature for 30 min. Subsequently, they were digested by trypsin (Promega), chymotrypsin, endoprotease Glu-C, and protease double combinations (*25*) at 37 °C for 1 h. The CBS:protease ratio (w/w) was 20:1. The protein digest was fractionated by ZipTip (Millipore), and each fraction was mixed with the matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid supplied by Bruker Daltonics, sample:matrix ratio of 1:1 v/v) and measured using Autoflex II (Bruker Daltonics) mass spectrometer equipped with a nitrogen laser (337 nm) in reflector positive mode (m/z range from 500 to 4500). The mass spectrometer was externally calibrated by peptide calibration standard II (Bruker Daltonics). All spectra were processed by Flex Analysis, Biotools 3.0 and mMass 3.0 (*26*); mass accuracy tolerance was set at 50 ppm for MS and ±0.5 Da for MS/MS analyses (*22*).

With the exception of labeling with NBS, all other modification sites were identified by detection of labeled peptides that were not detected in unmodified controls (*27*); expected mass shifts for each reaction are shown in Supporting Information Table S1. The labeling with NBS induces tryptophan oxidation (*19*) which is also considered to be a common artifact of sample handling (*28*). Since we observed tryptophan oxidation even in the unmodified controls, the residues labeled with NBS were determined by comparing peak intensities of the modified and the unmodified peptides (*29*). Tryptophan residues were classified as labeled if the relative intensity of modified peptide increased at least 1.5-fold compared to the unmodified control. Identity of the modified peptides generated from all labeling experiments was confirmed by MS/MS measurements (method LIFT).

In general, mass spectrometric measurements were satisfactorily reproducible; i.e., modification sites were determined identically in the repeated experiments.

Thermal Activation of wtCBS. The wtCBS diluted in the reaction buffer was incubated at 55 °C for 10 min and then chilled on ice (*3*, *5*). Thermally activated proteins were labeled and analyzed by native electrophoresis, activity assay, and mass spectrometry as described above.

Protein Structure Modeling. Model of the C-terminal regulatory module was built by homology modeling package Modeller 9v3 (*30*) using the structure of CBS-domain containing protein MJ0100 from *Methanocladococcus jannaschii* (PDB ID 3KPB) (*31*) as a template. The initial sequence−sequence alignment was processed by the web services of PHYRE (*32*) and PSI-BLAST (*33*) and further modified manually. The resulting model was evaluated using Prosa web service (*34*) and statistical coupling/protein sector analysis (*35*). For this purpose, 6983 protein sequences from CBS subfamily were taken from the Pfam database (*36*) and analyzed using a Python script based on the procedure introduced by Halabi and co-workers (*35*).

Model of wtCBS dimer was obtained by docking of a single C-terminal regulatory domain to 45CBS dimer (PDB ID 1JBQ, with missing loops reconstructed by Modeller package) using the

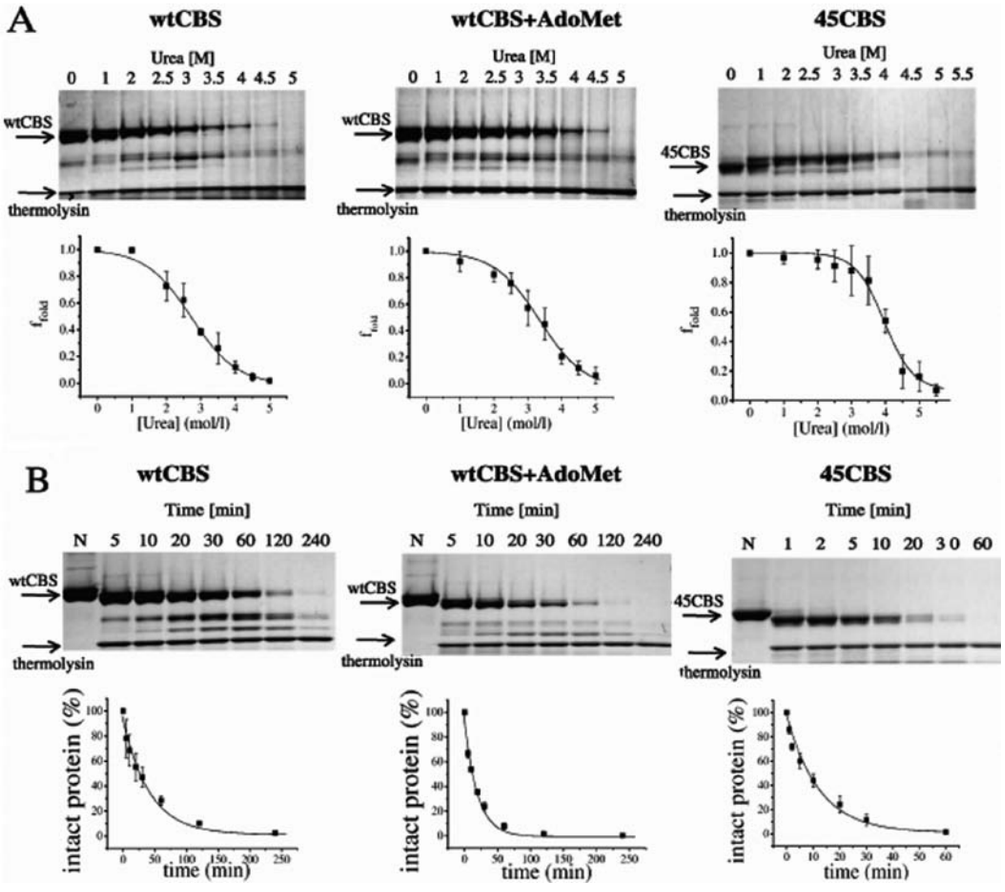


FIGURE 1: Pulse proteolysis in urea gradient (A) and proteolytic kinetics by thermolysin under native conditions (B) of CBS proteins. Below the representative SDS−PAGE gels, the corresponding plots are shown. Points are depicted as a mean with standard deviations; curves were fitted by nonlinear regression. (A) Molar concentration of urea for proteolytic pulse is indicated at the top of each line at the gels. F_{fold} values which represent fraction of remaining intact protein after the proteolytic pulse are plotted against urea concentration. (B) Portion of remaining protein is plotted against the incubation time. Each line of the gels is marked by designed time point of proteolysis in minutes; "N" refers to uncleaved control.

program ZDOCK (*37*). This was followed by addition of the C-terminal domain to the second subunit that was driven by symmetry. Differentially modified residues from the experimental results were forced to be involved in an interdomain interaction during the docking process.

Structural models generated by this approach were visually inspected on the basis of several criteria, namely, involvement of differentially modified resides in the interaction, dimer symmetry, and protein stereochemistry. The full-length dimer was built by Modeller using the best-suited structure from docking procedures.

RESULTS

Pulse Proteolysis and Proteolytic Kinetics under Native Conditions. Using pulse proteolysis we determined the global conformational stability and unfolding cooperativity of CBS proteins (Figure 1 and Table 1). The wtCBS exhibited lower resistance to urea-induced denaturation and lower degree of unfolding cooperativity compared to 45CBS. Binding of AdoMet to wtCBS moderately increased the protein stability toward urea, although it remained lower than the 45CBS resistance. On the other hand, unfolding cooperativity of wtCBS did not differ from wtCBS in the presence of AdoMet. These data show that CBS proteins adopt variant conformational states characterized by different degree of the stability.

Proteolytic kinetics by thermolysin under native conditions (Figure 1B) revealed slower cleavage of wtCBS compared to the 45CBS. AdoMet binding to wtCBS accelerated proteolysis; however, it was still slower than cleavage of 45CBS.

Table 1: Results from Pulse Proteolysis in Urea Gradient and Proteolytic Kinetics under Native Conditions[a]

	pulse proteolysis		proteolytic kinetics under native conditions
protein	c_m (mol/L)	p	k_p (min^{-1})
wtCBS	2.70 ± 0.08	1.51 ± 0.15	0.026 ± 0.005
wtCBS + AdoMet	3.26 ± 0.08	1.48 ± 0.13	0.056 ± 0.005
45CBS	4.08 ± 0.07	2.20 ± 0.28	0.075 ± 0.008

[a]The CBS proteins (0.5 mg/mL) were digested with thermolysin (0.1 mg/mL). Data were evaluated by nonlinear curve fitting. Value of c_m reflects conformational stability, and value of p is informative about unfolding cooperativity; the constant k_p was acquired from the equation of first-rate kinetics.

These results are concordant with previously proposed regulation mechanisms; i.e., catalytic core is sterically hindered in the full-length protein by the C-terminal regulatory domain, and the hindrance is partially cleared off upon AdoMet binding (*3*, *38*). To verify this hypothesis at the atomic level, we compared three-dimensional structures of 45CBS and wtCBS using protein surface mapping.

Surface Mapping of CBS. (*i*) *Sequence Coverage*. To reach high degree of protein sequence coverage, 45CBS and wtCBS were digested by three proteolytic enzymes, namely, chymotrypsin,

endoprotease Glu-C, and trypsin and by their double combinations; we obtained the sequence coverage of unmodified proteins 89% and 94% for 45CBS and wtCBS, respectively. For each labeling experiment, we selected the digests that yielded the highest amount of reliably identified modified amino acid residues (mass spectrometric data set available in the Supporting Information).

(*ii*) *Modifier Concentrations*. In the next step, we optimalized conditions of each labeling reaction as the excess of a modification agent may disrupt the spatial arrangement of a protein (*39*). Therefore, the lowest concentration of the modifier that enabled efficient mass spectrometric detection of the modified residues was chosen. The integrity of modified proteins was monitored by disruption of their structure manifested by smears and lack of sharp bands on native gels along with complete loss of enzymatic activity. These effects were observed in the case of labeling with tyrosine modifiers tetranitromethane and iodine, and tryptophan modifier 2-hydroxy-5-nitrobenzyl bromide (data not shown). Six other modification agents (Supporting Information Table S1) were feasible for this study since modified CBS proteins migrated as sharp bands on native gels and retained high levels of enzymatic activity (Figure S1 and Table S2 in the Supporting Information). These data showed that most of the modification reactions did not even partially disturb integrity of CBS proteins, with the exception of the 45CBS labeled with NBS. In this case, modification procedure decreased enzyme activity to 43% of the unmodified control. This observation indicated that the labeling reaction may partially affect the catalytic activity. Despite this obstacle, we utilized NBS labeling since it was the only suitable compound for the detecting of solvent-exposed tryptophans. The eventual impact of the modification procedure on the protein integrity should be thus taken into account during structural interpretation.

Modification Sites in CBS. Modification reactions were examined by mass spectrometry, and residues labeled by six different agents were determined. The labeling was monitored qualitatively; i.e., the evaluation was based on the presence/absence of the modified peptides in 45CBS and wtCBS. This approach is commonly known as chemical footprinting (*40*), a suitable technique for study of protein/protein and protein/DNA interactions (*41*).

Mass spectrometric analysis revealed 50 and 70 modification sites in 45CBS and wtCBS, respectively (Table 2). Identity of the modified peptides was verified by MS/MS sequencing. However, several sites could not be confirmed due to insufficient fragmentation of the modified peptides (see Table 2). The majority of the unconfirmed peptides contain a modified arginine residue since their tagging may affect the fragmentation process as previously reported (*42*). Nevertheless, these peptides were included in the data set, since their observed masses were unambiguously assigned against *in silico* generated digests.

In wtCBS, 46 labeled residues were identified in the active core (region 1−413), and 24 sites were located in the regulatory domain (414−551). Comparing the side-chain reactivity of 45CBS and wtCBS in the region 1−413, we found four sites that were differentially labeled, i.e., modified in 45CBS and not wtCBS (Figure 2; MS/MS spectra are shown in Figures S2−S4 in the Supporting Information). Differentially modified amino acid side chains were found in the peptide 164−181 (residue K172 and/or K177 modified by NHS), in the peptide 326−345 (residue R336 modified by HPG), in the peptide 380−389 (residue K384 modified by NAI), and in the peptide 406−413 (residue W408 and/or W409 and/or W410 modified by NBS).

Table 2: Modification Profile of CBS: List of Modified Amino Acids[a]

modifier					
DEP	NBS	NEM	NAI	NHS	HPG
G1	W43	C15	G1	G1	R18
C15	W54	C52	K25	K25	R45
H17	W208	C272	K39	K39	R182/R190/R196[b]
H22	W408/W409/W410	C370	K72	K72	R209
H65/H66/H67	M505[b]	C431	K137	K137	R336[b]
H203	M529[b]		K172/K177[c]	K172/K177	R369[b]
K211			K211	K211	R389
K406			K271	K271	R413
H411			Y308	K322	R439
H433			K322	K405	R491[b]
H501			K359	K406	R498
H507			K384	K441	R527
			K398/K405[b]	K472	R548
			K406	K481	
			K441	K485	
			K472	K488	
			Y484/K485	K523	
			K488	K551	
			K523		
			K551		

[a]Differentially reactive residues (modified in 45CBS but not in wtCBS) are underlined. [b]Identity of modified peptide could not have been confirmed by MS/MS due to insufficient fragmentation. [c]Reactivity of these residues could have been confirmed by MS/MS only in the case of 45CBS.

Differentially Reactive Peptides in Thermally Activated CBS. Lack of reactivity of the above four peptides in wtCBS could be explained by interdomain sterical hindrance that is independent of the regulatory motions or by conformational changes which modulate the enzymatic activity. Thus, we tested whether the reactivity of the residues can be restored by stimulating activity of wtCBS. If so, such a result would suggest that the residues are involved in conformational motions; on the other hand, persistent unreactivity of these side chains would indicate their location at the fixed interdomain interface. Since the surface mapping in the presence of AdoMet could not be performed due to this ligand's reactivity toward most of the modifiers, thermally activated wtCBS was analyzed as a surrogate. This approach is feasible since allosteric changes due to AdoMet binding and partial heat denaturation share a common mechanism (*3*).

For covalent labeling of the stimulated wtCBS we applied only the modifiers and the digestions which provided differentially reactive peptides. Structural integrity of thermally activated wtCBS was preserved after the labeling reactions to an extent similar to the nondenatured wtCBS at the same concentration of modifier (enzyme activities are shown in Table S3, Supporting Information). The restoration of the residue reactivity was observed only for peptide 164−181 labeled by NHS, while the other three peptides were not labeled in thermally activated wtCBS (Figure 2). These findings indicate that both sterical hindrance and regulatory motions are responsible for the differential reactivity of the residues.

Structural Prediction Using Computational Modeling. Initially, homology model of the C-terminal regulatory domain was built using archeal CBS domain as a template (*31*). The amino acid sequence identity was 16% for the template-model pair. Nevertheless, CBS domains form a conserved tertiary structure despite rather low sequence identity of individual proteins (*43*).

A

Modifier	Peptide	Protease	Sequence	Modification site	Calculated molecular weight	Error of observed weight in ppm		
						45CBS	wtCBS	thermally activated wtCBS
NBS	406-413	Trypsin	KPWWWHLR	W408 and/or W409 and/or W410	1224.63	-24	N.D.	N.D.
NAI	380-389	Trypsin	NYMTKFLSDR	K384	1316.63	0	N.D.	N.D.
NHS	164-181	Chymotrypsin	RCIIVMPEKM SSEKVDVL	K172 and/or K177	2176.12	24	N.D.	-27
HPG	326-345	Trypsin+ chymotrypsin	SNDEEAFTFAR MLIAQEGLL	R336	2387.34	-3	N.D.	N.D.

B

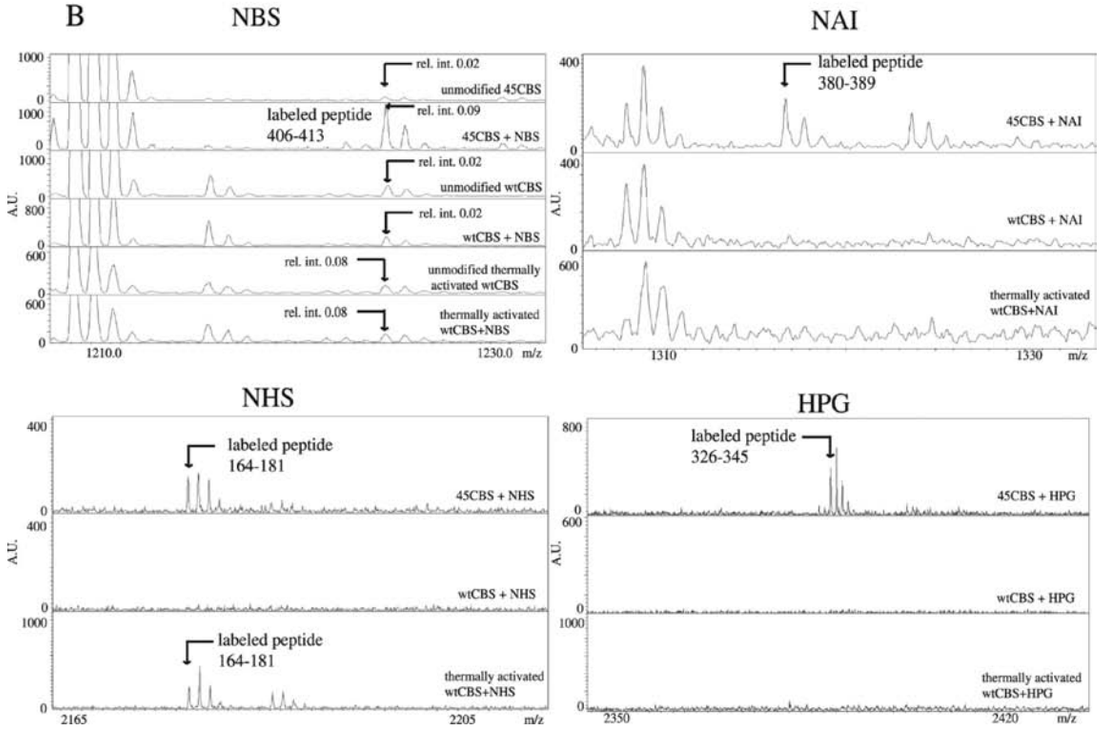


FIGURE 2: Differentially reactive peptides and their modification sites (A) together with corresponding representative spectra (B). Reactivity of the peptides is shown in 45CBS, wtCBS, and thermally activated wtCBS.

Since this structural module contains many flexible loops, additional restraints were applied; namely, residues 466−472 and 537−549 of the CBS were forced to α-helix formation according to secondary structure prediction, and the distance between C_α atom of L412 and CD2 of L539 was restrained to 4 Å according to structural data of the template.The resulting structure was evaluated by Prosa and yielded a value of −6.38 which is comparable to values usual for experimental structures of the same size and similar to the score for the single subunit of the template (−7.50). Moreover, statistical coupling/protein sector analysis was used for evaluation of the model. Protein sectors are coevolving networks of residues supposed to play a common role (i.e., catalytic, stabilizing etc.) and thus showing a spatial proximity. Two sectors with evolutionary coupled residues were identified in the autoregulatory domain (Figure 3A), showing a strong coevolution within each sector but a loose one between each other (Figure 3B). The residues from the particular sector were found next to each other in the homology model which indicated a high plausibility of the resulting structure. In addition, the residue I483 was located at the sector interface and revealed strong coupling with both sectors (further details on SCA/sectors can be obtained in the Supporting Information).

In the next step, the modeled C-terminal domain (residues 410−545) was docked onto the available structure of 45CBS. The initial docking was not successful, indicating that certain conformational changes in the catalytic domain may be associated

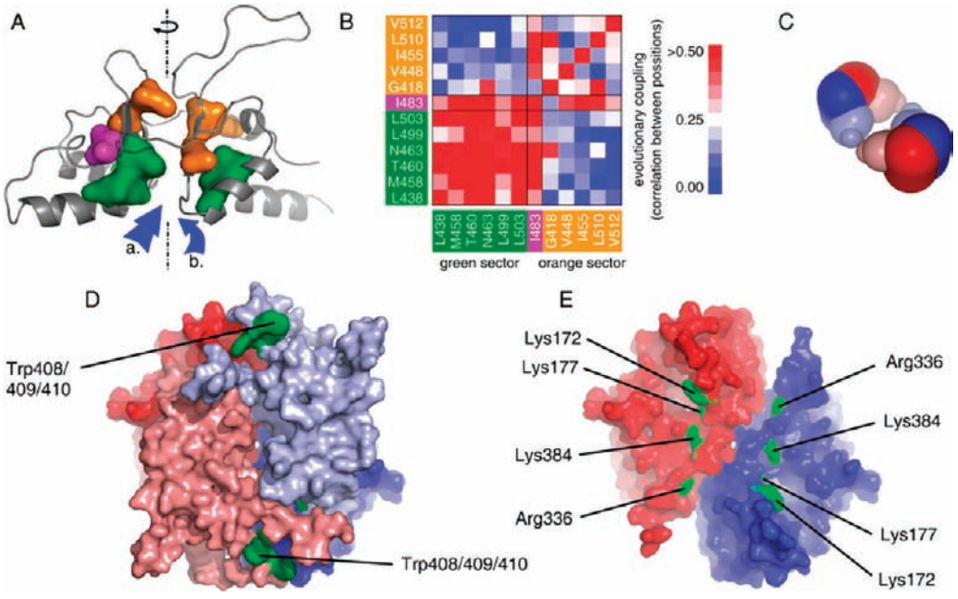


FIGURE 3: Computational modeling of CBS structure. (A) Model of C-terminal domain generated by homology modeling. Reliability of the built structure was assessed by protein sector analysis. Each sector is depicted by its particular color (green and orange, respectively); residue I483, coupled in both sectors, is indicated in magenta. The dashed line indicates the axis of pseudo-2-fold symmetry of the subunit; arrows show the potential binding sites for AdoMet. (B) Statistical coupling between sector residues. It illustrates that these positions in the structure of the autoregulatory domain are strongly coupled within each sector but loosely coupled between the two sectors. Colors of the sectors are consistent with panel A. (C) Scheme of tetrameric assembly in CBS using available structural data. Dimer−dimer interface is located between the autoinhibitory domains. Dimers of catalytic core are colored in dark color, autoinhibitory modules are depicted in light colors. (D) Structural model of dimeric wtCBS. Position of differentially reactive cluster W408/9/10 is indicated in green. Each subunit is depicted in particular color, red and blue, respectively. Autoinhibitory module is colored in light colors; catalytic core is depicted darkly. (E) Differentially reactive residues located in crystal structure of 45CBS, indicated in green. Each subunit in dimer is colored in blue and red, respective.

with the binding of the regulatory domain. Therefore, the sterically hindering C-terminal helical region of the catalytic domain (residues 385−397) was deleted, and the truncated structure was used as bait with residues K172, K177, and R336 being forced to interaction. Other differentially reactive residues (K384 and W408/409/410) were not involved in docking as they are located in the linker between the catalytic and autoregulatory domains; their topology was used rather for verification of the resulted models. This modified docking procedure resulted in the generation of 79 structures; model no. 32 was selected by visual inspection on the basis of the location of the differentially modified residues, protein symmetry, and general stereochemistry. Using the result from the docking procedures, the structure of the full-length dimer was built. Plausibility of the structural model is greatly supported by data from surface mapping experiments: differentially reactive residues are located at the regulatory interface while residues modified in 45CBS as well as in wtCBS are still solvent-accessible (see Figure 3D,E; structural model is available in the Supporting Information). However, the resulting model represents a possible structural interpretation of our experimental data and should not be interpreted as an atom-resolved structure due to limitations of homology modeling and protein−protein docking procedures.

DISCUSSION

Regulatory Interface in CBS. A cross-talk between the active core and the regulatory domain in CBS modulates its enzyme activity. The main aim of the study was to compare residue reactivity in 45CBS and wtCBS as the differences may reveal the regulatory network. In 45CBS, we identified 50 labeled residues in total, and we found only 4 modification sites which were not detected in wtCBS (Table 2). Using the thermally activated wtCBS as a surrogate of the AdoMet activated enzyme, we tested whether the abolished side-chain reactivity could be restored by the allosteric stimulation. The only differentially reactive peptide 164−181 was labeled in the thermally activated wtCBS, suggesting that this region (namely, residues K172 and/or K177) increases surface accessibility during enzyme stimulation and that it is involved in regulatory motions of CBS. Three other differentially reactive peptides were not labeled in wtCBS even upon thermal activation. Therefore, these side chains (R336, K384, W408 and/or W409 and/or W410) are probably localized at the fixed domain interface.

Contact Area between the Catalytic Core and the Regulatory Domain. Three differentially modified residues (K172 and/or K177, R336, K384) were located in the same region of the 45CBS crystal structure (see Figure 2E), and docking procedure showed that these residues may form an interface between the catalytic core and the regulatory domain. The region possessing differentially reactive sites was also predicted to form interdomain contact area due to the presence of hydrophobic residues on the surface of the 45CBS crystal structure (*44*). Several CBS patient-derived mutations, namely, the p.V173M (*45*), the p.E176K (*46*), and the p.E302K (*47*), which are located at this putative interface, exhibited enzyme activity similar to wtCBS and failed to be allostericaly stimulated by AdoMet. These observations indicate that mutations of these residues affect interdomain interactions and the CBS allostery.

Another differentially reactive site, the tryptophan cluster W408/9/10, was not previously assigned by the diffraction analysis of the 45CBS crystal; thus we propose that it forms a flexible region in 45CBS and a loop between the active core and the C-terminal domain which is sterically hindered in the wtCBS. As mentioned in the Results, findings dealing with the residues W408/9/10

should be taken with care as modification of 45CBS with NBS led to apparent decrease in enzymatic activity. On the other hand, the electromigration of modified 45CBS was undistinguishable from the unmodified control, indicating that quaternary structure was preserved after the labeling. We assume that the protein structural integrity was not essentially disrupted and that the enzyme activity was affected due to the local conformational changes. Moreover, conclusions about different microenvironment along the tryptophan cluster are also supported by changes in tryptophan flourescence spectra reported previously (*3*, *4*).

However, a previously published study involving H/D exchange (*10*) revealed the interdomain contact at a different region of the CBS structure. Although the changes in microenvironment of K384 were observed by both the H/D exchange and the covalent labeling in the present study, other differentially solvent-accessible regions were found using just single technique. We observed changes in residues K172/K177, R336, and W408/W409/W410, but they were not reported by Sen et al. On the contrary, H/D exchange study revealed differences in the segment of 359−370, but our covalent labeling experiments did not confirm them; in this region, three modification sites (K359, R369, and C370) were identically observed in the both proteins, 45CBS and wtCBS. Similarly to our study, results from H/D exchange were further supported by properties of certain mutant proteins, namely of double-linked mutant p.P78R/K102N (*48*). Its amino acid substitutions are located in the proximity of the differentially solvent-accessible region 359−370, and this mutant affects the protein allostery driven by AdoMet binding.

The discrepancies between results of covalent labeling and H/D exchange are unclear. The inconsistency may reflect the methodological limitations of each technique. Our experimental setup was designed for identification of differentially reactive sites rather than for quantification of small changes in extent of modification. Mass spectrometry analyses of the reactions were performed qualitatively (with exception of labeling with NBS; see Experimental Procedures) which enabled determination of totally blocked residues only. On the other hand, we might have lost information about subtle conformational motions that would be revealed by quantitative evaluation. Conformational study using H/D exchange has its own limitations as well. It determines the rate of deuterium incorporation to protein backbone from several seconds to hours, and consequently any differences on a short time scale of the exchange may be missed. Therefore, each of these two approaches might locate only particular changes in the CBS protein. Unfortunately, an attempt to generate a model consistent with both data sets was not successful (data not shown). We can speculate that the discrepancies between these studies might arise from different conditions and procedures during preparation of CBS proteins. Consequently, each study would have analyzed only limited set of all possible states from the conformational ensemble. Nevertheless, the inconsistency needs to be examined by additional structural techniques.

Allostery of CBS Is Not Associated with Extensive Conformational Changes. Covalent labeling as well as H/D exchange showed that autoinhibition of the active core by the regulatory domain is associated with only subtle changes at the protein surface. These observations indicate that the CBS allostery is not necessarily directed by extensive conformational motions, suggesting that other factors may play an important role. Changes in structural flexibility and "population shift" as determinants for protein allostery were proposed in the past decade (*49*−*51*); it has been shown that the ligand binding often leads to stabilization and/or rigidification of certain conformations (*52*). As the enzyme activity of CBS proteins (*14*) is directly proportional to the conformational stability, as determined by pulse proteolysis (Table 1; 45CBS, wtCBS in the presence of AdoMet, wtCBS in the absence of AdoMet, in descending order), it is tempting to speculate that CBS regulation may be driven by changes in protein dynamics. However, detailed knowledge of this type of CBS allostery is limited since the 3-D structure of the protein has not yet been reliably described in sufficient resolution.

AdoMet Binding Site. Furthermore, the designed model of wtCBS provides information about the structural features of several sites with putative regulatory function. Since the spatial arrangement of archeal CBS-domain pair in complex with AdoMet was solved recently (*31*) and we used this structure as a template for homology modeling of the C-terminal regulatory domain, the possible AdoMet binding site can be proposed. An interesting feature of the C-terminal autoinhibitory domain is its pseudo-2-fold symmetry (the axis indicated in Figure 3A) which forms the basis for two ligand binding sites in each regulatory subunit (a and b in Figure 3A). The experimental structures of the template−ligand complexes showed that the ligands bind to either one of these sites. Sequence similarity does not provide enough information to precisely identify the AdoMet binding site in CBS. However, AdoMet is likely bound in site b (Figure 3A) including the residue D444 that has been identified to be involved in the autoinhibitory function (*38*).

CXXC Oxidoreductase Motif. CBS also contains the CXXC oxidoreductase motif which spans residues 272−275. Here we identified the C272 as a solvent-exposed residue both in the 45CBS and in the wtCBS. This observation disagrees with a previous study that used three different cysteine modifying agents and an N-terminal sequencing of carboxymethylated peptides (*5*). However, our findings are in agreement with the crystal structure of the 45CBS. The solvent accessibility of CXXC motif observed in our study may thus support the notion of its possible role in redox sensing (*53*), although the biological relevance of this observation remains to be answered.

Residues Responsible for Aggregation and Allostery. Other residues, which play important role in CBS function, were revealed by labeling with NAI; this modification decreased the tendency of wtCBS to form higher order oligomers and increased its catalytic activity (Table S2 and Figure S1 in the Supporting Information). Similar effect was also observed after modification by NEM as reported previously (*5*). Frank et al. explained the stabilizing action of the NEM by covalent blocking of C15, the residue responsible for aggregation of wtCBS.

Interestingly, wtCBS labeled with NAI failed to be fully activated upon AdoMet binding while modification of wtCBS by NHS, which exhibited similar modification pattern as NAI (Table 2), did not cause such effects. These data indicate that certain modified residues are responsible for CBS aggregation and also for allosteric activation, and their function can be repressed by covalent blocking of the reactive groups. Comparing the results from labeling with NAI and NHS, we can point out three candidate residues, namely, Y308, K359, and Y484, that are modified by NAI and not by NHS. However, we could not specify the "aggregation inducing" and "regulation networking" side chains in this study.

Quarternary Structure of wtCBS. The relevance of the structural model proposed in this paper is limited as dimeric full-length CBS does not explain the atomic basis of the protein tetramerization. Our results from surface mapping revealed a single

contact area between the catalytic core and the C-terminal regulatory domain; this is in agreement with the solved structure of the dimeric protein MJ0100 from *M. jannaschii* containing CBS pairs binding AdoMet (*31*) and suggests that the autoinhibitory module contains dimer−dimer interface responsible for the CBS tetramer assembly (scheme in Figure 3C). This proposal is also in agreement with the previously built structural model of wtCBS derived from H/D exchange (*54*).

In summary, we covalently labeled solvent-exposed side chains in CBS, and we identified the interface between the active core and the regulatory domain. The data were applied for generation of the refined full-length CBS structural model. Our results also indicate that the allostery of CBS is not associated with extensive conformational conversion but rather with changes in protein dynamics.

ACKNOWLEDGMENT

We thank Petr Prikryl, Ph.D., for helpful assistance with the mass spectrometry measurements.

SUPPORTING INFORMATION AVAILABLE

(i) Detailed results from the labeling procedures and from the statistical coupling analysis of the autoinhibitory domain; (ii) an overview of the modified peptides that were identified by mass spectrometry; (iii) the generated structural model for the wtCBS illustrating the experimental data. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Mudd, S. H., Finkelstein, J. D., Irreverre, F., and Laster, L. (1965) Threonine dehydratase activity in humans lacking cystathionine synthase. *Biochem. Biophys. Res. Commun. 19*, 665–670.
2. Oliveriusova, J., Kery, V., Maclean, K. N., and Kraus, J. P. (2002) Deletion mutagenesis of human cystathionine beta-synthase. Impact on activity, oligomeric status, and S-adenosylmethionine regulation. *J. Biol. Chem. 277*, 48386–48394.
3. Janosik, M., Kery, V., Gaustadnes, M., Maclean, K. N., and Kraus, J. P. (2001) 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*, 10625–10633.
4. Kery, V., Poneleit, L., and Kraus, J. P. (1998) Trypsin cleavage of human cystathionine beta-synthase into an evolutionarily conserved active core: structural and functional consequences. *Arch. Biochem. Biophys. 355*, 222–232.
5. Frank, N., Kery, V., Maclean, K. N., and Kraus, J. P. (2006) Solvent-accessible cysteines in human cystathionine beta-synthase: crucial role of cysteine 431 in *S*-adenosyl-L-methionine binding. *Biochemistry 45*, 11021–11029.
6. Skovby, F., Kraus, J. P., and Rosenberg, L. E. (1984) Biosynthesis and proteolytic activation of cystathionine beta-synthase in rat liver. *J. Biol. Chem. 259*, 588–593.
7. Zou, C. G., and Banerjee, R. (2003) Tumor necrosis factor-alpha-induced targeted proteolysis of cystathionine beta-synthase modulates redox homeostasis. *J. Biol. Chem. 278*, 16802–16808.
8. Meier, M., Janosik, M., Kery, V., Kraus, J. P., and Burkhard, P. (2001) Structure of human cystathionine beta-synthase: a unique pyridoxal 5′-phosphate-dependent heme protein. *EMBO J. 20*, 3910–3916.
9. Taoka, S., Lepore, B. W., Kabil, O., Ojha, S., Ringe, D., and Banerjee, R. (2002) Human cystathionine beta-synthase is a heme sensor protein. Evidence that the redox sensor is heme and not the vicinal cysteines in the CXXC motif seen in the crystal structure of the truncated enzyme. *Biochemistry 41*, 10454–10461.
10. Sen, S., Yu, J., Yamanishi, M., Schellhorn, D., and Banerjee, R. (2005) Mapping peptides correlated with transmission of intrasteric inhibition and allosteric activation in human cystathionine beta-synthase. *Biochemistry 44*, 14210–14216.
11. Suckau, D., Mak, M., and Przybylski, M. (1992) Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping. *Proc. Natl. Acad. Sci. U.S.A. 89*, 5630–5634.
12. Mendoza, V. L., and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev. 28*, 785–815.
13. Janosik, M., Meier, M., Kery, V., Oliveriusova, J., Burkhard, P., and Kraus, J. P. (2001) Crystallization and preliminary X-ray diffraction analysis of the active core of human recombinant cystathionine beta-synthase: an enzyme involved in vascular disease. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 289–291.
14. Frank, N., Kent, J. O., Meier, M., and Kraus, J. P. (2008) Purification and characterization of the wild type and truncated human cystathionine beta-synthase enzymes expressed in E. coli. *Arch. Biochem. Biophys. 470*, 64–72.
15. Park, C., and Marqusee, S. (2005) Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. *Nat. Methods 2*, 207–212.
16. Prudova, A., Bauman, Z., Braun, A., Vitvitsky, V., Lu, S. C., and Banerjee, R. (2006) S-adenosylmethionine stabilizes cystathionine beta-synthase and modulates redox capacity. *Proc. Natl. Acad. Sci. U.S.A. 103*, 6489–6494.
17. Park, C., and Marqusee, S. (2004) Probing the high energy states in proteins by proteolysis. *J. Mol. Biol. 343*, 1467–1476.
18. Hnizda, A., Santrucek, J., Sanda, M., Strohalm, M., and Kodicek, M. (2008) Reactivity of histidine and lysine side-chains with diethylpyrocarbonate—a method to identify surface exposed residues in proteins. *J. Biochem. Biophys. Methods 70*, 1091–1097.
19. Hirasawa, M., Kleis-SanFrancisco, S., Proske, P. A., and Knaff, D. B. (1995) The effect of N-bromosuccinimide on ferredoxin:$NADP^+$ oxidoreductase. *Arch. Biochem. Biophys. 320*, 280–288.
20. Hubalek, F., Pohl, J., and Edmondson, D. E. (2003) Structural comparison of human monoamine oxidases A and B: mass spectrometry monitoring of cysteine reactivities. *J. Biol. Chem. 278*, 28612–28618.
21. Wells, I., and Marnett, L. J. (1992) Acetylation of prostaglandin endoperoxide synthase by *N*-acetylimidazole: comparison to acetylation by aspirin. *Biochemistry 31*, 9520–9525.
22. Gabant, G., Augier, J., and Armengaud, J. (2008) Assessment of solvent residues accessibility using three sulfo-NHS-biotin reagents in parallel: application to footprint changes of a methyltransferase upon binding its substrate. *J. Mass Spectrom. 43*, 360–370.
23. Carven, G. J., and Stern, L. J. (2005) Probing the ligand-induced conformational change in HLA-DR1 by selective chemical modification and mass spectrometric mapping. *Biochemistry 44*, 13625–13637.
24. Kozich, V., and Kraus, J. P. (1992) Screening for mutations by expressing patient cDNA segments in *E. coli*: homocystinuria due to cystathionine beta-synthase deficiency. *Hum. Mutat. 1*, 113–123.
25. Biringer, R. G., Amato, H., Harrington, M. G., Fonteh, A. N., Riggins, J. N., and Huhmer, A. F. (2006) Enhanced sequence coverage of proteins in human cerebrospinal fluid using multiple enzymatic digestion and linear ion trap LC-MS/MS. *Briefings Funct. Genomics Proteomics 5*, 144–153.
26. Strohalm, M., Hassman, M., Kosata, B., and Kodicek, M. (2008) mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun. Mass Spectrom. 22*, 905–908.
27. Turner, B. T., Jr., Sabo, T. M., Wilding, D., and Maurer, M. C. (2004) Mapping of factor XIII solvent accessibility as a function of activation state using chemical modification methods. *Biochemistry 43*, 9755–9765.
28. Perdivara, I., Deterding, L. J., Przybylski, M., and Tomer, K. B. (2010) Mass spectrometric identification of oxidative modifications of tryptophan residues in proteins: chemical artifact or post-translational modification? *J. Am. Soc. Mass Spectrom. 21*, 1114–1117.
29. Sharp, J. S., Becker, J. M., and Hettich, R. L. (2004) Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectrometry. *Anal. Chem. 76*, 672–683.
30. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779–815.
31. Lucas, M., Encinar, J. A., Arribas, E. A., Oyenarte, I., Garcia, I. G., Kortazar, D., Fernandez, J. A., Mato, J. M., Martinez-Chantar, M. L., and Martinez-Cruz, L. A. (2010) Binding of S-methyl-5′-thioadenosine and S-adenosyl-L-methionine to protein MJ0100 triggers an open-to-closed conformational change in its CBS motif pair. *J. Mol. Biol. 396*, 800–820.
32. Kelley, L. A., and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc. 4*, 363–371.
33. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389–3402.

34. Sippl, M. J. (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins 17*, 355–362.
35. Halabi, N., Rivoire, O., Leibler, S., and Ranganathan, R. (2009) Protein sectors: evolutionary units of three-dimensional structure. *Cell 138*, 774–786.
36. Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res. 38*, D211–D222.
37. Chen, R., Li, L., and Weng, Z. (2003) ZDOCK: an initial-stage protein-docking algorithm. *Proteins 52*, 80–87.
38. Evande, R., Blom, H., Boers, G. H., and Banerjee, R. (2002) Alleviation of intrasteric inhibition by the pathogenic activation domain mutation, D444N, in human cystathionine beta-synthase. *Biochemistry 41*, 11832–11837.
39. Mendoza, V. L., and Vachet, R. W. (2008) Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem. 80*, 2895–2904.
40. Kvaratskhelia, M., Miller, J. T., Budihas, S. R., Pannell, L. K., and Le Grice, S. F. (2002) Identification of specific HIV-1 reverse transcriptase contacts to the viral RNA:tRNA complex by mass spectrometry and a primary amine selective reagent. *Proc. Natl. Acad. Sci. U.S.A. 99*, 15988–15993.
41. Blair, L. P., Tackett, A. J., and Raney, K. D. (2009) Development and evaluation of a structural model for SF1B helicase Dda. *Biochemistry 48*, 2321–2329.
42. Leitner, A., and Lindner, W. (2005) Effect of an arginine-selective tagging procedure on the fragmentation behavior of peptides studied by electrospray ionization tandem mass spectrometry (ESI-MS/MS). *Anal. Chim. Acta 528*, 165–173.
43. Ignoul, S., and Eggermont, J. (2005) CBS domains: structure, function, and pathology in human proteins. *Am. J. Physiol. 289*, C1369–C1378.
44. Meier, M., Oliveriusova, J., Kraus, J. P., and Burkhard, P. (2003) Structural insights into mutations of cystathionine beta-synthase. *Biochim. Biophys. Acta 1647*, 206–213.
45. Urreizti, R., Asteggiano, C., Cozar, M., Frank, N., Vilaseca, M. A., Grinberg, D., and Balcells, S. (2006) Functional assays testing pathogenicity of 14 cystathionine-beta synthase mutations. *Hum. Mutat. 27*, 211.
46. Majtan, T., Liu, L., Carpenter, J. F., and Kraus, J. P. (2010) Rescue of cystathionine beta-synthase (CBS) mutants with chemical chaperones: purification and characterization of eight CBS mutant enzymes. *J. Biol. Chem. 21*, 15866–15873.
47. Kozich, V., Sokolova, J., Klatovska, V., Krijt, J., Janosik, M., Jelinek, K., and Kraus, J. P. (2010) Cystathionine beta-synthase mutations: effect of mutation topology on folding and activity. *Hum. Mutat. 7*, 809–819.
48. Sen, S., and Banerjee, R. (2007) A pathogenic linked mutation in the catalytic core of human cystathionine beta-synthase disrupts allosteric regulation and allows kinetic characterization of a full-length dimer. *Biochemistry 46*, 4110–4116.
49. Tsai, C. J., del Sol, A., and Nussinov, R. (2008) Allostery: absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol. 378*, 1–11.
50. Kern, D., and Zuiderweg, E. R. (2003) The role of dynamics in allosteric regulation. *Curr. Opin. Struct. Biol. 13*, 748–757.
51. Laskowski, R. A., Gerick, F., and Thornton, J. M. (2009) The structural basis of allosteric regulation in proteins. *FEBS Lett. 583*, 1692–1698.
52. Goodey, N. M., and Benkovic, S. J. (2008) Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol. 4*, 474–482.
53. Singh, S., Madzelan, P., and Banerjee, R. (2007) Properties of an unusual heme cofactor in PLP-dependent cystathionine beta-synthase. *Nat. Prod. Rep. 24*, 631–639.
54. Yamanishi, M., Kabil, O., Sen, S., and Banerjee, R. (2006) Structural insights into pathogenic mutations in heme-dependent cystathionine-beta-synthase. *J. Inorg. Biochem. 100*, 1988–1995.

SUPPLEMENTARY INFORMATION

A. Hnízda et al. (2010): Cross-Talk between the Catalytic Core and the Regulatory Domain in Cystathionine β-Synthase: Study by Differential Covalent Labeling and Computational Modeling

Table S1: Conditions of modification reactions used in this study. The most reactive amino acid is written in the first position, other potentially reactive amino acids are in the brackets.

Modifier	Modification buffer	Reaction time [min]	Reactive amino acid residues[a]	Mass shift in modified peptides	References
DEP	50 mM potassium phosphate; pH 6.8	60	H (K, C, N-terminal amino acid)	+72; +15	(*18*)
NBS	50 mM potassium phosphate; pH 6.5	5	W (M)	+16; +32	(*19*)
NEM	20 mM Tris-HCl, pH 8.0	30	C	+125	(*5, 20*)
NAI	10 mM potassium phosphate, pH 7.5	60	Y (K, N-terminal amino acid)	+42	(*21*)
NHS	100 mM sodium bicarbonate, pH 8.5	30	K (N-terminal amino acid)	+42	(*22*)
HPG	50 mM sodium bicarbonate, pH 8.0	60	R	+132	(*23*)

Table S2: Enzymatic activities of modified CBS for molar ratio of protein:modifier used in the study. Relative activities are refered to enzymatic activities of unmodifed control in absence of AdoMet; values are expressed as a mean of 3 measurements with standard deviation. Activation of CBS activity by AdoMet in the unmodified control differed for each labelling compound due to variant sample preparation.

			Relative activity of modified protein		Activation of CBS activity by AdoMet	
Modifier	Protein	CBS:modifier molar ratio	-AdoMet	+AdoMet	unmodified control	modified protein
DEP	45CBS	1-5	0.77 ± 0.13	N/A	N/A	N/A
	wtCBS		0.90 ± 0.16	2.45 ± 0.47	3.3	2.7
NBS	45CBS	1-1	0.43 ± 0.05	N/A	N/A	N/A
	wtCBS		1.04 ± 0.20	1.63 ± 0.35	2.3	1.6
NEM	45CBS	1-1	0.98 ± 0.18	N/A	N/A	N/A
	wtCBS		1.36 ± 0.23	3.18 ± 0.33	2.9	2.3
NAI	45CBS	1-250	1.09 ± 0.15	N/A	N/A	N/A
	wtCBS		1.18 ± 0.16	1.67 ± 0.36	3.0	1.3
NHS	45CBS	1-10	0.77 ± 0.17	N/A	N/A	N/A
	wtCBS		0.73 ± 0.13	2.08 ± 0.38	2.7	2.7
HPG	45CBS	1-250	0.69 ± 0.08	N/A	N/A	N/A
	wtCBS	1-500	0.68 ± 0.10	1.26 ± 0.10	3.1	1.9

Table S3: Relative enzymatic activities of modified thermally activated wtCBS. Relative activities are given as ratios of modified enzyme to an unmodifed control; values are expressed as a mean with standard deviation. Heat stimulation shows increase in enzymatic activity upon partial thermal denaturation.

Modifier	Rel. Activity	Heat stimulation
NAI	0.46 ± 0.12	2.1
HPG	0.60 ± 0.07	1.5
NBS	0.56 ± 0.09	1.6
NHS	0.77 ± 0.19	1.9

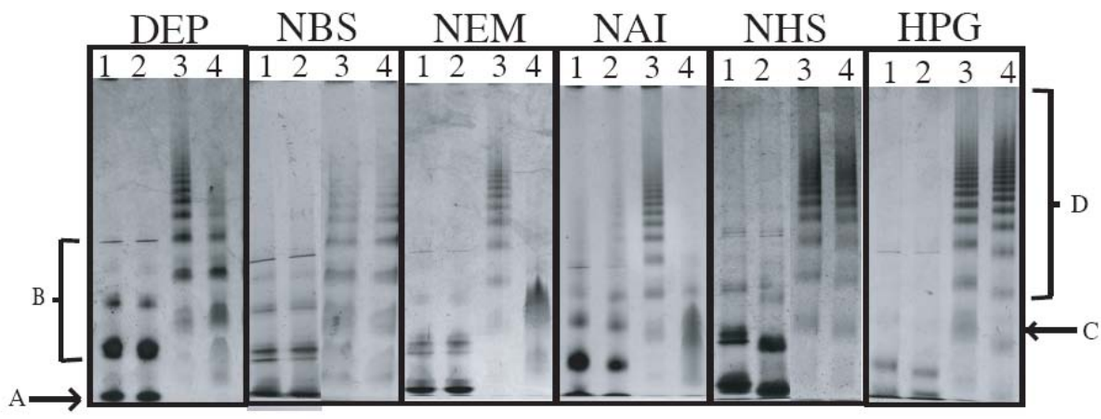


Figure S1: Analysis of labelled CBS by native electrophoresis. Arrows indicate bands corresponding to common oligomeric forms of the unmodified proteins (A – dimeric 45CBS; B – higher order oligomers of 45CBS; C – tetramer of wtCBS; D – higher order oligomers of wtCBS). Molecular weight markers were not used since migration distance of CBS oligomers is not determined by their masses only; the expected migration of oligomers is shown by arrows in accordance with previous reports on behaviour of CBS in native electrophoretic gels (*5*, *47*).

Each labelling experiment is placed in separated box with identical numbering: 1 – unmodified 45CBS; 2 – labelled 45CBS; 3 – unmodified wtCBS; 4 – labelled wtCBS. Proteins shown in this figure were labeled by modifier at the same concentrations used for surface mapping (ratio of protein:modifier are shown in Table S2).

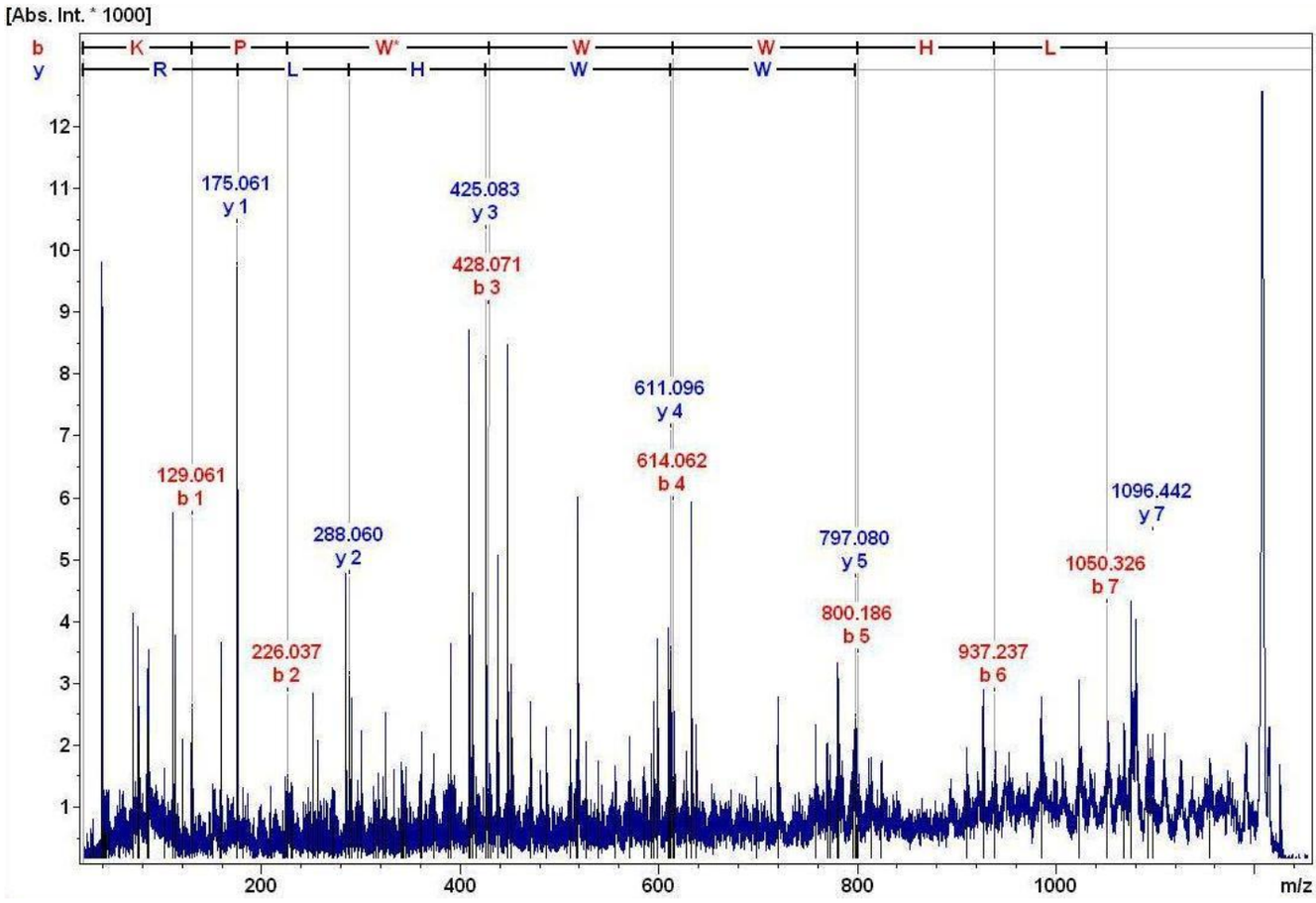


Figure S2: MS/MS spectrum of differentially reactive peptide 406-413 (KPWWWHLR) that contained modified tryptophane cluster W408/409/410. The labelling was perfomed with NBS.

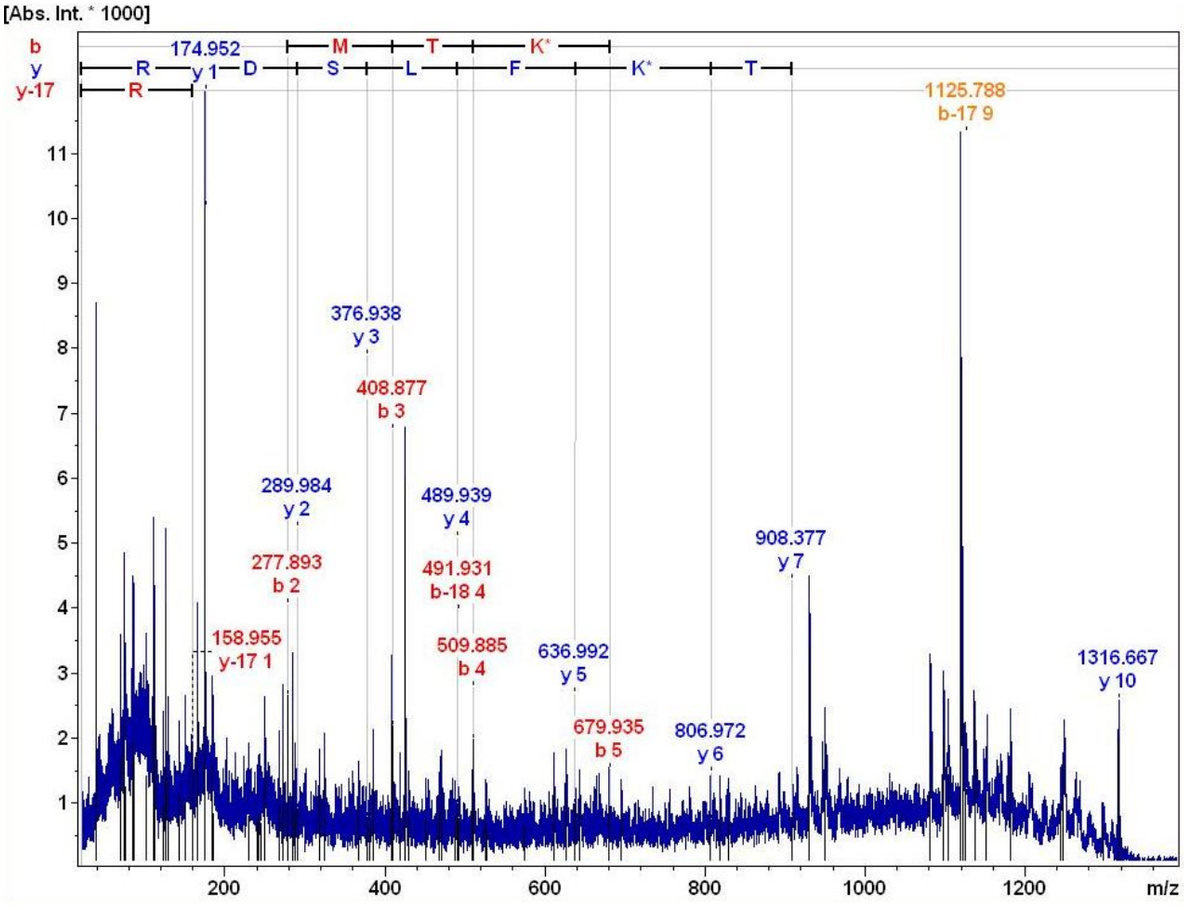


Figure S3: MS/MS spectrum of differentially reactive peptide 380-389 (NYMTKFLSDR) that contained modified K384. The labelling was performed with NAI.

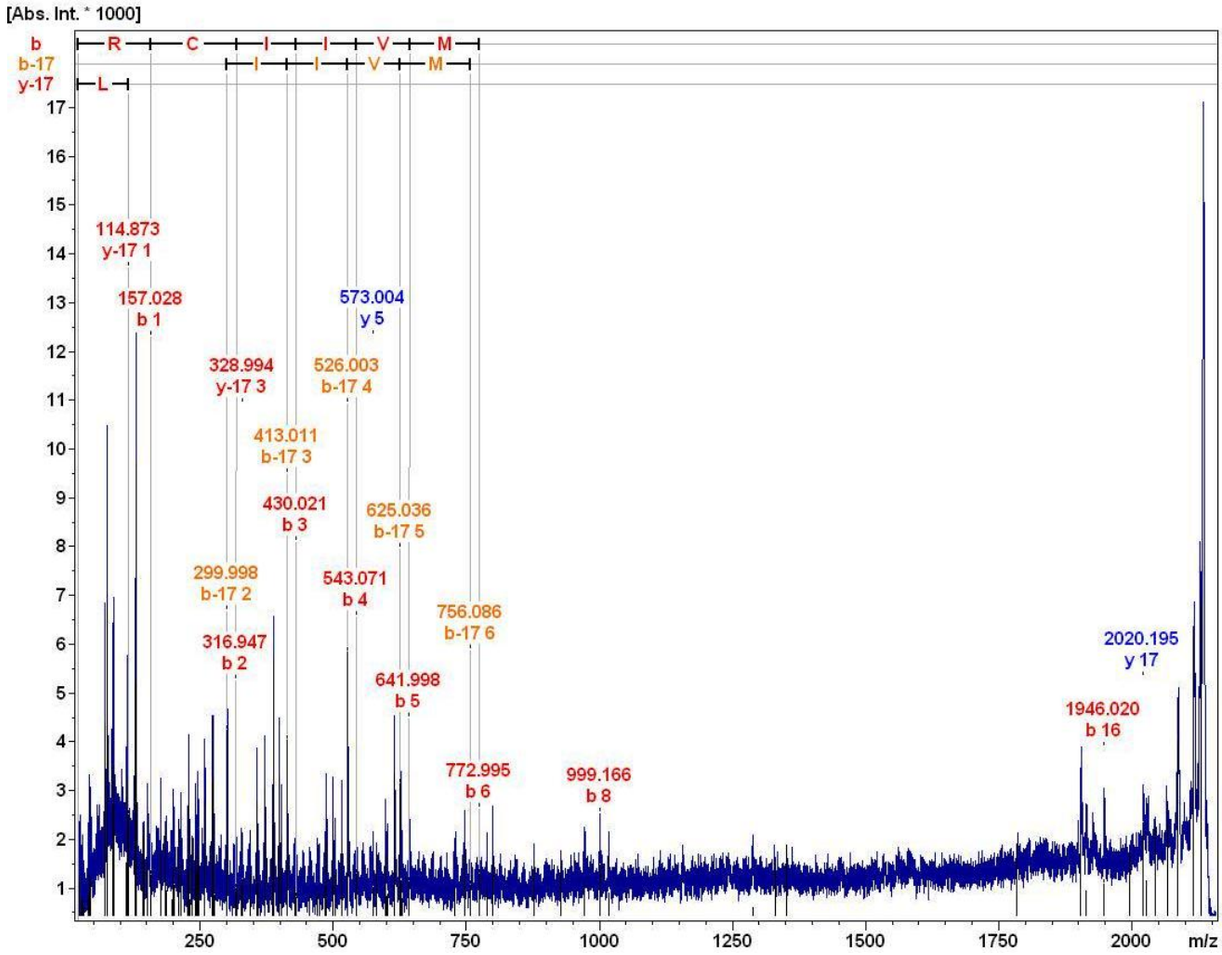


Figure S4: MS/MS spectrum of differentially reactive peptide 164-181 (RCIIVMPEKMSSEKVDVL) that contained modified K172 and/or K177. The labelling was performed with NHS.

Evaluation of the model of the autoinhibitory domain

using statistical coupling analysis and protein sector analysis

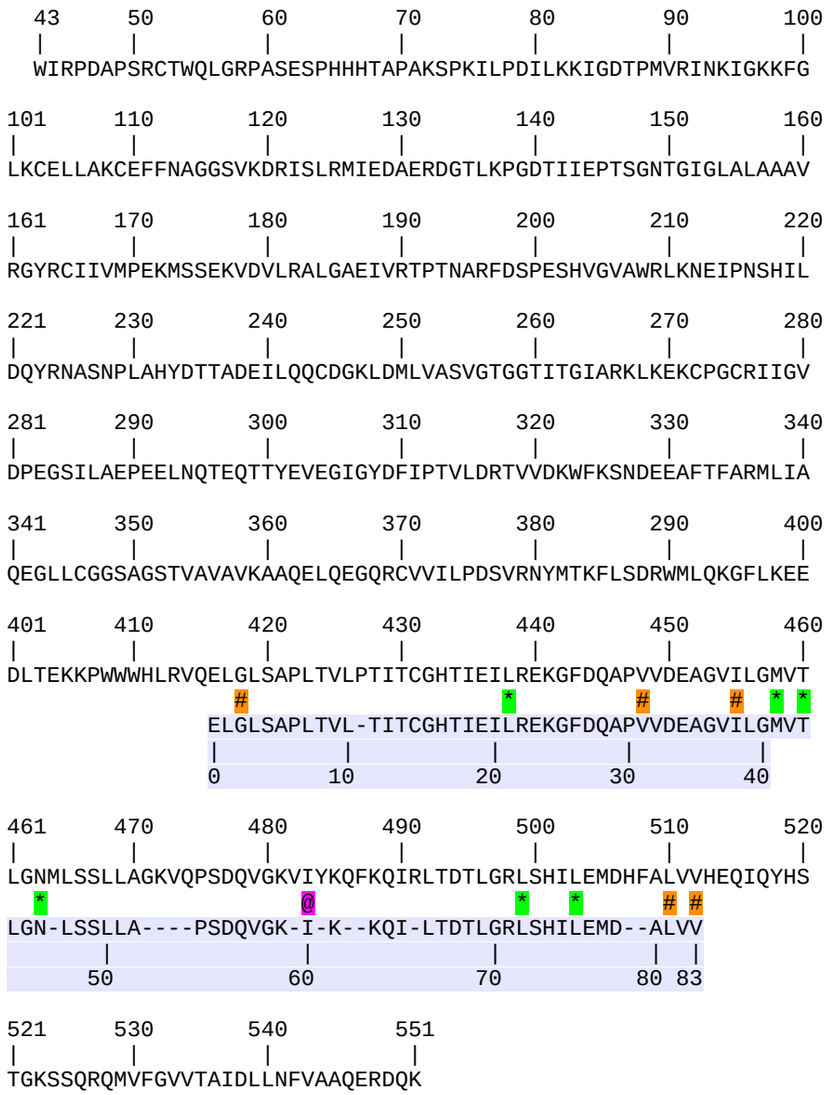
Statistical coupling analysis (SCA) and protein sector identification was performed using our in-house Python script on the basis of procedures introduced by Halabi an co-workers [Halabi 2009]. Data were visualized using Matplotlib library (matplotlib.sourceforge.net). The multiple sequence alignment of CBS domain sequences was obtained from Pfam database [Finn, 2010]. Incomplete sequences were manually removed, resulting to 6,983 sequences. Positions with a high number of gaps (> 20 %) were excluded from the analysis (Scheme SI). The multiple sequence alignment (6,983 x 84) was then converted to a binary form (1 = the prevalent amino acid in the position, 0 = otherwise). Figure S5 shows that application of a binary alignment does not cause a significant loss of information. Figure S6 shows relative entropies (conservation of individual alignment positions, top) and statistical coupling matrix (evolutionary correlation between positions, bottom). High coupling values indicate that a presence of amino acid residue in one position influences the presence of another amino acid residue in another position. Eigenvalues of this matrix are shown as a histogram in Figure S7. The first eigenvector correlates with a net contribution of each position (Figure S8) and it is therefore not used in further analysis. The second, third and fourth eigenvectors show an evolutionary coupling are is illustrated in Figure S9. The weight of the second eigenvector of the position corresponding to Ile483 in human CBS is high as this position is strongly correlated with others. It was therefore identified as a separate sector (magenta). Other two sectors were identified as positions with weights of the third eigenvector lower than –0.12 (green) or higher than 0.17 (orange). Figure 3A illustrates the model of CBS domain with sector residues shown in surface representation. The matrix in Figure 3B was obtained by filtering sector position from the matrix in Figure S6. Detailed explanation of statistical

coupling analysis and protein sectors can be obtained in the article of Halabi and co-workers [Halabi, 2009].

References:

[Halabi, 2009] Halabi, N., Rivoire, O., Leibler, S., Ranganathan R. (2009) Protein Sectors: Evolutionary Units of Three-Dimensional Structure. *Cell* **138**, 774-786.

[Finn, 2010] Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunesekaran, P., G. Ceric, G., K. Forslund, K., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman A. (2010) The Pfam Protein Families Database. *Nucleic Acids Res.* **38**, D211-D222.



Scheme SI: Amino acid sequence of human CBS. The autoregulatory domain residues analysed by statistical coupling analysis are shown as an alignment with the full sequence (in grey, missing positions were removed prior to analysis because they contained a large number of gaps in the multiple sequence alignment of the family). Identified sectors are highlighted (in green, orange and the residue Ile483 in magenta).

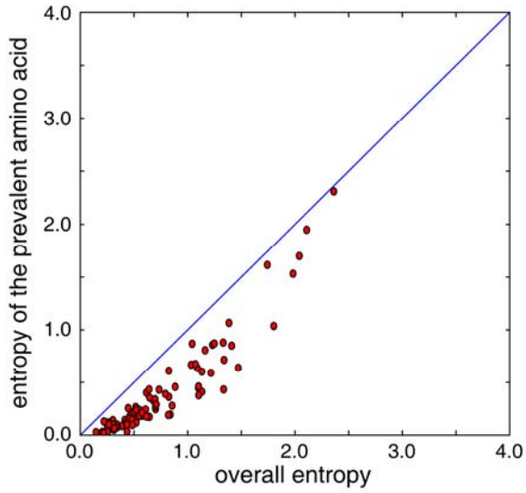


Figure S5: Agreement between the entropy of the prevalent amino acid and the overall relative entropy.

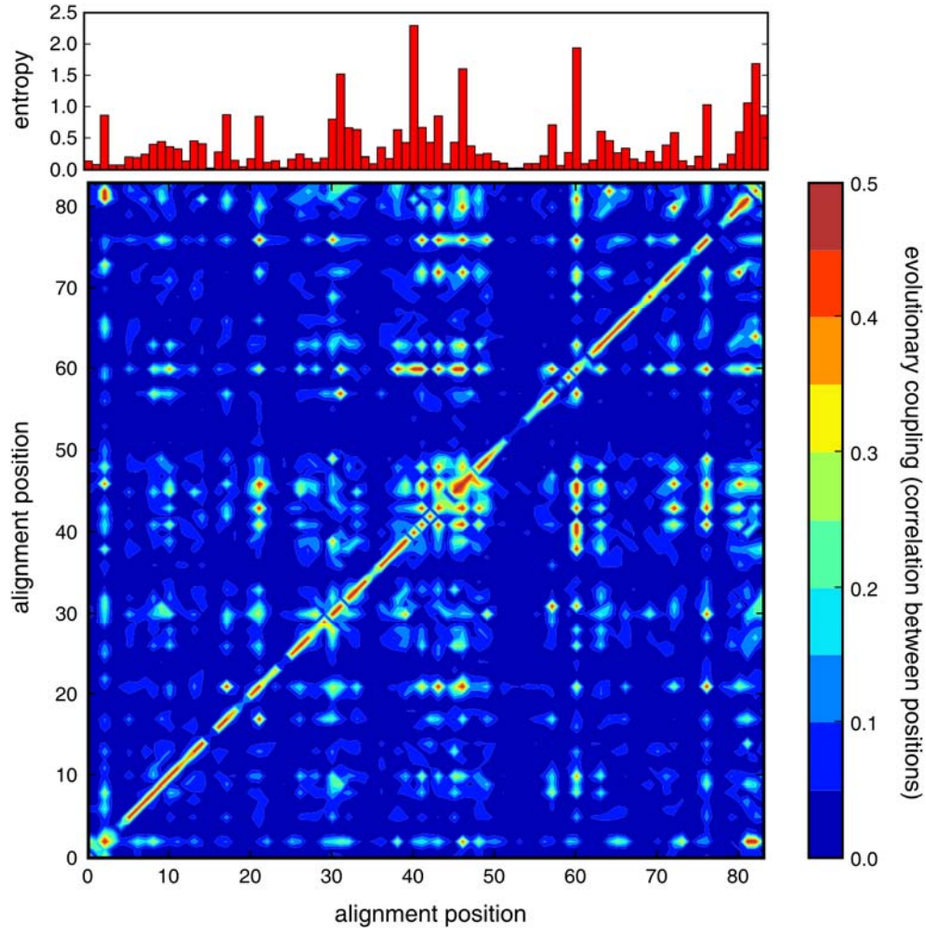


Figure S6: Conservation of individual positions in the multiple sequence alignment of CBS domains, described as relative entropy (top) and SCA matrix (bottom). Alignment positions correspond to Scheme SI.

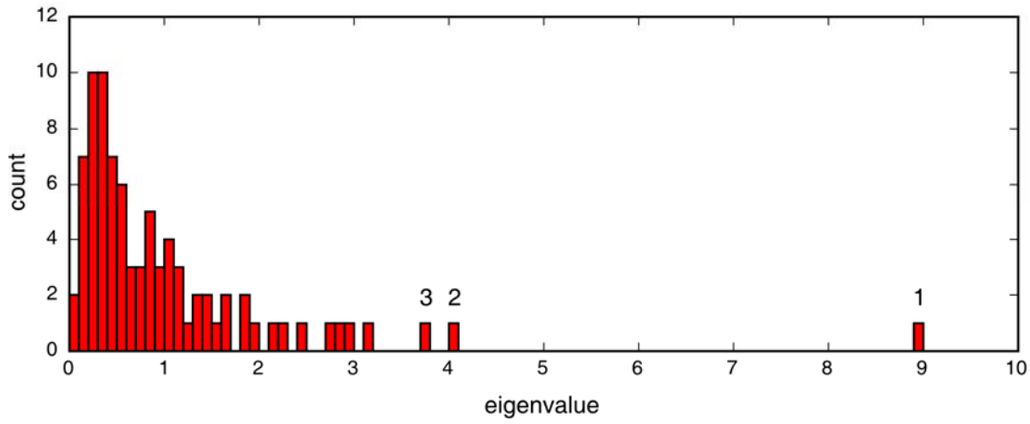


Figure S7: Eigenvalue histogram for the SCA matrix. The first, second and third eigenvalues are indicated.

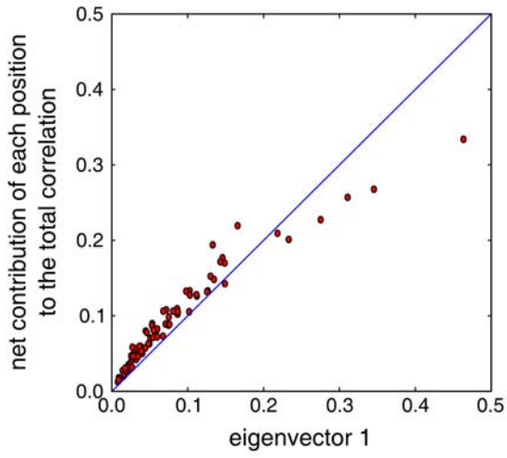


Figure S8: The first eigenvector of SCA matrix against the net contribution of each position to the total correlation.

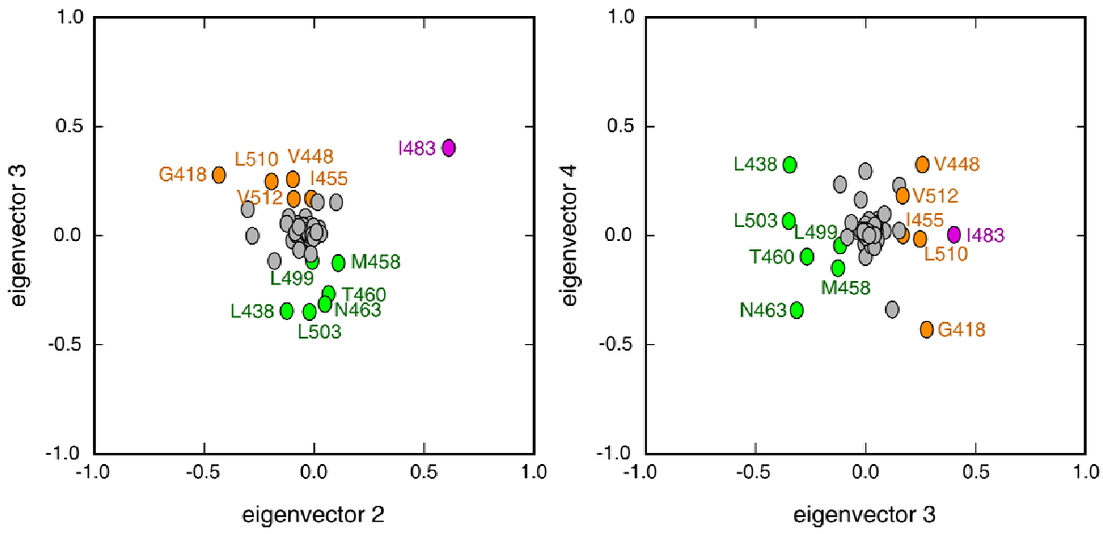


Figure S9: Residue weights along eigenvectors 2, 3 and 4. Sector residues are indicated.

Publication 5.1.3

R. Vozdek et al.

Novel structural arrangement of nematode cystathionine beta-synthases: characterization of Caenorhabditis elegans CBS-1

Biochemical Journal, 2012

BJ www.biochemj.org

Biochem. J. (2012) **443**, 535–547 (Printed in Great Britain) doi:10.1042/BJ20111478

Novel structural arrangement of nematode cystathionine β-synthases: characterization of *Caenorhabditis elegans* CBS-1

Roman VOZDEK, Aleš HNÍZDA, Jakub KRIJT, Marta KOSTROUCHOVÁ and Viktor KOŽICH[1]

Institute of Inherited Metabolic Disorders, Charles University in Prague, First Faculty of Medicine and General University Hospital, Ke Karlovu 2, 128 08, Praha 2, Czech Republic

CBSs (cystathionine β-synthases) are eukaryotic PLP (pyridoxal 5 *-phosphate)-dependent proteins that maintain cellular homocysteine homoeostasis and produce cystathionine and hydrogen sulfide. In the present study, we describe a novel structural arrangement of the CBS enzyme encoded by the *cbs-1* gene of the nematode *Caenorhabditis elegans*. The CBS-1 protein contains a unique tandem repeat of two evolutionarily conserved catalytic regions in a single polypeptide chain. These repeats include a catalytically active C-terminal module containing a PLP-binding site and a less conserved N-terminal module that is unable to bind the PLP cofactor and cannot catalyse CBS reactions, as demonstrated by analysis of truncated variants and active-site mutant proteins. In contrast with other metazoan enzymes, CBS-1 lacks the haem and regulatory Bateman domain essential for activation by AdoMet (*S*-adenosylmethionine) and only forms monomers. We determined the tissue and subcellular distribution of CBS-1 and showed that *cbs-1* knockdown by RNA interference leads to delayed development and to an approximately 10-fold elevation of homocysteine concentrations in nematode extracts. The present study provides the first insight into the metabolism of sulfur amino acids and hydrogen sulfide in *C. elegans* and shows that nematode CBSs possess a structural feature that is unique among CBS proteins.

Key words: cystathionine β-synthase (CBS), *Caenorhabditis elegans*, domain architecture, homocysteine, hydrogen sulfide, knockdown.

INTRODUCTION

Methionine and cysteine are sulfur amino acids that play important roles in many biochemical reactions. Methionine, an essential amino acid, can be irreversibly converted into cysteine in a series of reactions. Methionine is first converted into AdoMet (*S*-adenosylmethionine), which serves as a methyl donor in various transmethylation reactions. A product of these transmethylations, *S*-adenosylhomocysteine, is further converted into homocysteine, which is a key intermediate in the metabolism of sulfur amino acids. In animal tissues, homocysteine is universally remethylated to methionine by methionine synthase using methyltetrahydrofolate as the methyl donor. In addition, a number of tissues can convert homocysteine into cystathionine and further to cysteine via the transsulfuration pathway through two PLP (pyridoxal 5*-phosphate)-dependent enzymes, CBS (cystathionine β-synthase) and CGL (cystathionine γ-lyase) [1].

CBS is a cytosolic enzyme that catalyses the formation of cystathionine with the release of water or hydrogen sulfide, depending on whether homocysteine is condensed with serine or cysteine. The human and rodent CBSs that have been characterized are tetrameric enzymes and each of their ~63 kDa polypeptide chains contains three different domains. The N-terminal domain binds haem, the presence of which has been suggested to increase CBS activity during the oxidation of the intracellular environment [1a,2]. Other studies suggest that haem may play a structural role that is necessary for the correct folding of the CBS protein [3,4]. The middle portion of the polypeptide chain forms the catalytic domain and is well conserved among the fold-type II PLP-dependent proteins [5]. The C-terminal domain possesses two defined CBS domains, a hydrophobic domain, CBS1, and a less conserved domain, CBS2, that are together referred to as the Bateman domain. Together, these two CBS domains bind AdoMet, an allosteric activator of mammalian CBS [6]. The C-terminal domain of mammalian CBS is also thought to be responsible for multimerization of the enzyme into homotetramers and higher oligomeric forms [6,7]. The C-terminal autoinhibitory domain of mammalian CBS can be removed by *in vitro* or *in vivo* proteolytic processing, yielding a ~45 kDa truncated form (45CBS) that forms dimers and is more active than the full-length enzyme [6,8,9].

The canonical domain architecture of mammalian CBSs is not conserved across phyla. The CBS enzymes of *Saccharomyces cerevisiae, Trypanosoma cruzi* and *Drosophila melanogaster* have been experimentally characterized; the N-terminal haem-binding domain is absent in yeast and protozoan CBS, in contrast with its presence in *Drosophila* [10–12], whereas the catalytic domain is conserved in the CBS enzymes of all three of these species. The C-terminal portion exhibits the highest degree of variability. The yeast and *Drosophila* CBS proteins contain the Bateman domain, but lack a response to AdoMet. Interestingly, although the C-terminal portion of the yeast CBS inhibits the activity of the enzyme and supports the formation of tetramers and octamers [13], *Drosophila* CBS forms only dimers [12]. In contrast, the protozoan CBS does not contain the Bateman domain and is not activated by AdoMet. Although its C-terminus is shortened, the protozoan CBS is still able to form tetramers [11]. The phylogenetic variability in the domain architecture of CBSs suggests that the activity of these enzymes is regulated differently in evolutionarily distant organisms.

In the present study, we characterized the structural and functional properties of the CBS in *Caenorhabditis elegans*, a

Abbreviations used: AdoMet, *S*-adenosylmethionine; BN, blue native; BS^3, bis(sulfosuccinimidyl) suberate; CBS, cystathionine β-synthase; CGL, cystathionine γ-lyase; DTT, dithiothreitol; EST, expressed sequence tag; GFP, green fluorescent protein; LC–MS/MS, liquid chromatography–tandem MS; PLP, pyridoxal 5*-phosphate; RNAi, RNA interference; RT, reverse transcription; SEC, size-exclusion chromatography; UTR, untranslated region; WT, wild-type.

[1] To whom correspondence should be addressed (email Viktor.Kozich@LF1.cuni.cz).

well-established model organism used to study human diseases. We first identified a transcriptionally active gene encoding CBS in *C. elegans*, we then determined its pattern of expression and characterized the enzymatic and structural properties of the encoded protein. Finally, we determined the phenotypic effects of *cbs-1* inhibition using RNA-mediated interference. These data describe novel structural features that are unique among CBS enzymes and provide the first insight into the metabolism of sulfur amino acids and hydrogen sulfide in *C. elegans*.

EXPERIMENTAL

C. elegans strains

The WT (wild-type) *C. elegans* Bristol strain N2 was obtained from the *C. elegans* Stock Center (University of Minnesota, Minneapolis, MN, U.S.A.), and the RB839 strain carrying the *F54A3.4* (*ok666*) allele was provided by the *C. elegans* Gene Knockout Consortium (Oklahoma Medical Research Foundation, Oklahoma City, OK, U.S.A.). Worm cultures were maintained as described previously [14].

Bioinformatics

BLASTp searches were performed by online BLAST software using the *C. elegans* protein database (release WS215). Protein domain modelling was performed by Swiss-model (automatic modelling mode) using the crystal structure of human 45CBS (PDB code 1JBQ, chain A) as a template [15]. PDB structures were subsequently evaluated in the Prosa program [16] and visualized in Swiss-PDBViewer 4.0.4 [17]. Phylogenetic trees were constructed in the online portal system Mobyle [18]. Multiple alignments of amino acid sequences were performed using ClustalW2 online software with default parameters [19]. Conserved regions were also separated for further analysis by ClustalW2. For phylogenetic analysis, alignment was bootstrapped 100 times and analysed by the maximal likelihood method using the PHYML 3.0 program [20]. Bootstrap output trees were analysed by the PHYLIP 3.67 CONSENSE program; the final tree shape was visualized in the Dendroscope program [21].

PCR amplification and DNA sequencing

Nematode cDNA was prepared by RT (reverse transcription) using isolated total RNA from mixed stages of N2 worms and a RT kit with an oligo(dT) primer (Promega). Open reading frames of *ZC373.1* and *F54A3.4* were amplified by PCR using either cDNA prepared by RT–PCR or a *C. elegans* cDNA library (Invitrogen) as the template (a list of the primers is given in Supplementary Table S1 at http://www.BiochemJ.org/bj/443/bj4430535add.htm). PCR products were cloned into the pCR4-TOPO vector (Invitrogen), and the authenticity of the DNA sequence was verified by dideoxy sequencing using an ABI PRISM 3100-Avant sequencer (Applied Biosystems).

GFP (green fluorescent protein) reporter assay

To determine the expression pattern of *cbs-1*, we generated a translational fusion vector using the PCR fusion technique described previously [22]. The 1.8 kb of 5′ upstream sequence and the entire coding region of *ZC373.1* were amplified by PCR using primers A and B (Supplementary Table S1), and genomic *C. elegans* DNA as a template. The vector pPD95.75 was used as a template for amplification of the GFP-coding sequence using primers C and D (Supplementary Table S1). The two PCR products were mixed and used as a template for PCR fusion using nested primers E and F (Supplementary Table S1). The 6.8-kb PCR product was injected into *C. elegans* hermaphrodite gonads together with the plasmid pRF4 as a phenotypic marker for injection. Transgenic animals were separated, and the F2 progeny were screened for the GFP signal. An Olympus BX60 microscope and a Nikon Eclipse E800 with C1 confocal module and 488 nm laser and differential interference contrast optics were used for specimen examination.

Bacterial expression and protein purification

Initially, recombinant CBS-1 was expressed as a fusion protein with an N-terminal GST tag and further purified by affinity chromatography to 75 % purity (see Supplementary Figure S4, lane 6) according to a previously described procedure for human CBS [23]. The contaminating polypeptide with the highest abundance, a 40-kDa fragment that represented approximately 20 % of the total protein, was identified as the N-terminal portion of CBS-1 (residues 1–375) by peptide mass fingerprinting using MS detection (results not shown). This N-terminal fragment was observed with similar abundance even when the purification procedure was modified to limit proteolytic cleavage of the recombinant protein (the modification involved performing affinity chromatography at 4 °C and increasing the concentration of protease inhibitors in the bacterial crude extract). To overcome this obstacle that was not previously reported for other CBS orthologues, we constructed a new vector that produced double-tagged CBS-1 with a cleavable N-terminal GST tag and a C-terminal His tag. The open reading frame of the *ZC373.1* gene encoding CBS-1 was amplified by PCR using a *C. elegans* cDNA library as the template. PCR was performed with Taq polymerase using primers P and R (Supplementary Table S1). The 2.1-kb DNA fragment obtained by digestion of the PCR product with BamHI and XhoI was cloned into the BamHI- and XhoI-digested pGEX-6p-1 vector. Express Competent *Escherichia coli* cells (New England Biolabs) were transformed with the plasmid that encodes double-tagged CBS-1 (GST–CBS–1–His_6) and cultured in the presence of 100 μM IPTG (isopropyl β-D-thiogalactopyranoside) at 18 °C for 24 h. The GST–CBS-1 fusion protein was purified according to the purification protocol for human CBS described previously [24] with the following modifications: after cleavage by the PreScission protease (GE Healthcare), recombinant CBS-1 was loaded on to a Ni-Sepharose column that had been equilibrated with IMAC buffer [20 mM phosphate (pH 7.5), containing 0.5 M NaCl, 20 mM imidazole and 1 mM DTT (dithiothreitol)]. The column was washed with IMAC buffer containing 50 mM imidazole. CBS-1 was then eluted with IMAC buffer containing 75 mM imidazole. The protein enrichment procedure yielded approximately 1 mg of CBS-1 per litre of bacterial culture. The purity of isolated CBS-1 was analysed by SDS/PAGE [pre-cast 3–8 % gradient gel (Invitrogen)] with Coomassie Brilliant Blue staining. The protein concentration was determined using Bradford reagent (Sigma–Aldrich) with BSA as the standard. The absorption spectrum of CBS-1 was recorded using a UV–visible spectrophotometer (Shimadzu UV-2550) at room temperature (25 °C).

SEC (size-exclusion chromatography)

SEC was performed on an HPLC platform (Shimadzu LC-10A system). Recombinant purified CBS-1 was loaded on to a Bio-Sil SEC HPLC column (catalogue number 125-0060, Bio-Rad

Laboratories) that had been previously equilibrated with buffer containing 50 mM Tris/HCl (pH 8.0), 1 mM DTT and 100 mM NaCl. The analysis was performed at a flow rate of 1.0 ml/min at 25 °C; the elution profile was obtained by measurement of the absorbance at 280 nm. Calibration was performed using ferritin, aldolase, conalbumin (GE Healthcare), BSA (Thermo Fisher Scientific) and human 45CBS produced in *E. coli* and purified as described previously [24].

Native PAGE, BN (blue native)-PAGE and chemical cross-linking

Native electrophoresis was performed on 8 % polyacrylamide gels using the Laemmli buffer system without SDS [25]. Per lane, 5 μg of CBS-1 and of the standards (BSA and human 45CBS) were loaded. BN electrophoresis was performed as described previously [26] with the High Molecular Weight Calibration kit for electrophoresis (GE Healthcare) and rabbit aldolase as the protein marker. Chemical cross-linking was performed using three different concentrations of BS^3 [bis(sulfosuccinimidyl) suberate]; the molar ratios of CBS-1 (0.5 mg/ml) to the cross-linker were 1:10, 1:50 and 1:100. Cross-linked proteins were analysed using precast 3–8 % gradient polyacrylamide gels. As a positive control for efficient cross-linking, we used dimeric human 45CBS reacted with BS^3 at a protein/cross-linker molar ratio of 1:10. All of the proteins analysed by electrophoretic techniques were stained with EZ Blue Gel reagent (Sigma–Aldrich).

Pulse proteolysis

Pulse proteolysis of CBS-1 in a urea gradient was performed with thermolysin as described previously for human CBS [27].

Fluorescence-based thermal-shift assay

Protein samples (0.5 mg/ml) were dissolved in 20 mM Tris/HCl (pH 8.0), and 5×Sypro Orange dye (Bio-Rad Laboratories). Using the real-time PCR Detection System CFX96 Touch (Bio-Rad Laboratories), the proteins were incubated in a thermal gradient from 25 °C to 70 °C at increments of 0.5 °C and with 1-min-hold intervals. The degree of protein unfolding was monitored by a FRET (fluorescence resonance energy transfer) channel that captured the spectral properties of Sypro Orange unfolded protein complexes (excitation wavelength ≈ 470 nm and emission wavelength ≈ 570 nm). The data were analysed by CFX Manager software, and the melting temperatures were determined using the first derivative spectra.

CD and fluorescence spectroscopy

The CD spectra of CBS-1 protein variants [0.5 mg/ml in 50 mM phosphate buffer (pH 7.5)] were recorded using a Jasco J-810 chiroptic spectrometer. The intrinsic fluorescence of CBS proteins in 50 mM Tris/HCl (pH 8.0), was measured in the same buffer using a PerkinElmer LS55 fluorescence spectrometer. The excitation wavelength for tryptophan was 298 nm (slit width of 5 nm) with an emission signal scanned from 300 to 700 nm (slit width of 5 nm).

Determination of substrate specificity

All enzyme assays were performed at 25 °C with an incubation time of 10 min to ensure a linear increase in cystathionine or cysteine production. The reaction mixtures (50 μl) contained 1 μg/ml purified recombinant CBS-1, 10 mM tested substrates in the combinations shown in the Results section, 1 mM PLP, 1 mM DTT, 1 mg/ml BSA and 150 mM Tris/HCl (pH 7.0). The reactions were stopped by the addition of 25 μl of 1 M trichloracetic acid, and the reaction products were determined by HPLC [28] or LC–MS/MS (liquid chromatography–tandem MS) analysis [29] with the modifications described below.

Temperature and pH optima and kinetic analysis

We measured cystathionine production using LC–MS/MS analysis [29] with the following modifications: assays were performed in 100 mM Bis/Tris buffer with 2 μg/ml purified recombinant CBS-1 and unlabelled serine as the substrate. The temperature optimum for CBS-1 activity was determined in 5 °C temperature intervals from 5 °C to 80 °C at pH 8.0, and the pH optimum of CBS-1 was determined at 25 °C in 0.5 pH unit intervals using 100 mM Bis/Tris buffer at pH 6–10. Kinetic analyses at different concentrations of serine or homocysteine were performed at 25 °C and pH 8.0, and the data were evaluated by non-linear data fitting using software Origin 8 (OriginLab). All measurements were repeated four times and the results are shown as means ± S.D.

Site-directed mutagenesis and preparation of CBS-1 protein variants

We prepared and analysed a series of mutant CBS-1 enzymes that included two missense variants of full-length CBS-1 (E62K and K421A) and six truncated CBS-1 variants (CBS-1b, Δ1–372, Δ1–322, Δ1–299, b/360 and b/375) (Figure 4A). All CBS-1 variants were cloned into the pGEX vector, which produces GST-tagged proteins. The sequences of primers used for cloning and site-directed mutagenesis are shown in Supplementary Table S1. Proteins were expressed and purified according to the procedure developed for CBS-1. The yields of purified mutant proteins were slightly lower, typically approximately 0.5 mg per litre of bacterial culture. The mutant proteins were analysed by UV–visible spectroscopy, CD and fluorescence spectroscopy and by BN-PAGE as described above for CBS-1. The catalytic activities for the reaction of serine with homocysteine were assessed in 100 mM Tris/HCl (pH 8.5) at 25 °C.

RNA-mediated interference

The *cbs-1*-specific sequence (~ 350 bp in length) was prepared by PCR amplification of a *C. elegans* cDNA library primers G and H (Supplementary Table S1) and cloned into the pCR4-TOPO vector. Single-stranded RNAs were prepared from linearized DNA by *in vitro* transcription using T3 DNA-dependent RNA polymerase (construct DNA digested by NotI) and T7 DNA-dependent RNA polymerase (construct DNA digested by SalI). The sense and antisense single-stranded RNAs were mixed and incubated at 68 °C for 10 min, followed by incubation at 37 °C for 30 min. The double-stranded RNA was further purified by phenol/chloroform extraction and precipitated by ethanol; the RNA pellet was diluted in water to an approximate concentration of 2 μg of RNA/μl. The double-stranded RNA was injected into the gonads of young adult hermaphrodite worms as described previously [30]. The embryos of microinjected animals were synchronized in 9–12 h intervals. Nematodes were grown at 16 °C on nematode growth medium plates and fed with *E. coli* strain OP50. After RNAi (RNA interference), the nematodes were seeded in 1× PBS buffer on 2 % agarose and screened by their phenotype.

Determination of CBS-1 antigen levels and measurement of enzymatic activity in *C. elegans* extracts

Worms were grown at 16 °C as described above and collected 7 days after embryo microinjection as a mixed population of all larval stages. Worm lysates were prepared by sonication of worm pellets resuspended in 1 vol. of 100 mM PBS containing protease

inhibitor cocktails for prokaryotic (P8465, Sigma–Aldrich) and eukaryotic (P8340, Sigma–Aldrich) cells. Crude extracts were centrifuged for 1 h at 4 °C and 20000 ***g*** and the supernatants were used for the determination of CBS-1 levels and CBS activity. Western blotting was used to examine CBS-1 antigen levels after RNAi. The samples were submitted to SDS/PAGE (pre-cast 3–8 % gradient gel), and protein immunodetection was performed by Western blot analysis using custom-made rabbit polyclonal anti-CBS-1 antibody prepared against purified recombinant CBS-1 (Exbio Praha). Actin, which was detected using a rabbit anti-actin antibody (Abcam), was used for the normalization of protein loading. The signal levels of CBS-1 and actin were determined by chemiluminescence (Pierce) by employing the ChemiGenius station and Gene Tools software for semi-quantification [31]. The enzyme assay was performed according to a previously described protocol [29] with the modification that the reaction mixture was incubated at 16 °C for 30 min.

Measurement of metabolites in *C. elegans* extracts

Worm lysates were prepared by sonication of worm pellets that had been resuspended in 1 vol. of 100 mM PBS without protease inhibitors (Cole-Parmer GE130 Ultrasonic processor, amplitude 20 for 2 min with 1 s on/off pulses). The crude extracts were centrifuged at 20000 ***g*** for 1 h at 4 °C, and the supernatants were used for HPLC aminothiol determination as described previously [28]. The cystathionine concentration was determined by LC–MS/MS using the EZ:faast kit for amino acid analysis (Phenomenex) [29]. The concentrations of all metabolites measured were normalized to the amount of protein present in the sample.

RESULTS

CBS in *C. elegans* is encoded by *ZC373.1*

We used a BLASTp search as an *in silico* approach to identify genes that encode a CBS in *C. elegans*. Using the query sequences of three enzymes of the CBS family (human CBS, trypanosomal CBS and bacterial cysteine synthases), we identified ten genes with predicted amino acid sequences that are homologous with the catalytic domains of the known CBSs; these genes are annotated in WormBase (http://www.wormbase.org/) as 'CBS and related proteins'. Alignment of the predicted amino acid sequences of the *C. elegans* genes *ZC373.1* and *F54A3.4* revealed the highest homology with human CBS (UniProt entry P35520). These predicted proteins exhibited 54 % sequence identity with the human protein, whereas the other eight predicted proteins showed lower homology, with 21–44 % sequence identity (Supplementary Table S2 at http://www.BiochemJ.org/bj/443/bj4430535add.htm). The BLASTp searches using all ten nematode CBS homologues as the query sequences against the UniProt database, together with phylogenetic analysis, indicated that only ZC373.1 and F54A3.4 are homologous with CBS, whereas the remaining eight amino acid sequences are homologous with other proteins within the family of fold-type II PLP-dependent enzymes (Supplementary Figure S1 at http://www.BiochemJ.org/bj/443/bj4430535add.htm).

An annotation in the WormBase database shows that the *ZC373.1* gene is trans-spliced to SL1 and contains ten exons, including 23 bp of the 5′-UTR (untranslated region) and 149 bp of the 3′-UTR followed by a polyadenylation sequence. The *F54A3.4* gene is predicted to contain either eight exons without any 5′- or 3′-UTRs (http://www.wormbase.org/) or only seven exons terminated by a 77 bp 3′-UTR sequence (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/). To determine whether the *ZC373.1* and *F54A3.4* genes are transcribed and spliced into the predicted full-length mRNAs, we analysed their coding regions by RT–PCR and by sequencing of PCR products. We found two differently spliced variants of the *ZC373.1*; one sequence is identical with the WormBase annotation (*cbs-1*), and the other is a new *ZC373.1* splice variant (*cbs-1b*) containing a 5-bp shortening of exon 7 in its 5′-terminus that leads to a frame-shift with a premature stop codon at amino acid residue 377 (Figure 1A and Supplementary Figure S2 at http://www.BiochemJ.org/bj/443/bj4430535add.htm). In contrast, we were unable to amplify either of the two hypothetical full-length *F54A3.4* mRNAs using several PCR conditions, various primers and various cDNA templates.

Because we did not succeed in detecting the *F54A3.4* mRNA by RT–PCR, we used additional approaches to examine the possible role of this gene in *C. elegans*. *In silico* analysis of the GenBank® database revealed three ESTs (expressed sequence tags) of *F54A3.4*: CK587466.1, CB389123.1 and FN902238.1; however, only FN902238.1 has been mapped to the sense strand of the *F54A3.4* region (http://www.ncbi.nlm.nih.gov/nucleotide/). Furthermore, the proteomic database PeptideAtlas did not contain any peptide matches to the hypothetical protein F54A3.4 (http://www.peptideatlas.org/) [32]. Moreover, expression analysis using translational fusion proteins F54A3.4–GFP and ZC373.1–GFP (cbs-1–GFP) (see below) showed that the GFP signals reflecting the expression pattern of the appropriate genes were observed only in worms carrying ZC373.1–GFP, in contrast with the expression patterns observed in several worms carrying F54A3.4–GFP. Finally, *F54A3.4* does not appear to have functional significance in *C. elegans* because the mutant strain RB839, which carries a deletion of *F54A3.4*, showed CBS activity and homocysteine concentrations indistinguishable from those of the WT strain (results not shown), and did not exhibit abnormal behavioural or a developmental phenotype (results not shown).

On the basis of the findings listed above, *F54A3.4* appears to be a pseudogene and was not further examined in the present study. All of the data above strongly indicate that the *C. elegans* genome contains only one expressed orthologue of the human CBS gene, i.e. *ZC373.1*. In accordance with the recommended nomenclature, this gene was named *cbs-1*.

CBS-1 is a cytoplasmic enzyme that is expressed in the hypodermis and intestine, and in muscle cells

To determine the expression pattern and subcellular localization of *cbs-1*, we constructed the translational vector cbs-1–GFP, which contains the promoter and the entire CBS-1 sequence tagged at the C-terminus with GFP (Figure 1A). In worms expressing cbs-1–GFP, the GFP signal was observed in the hypodermis, intestine, body-wall muscle cells and pharyngeal muscles pm3, pm4, pm5, pm6, pm7 and pm8 in all larval stages as well as in adults (Figure 2). Our data using a translational reporter showed a similar expression pattern, as did previous transcriptional screens, and a novel expression of *cbs-1* in pharyngeal muscles. We did not observe a GFP signal in embryos, although previous transcriptional screens and peptide mapping studies have reported expression of *cbs-1* in this developmental stage [32,33]. The observed GFP signal was distributed diffusely within cells and spared the nucleus, suggesting that CBS-1 is localized in the cytoplasm. These data provide the first reported insight into the tissue and subcellular localization of nematode CBS-1 at the protein level and indicate which nematode tissues can metabolize homocysteine to cystathionine and/or cysteine to hydrogen sulfide.

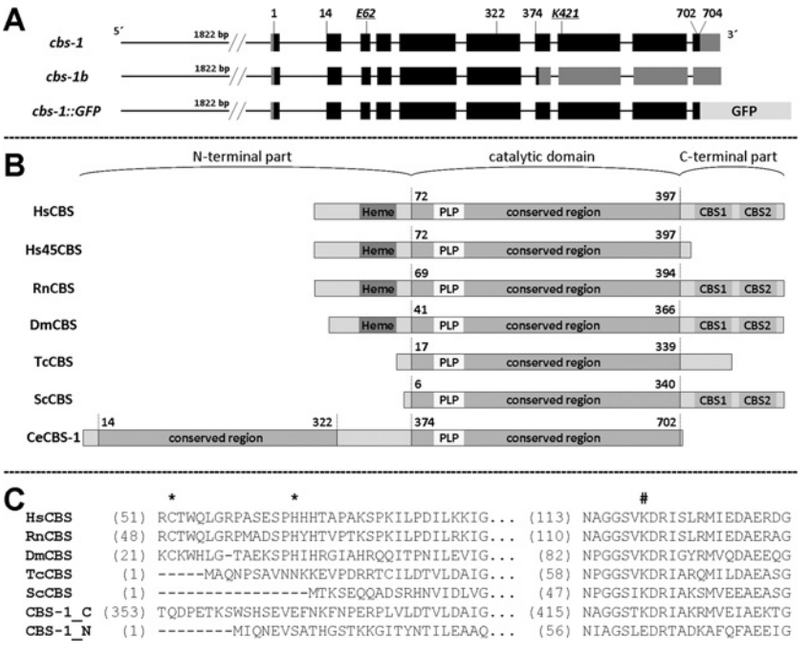


Figure 1 Organization of the *cbs-1* gene and domain architecture of CBS in various organisms

(**A**) Gene organization. The top diagram shows the organization of gene *ZC373.1* (*cbs-1*) encoding CBS-1. The numbers indicate the codon position encoding the appropriate amino acid. Exons are indicated as black boxes, and the 5′- and 3′-UTR sequences are indicated as grey boxes. The middle diagram shows the novel splice *ZC373.1* variant *cbs-1b*. The bottom diagram shows the translational fusion construct used in the GFP reporter assay. The length of the entire promoter used in the cbs-1–GFP construct is indicated by the number of base pairs of the 5′-upstream sequence. (**B**) Domain organization of various CBSs. The published structures of different CBSs were analysed for the presence of a haem-binding site (marked Heme), conserved catalytic regions with a PLP-binding site (marked PLP) and Bateman domain composed of two CBS domains (CBS1 and CBS2). The primary structures are aligned by the PLP-binding lysine residue; the numbers indicate the first and the last amino acid residues of conserved domains in the protein sequence. The aligned proteins are HsCBS (*Homo sapiens* CBS, UniProt entry P35520), Hs45CBS (truncated human CBS with 1–413 residues), RnCBS (*Rattus norvegicus* CBS, UniProt entry P32232), DmCBS (*D. melanogaster* CBS, UniProt entry Q9VRD9), TcCBS (*T. cruzi* CBS, UniProt entry Q9BH24), ScCBS (*S. cerevisiae* CBS, UniProt entry P32582) and CBS-1 (*C. elegans* CBS, UniProt entry Q23264). (**C**) Amino acid alignment of haem- and PLP-binding sites in various CBSs with separated N- and C-terminal conserved regions of CBS-1. The N-terminal region of CBS-1 does not contain the lysine residue that binds PLP. The cysteine and histidine residues that bind haem are indicated by asterisks, and the PLP-binding lysine residues are indicated by #.

CBS-1 is a haem-independent protein that lacks activation by AdoMet

To experimentally characterize the structural and enzymatic properties of CBS-1, *cbs*-1 cDNA was expressed in *E. coli*. Recombinant CBS-1 (704 residues of native CBS-1 with five additional amino acids at the N-terminus, six additional histidine residues at the C-terminus and a size of ~78 kDa) was purified to greater than 95% purity (Supplementary Figure S4 at http://www.BiochemJ.org/bj/443/bj4430535add.htm). The UV–visible absorption spectrophotometry of the purified recombinant CBS-1 showed a peak at 412 nm, confirming the presence of covalently bound PLP that forms an internal aldimine, but it did not reveal a Soret band associated with a haem moiety (Figure 4F). This analysis confirmed that, in contrast with other characterized metazoan CBS enzymes, the nematode enzyme is a haem-independent CBS.

We next tested four reactions that have been described for previously characterized CBSs: (i) cystathionine-synthesizing activity that produces cystathionine and water from homocysteine and serine; (ii) formation of cystathionine and hydrogen sulfide from homocysteine and cysteine; (iii) cysteine synthase activity that produces cysteine from *O*-acetylserine and hydrogen sulfide; and (iv) serine sulfhydrylase activity in which cysteine is synthesized from serine and hydrogen sulfide. CBS-1 exhibited high enzymatic activity for synthesis of cystathionine from homocysteine utilizing either serine or cysteine and considerably lower cysteine synthase and serine sulfhydrase activities for synthesis of cysteine. The specific activities of CBS for the production of cystathionine from serine and cysteine were $\sim 1500\ \mu mol \cdot h^{-1} \cdot mg^{-1}$ and $\sim 300\ \mu mol \cdot h^{-1} \cdot mg^{-1}$ respectively, and its specific cysteine synthase and serine sulfhydrase activities were $\sim 5\ \mu mol \cdot h^{-1} \cdot mg^{-1}$ and $\sim 30\ \mu mol \cdot h^{-1} \cdot mg^{-1}$ respectively. None of these activities were stimulated by 1 mM AdoMet (results not shown), which is consistent with the absence of a Bateman domain in CBS-1 (see below). These data show that the nematode CBS-1 enzyme exhibits typical CBS activity and that it is not activated by AdoMet.

CBS-1 has a unique structural arrangement

Alignment of the predicted amino acid sequence of CBS-1 with the sequences of previously characterized human, rat, *Drosophila*, trypanosome and yeast CBS enzymes revealed that the *C. elegans* enzyme possesses unique and novel domain architecture. In contrast with other CBSs, *C. elegans* CBS-1 lacks both the haem-binding N-terminus and the entire C-terminus found in other species (Figures 1B and 1C). Moreover, amino acid alignment together with protein modelling revealed that a single polypeptide chain of CBS-1 contains a unique tandem arrangement of two conserved CBS cores that belong to a family of fold-type II PLP-dependent proteins (Figure 1B and 3). Phylogenetic analysis of these two CBS-1 modules revealed that, in contrast with the C-terminal module, the N-terminal module has a lower homology with other CBS enzymes and does not belong to any of the fold-type II PLP-dependent protein families tested (Supplementary Figure S3 at http://www.BiochemJ.org/bj/443/bj4430535add.htm).

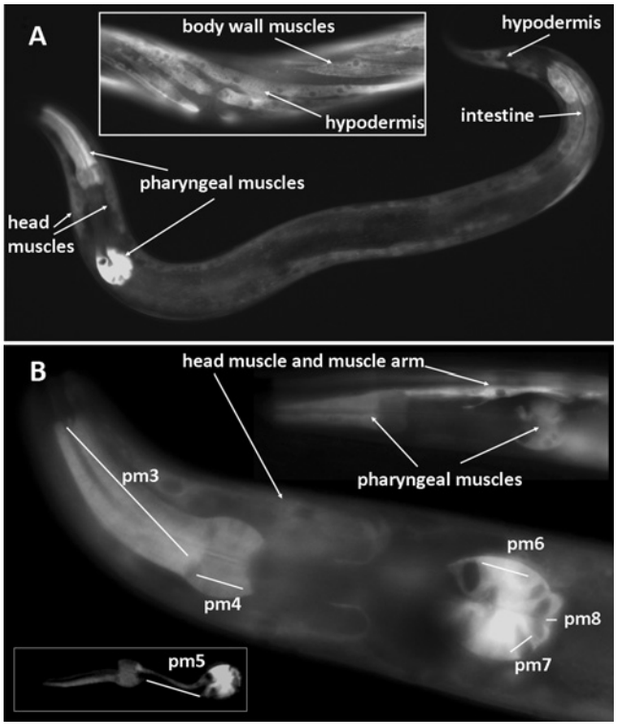


Figure 2 Expression pattern of *cbs-1* in worms

The images show transgenic worms that carry the translational fusion vector cbs-1–GFP. (**A**) L4 larval stage showing the distribution of the GFP signal in the pharyngeal muscles, intestine, hypodermis and muscle cells. The middle part of the adult body (inset) shows the GFP signal in the body wall muscles and hypodermis. (**B**) Head of the worm showing GFP signal in pharyngeal muscles and a head muscle cell with its muscle arm; the GFP signal is distributed in the pharyngeal muscles pm3, pm4, pm6, pm7 and pm8. Some worms also exhibited a GFP signal in pm5 (inset).

Furthermore, the critical PLP-binding lysine residue in the N-terminal module is replaced by a glutamic acid residue (Figures 1C and 3). Analysis of the PLP-binding site using homology modelling with the structure of 45CBS as the template revealed that the fully conserved glycine residue (Gly^{256} in human 45CBS) in the N-terminal module is replaced by a bulky asparagine residue (Asn^{197}) that may sterically affect PLP binding (Figure 3). Thus the *in silico* data strongly suggest that the N-terminal module of nematode CBS-1 cannot bind the PLP essential for the catalytic activity of the enzyme.

The catalytic activity of CBS-1 is mediated only by the C-terminal module

To confirm the hypothesis that the catalytic function of CBS-1 is mediated only by its C-terminal module, we generated individual CBS-1 modules in *E. coli*. While the CBS-1b variant, which lacks the C-terminal module of CBS-1, was highly soluble after expression in *E. coli*, all of the cloned CBS-1 variants without the N-terminal module showed substantially decreased solubility (Figure 4B) that prevented successful purification of proteins containing only the C-terminal module. However, we purified and characterized the CBS-1b and used UV–visible and fluorescence spectrometry to determine that it does not bind PLP; we also found that CBS-1b had virtually undetectable catalytic activity (Figures 4D, 4F and 4G and Table 1). These observations demonstrate that the PLP-binding site essential for catalytic activity is not located in the N-terminal module of the CBS-1 protein. Although CD spectrometry showed that the CBS-1b has a helical secondary structure similar to that of the WT CBS-1 protein (Figure 4E), BN-PAGE revealed that purified CBS-1b precipitates and may form higher-order oligomers (Figure 4C). These data indicate that the presence of both CBS-1 modules in one subunit is necessary for the maintenance of the global structural stability of CBS-1 and/or for proper folding of the recombinant protein.

Because analysis of the isolated N-terminal module indicated possible disruption of its native structure, we also generated and purified the CBS-1 mutants K421A, which abolishes a canonical PLP-binding site in the C-terminal module, and E62K, which creates a putative PLP-binding site in the N-terminal module. We observed altered fluorescence-based tryptophan spectra of the mutant proteins; their relative fluorescence showed different quenching of the tryptophan emission, and the existence of a wavelength maximum shift from ~340 to ~350 nm indicates higher accessibility of the tryptophan residues to the polar solvent in the K421A mutant (Figure 4G and Table 1). In contrast, both mutants retained oligomeric status identical with the WT as determined by BN-PAGE (Figure 4C), and CD measurements showed that the protein's secondary structure is not affected by either the K421A or E62K mutation (Figure 4E). The catalytic activity, PLP saturation and fluorescent properties of the E62K mutant were similar to those of WT (Figures 4D, 4F and 4G), supporting the idea that the structural properties of the N-terminal module do not permit PLP binding even if the canonical lysine residue is present. In contrast, the K421A mutant binds significantly less of the PLP cofactor, as determined by UV–visible absorption spectroscopy. The mutant enzyme's residual affinity for PLP probably results from the formation of an external aldimine; this affinity is manifested in its UV–visible spectrum by the presence of two bands with maxima at 403 and 418 nm and by the lack of the sharp maximum at 412 nm that is typical for internal aldimines (Figure 4F). The formation of an external aldimine in K421A was confirmed by fluorescence spectroscopy and, when excited at 298 nm, the emission spectrum of the mutant protein revealed a significantly higher extent of delayed fluorescence of Schiff bases for K421A in comparison with WT, E62K and CBS-1b (Figure 4G). Enhanced delayed fluorescence due to formation of external aldimines in the active site of the mutant enzyme has also been reported for the bacterial *O*-acetylsulfhydrylase mutant K42A [34]. Taken together, these experiments provide additional evidence that the catalytic activity of *C. elegans* CBS-1 is mediated only by the C-terminal module and that its N-terminal module cannot bind PLP cofactor either as an internal or as an external aldimine.

Analysis of the quaternary structure of nematode CBS-1 suggests a monomeric status of CBS-1

We analysed the quaternary structure of recombinant nematode CBS-1 to determine whether CBS-1 exists as a monomer with a structural arrangement similar to human 45CBS (the C-terminally truncated human CBS that lacks a Bateman domain and forms dimers of 90 kDa [15]), or whether CBS-1 forms dimers or higher-order oligomers. We first performed SEC using the standard proteins ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa) and BSA (66 kDa). To control for possible differences in the Stokes radii of the standard proteins and the CBS-related proteins that may influence their retention on the column, we analysed human 45CBS in parallel. Nematode CBS-1 exhibited a tailing peak with a retention time of 5.776 min (Figure 5A); on the basis of the calibration curve, the apparent native molecular mass of the protein was determined to be ~170 kDa. However, SEC of human 45CBS indicated a native molecular mass of

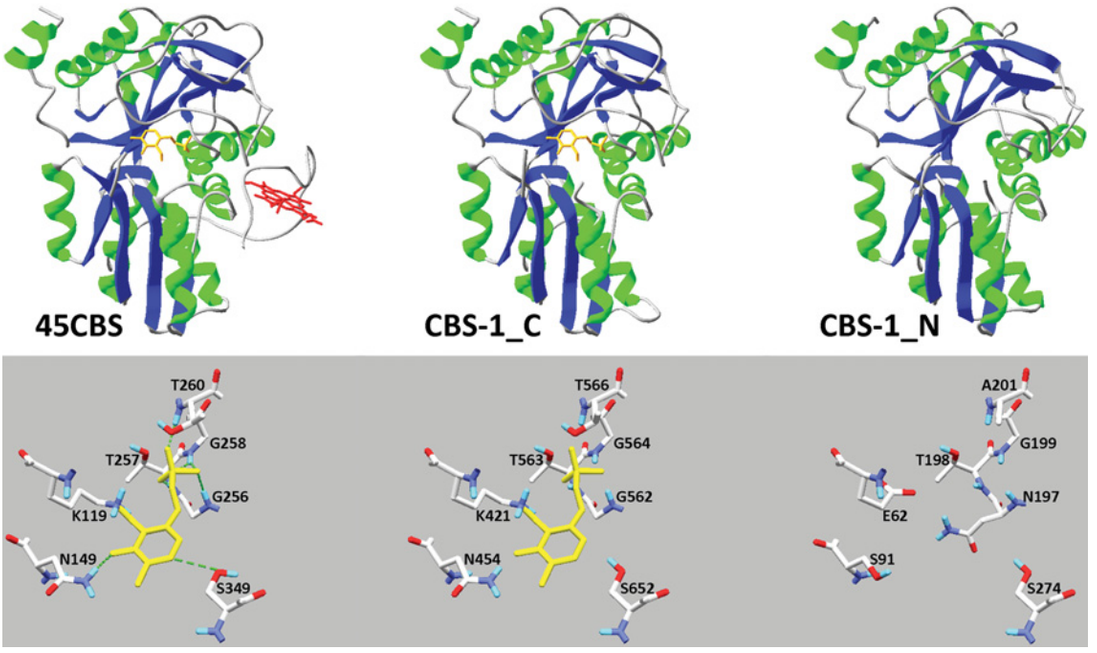


Figure 3 Computationally modelled CBS-1 domains

The images show the fold and PLP-binding site of human 45CBS, C-terminal module of CBS-1 (CBS-1_C) and N-terminal module of CBS-1 (CBS-1_N). The crystal structure of the human enzyme shows hydrogen bonds between amino acid residues and PLP, as indicated by broken green lines. Computational modelling of the individual CBS-1 modules revealed that both modules belong to the family of fold-type II PLP-dependent proteins and that the N-terminal module cannot bind PLP due to the absence of lysine and glycine residues in the consensus PLP-binding pocket.

approximately 150 kDa, suggesting that calibration with standard proteins may result in overestimation of the molecular mass of CBS proteins. According to the molecular mass markers used, SEC yielded ambiguous results compatible with both a monomeric and dimeric structure of CBS-1.

We next used additional techniques including native electrophoresis, BN electrophoresis and chemical cross-linking followed by SDS/PAGE to determine the most likely quaternary structure of CBS-1. These three techniques congruently showed that the 78 kDa nematode CBS-1 exists predominantly as a monomer. The evidence, which is shown in Figures 5(B)–5(D), is as follows: (i) in native PAGE, nematode CBS-1 migrates similarly to the 90 kDa marker of dimeric human 45CBS and between fractions containing monomeric and dimeric BSA respectively (66 kDa and 132 kDa); (ii) on BN electrophoresis, CBS-1 migrates between molecular mass markers of 66 and 140 kDa; and (ii) chemical cross-linking of CBS-1 did not result in changes in protein migration, suggesting modification of amino acid side chains within a single polypeptide chain, whereas human 45CBS readily formed a cross-linked dimeric product with a molecular mass of $\sim$100 kDa. On the basis of these results, we propose that, in contrast with CBS enzymes from other species, recombinant nematode CBS-1 does not form oligomeric structures *in vitro*. Because the conserved catalytic regions of CBS-1 are homologous with each other, we hypothesize that they form an internal interface similar to that formed by subunit dimerization of human 45CBS (Figure 5E).

CBS-1 is more sensitive to denaturation and is more active than human 45CBS

We hypothesized that the above-described differences in the oligomeric assembly of nematode CBS-1 and human 45CBS might result in differences in the energetics of the two proteins. To explore this hypothesis, we used a fluorescence-based thermal-shift assay and pulse proteolysis in a urea gradient. Both approaches revealed significantly lower stability of CBS-1 compared with the human 45CBS; the melting point of CBS-1 was 10 °C lower than that of human 45CBS, and the resistance of CBS-1 to urea-induced unfolding decreased by $\sim$2.8 M (Table 2 and Supplementary Figure S5 at http://www.BiochemJ.org/bj/443/bj4430535add.htm). These data show that the nematode CBS-1 is less energetically stable than the human 45CBS; this finding may be due to a lower energy of the interdomain interface or a higher structural flexibility of the worm CBS-1.

We considered the possibility that the observed structural and energetic differences between the nematode CBS-1 and human 45CBS result in different catalytic properties. We determined the temperature and pH optima and the kinetic parameters for the major CBS reaction, which produces cystathionine from serine and homocysteine. The CBS-1 protein exhibited the highest activity at pH 8.5 and 30 °C (Supplementary Figure S6 at http://www.BiochemJ.org/bj/443/bj4430535add.htm). These conditions are in accordance with the results of the thermal stability assay (see above). We speculate that the lower temperature optimum of CBS-1 compared with the human enzymes (37 °C) may reflect the lower body temperature of nematodes living in the soil. We also found that the affinity of CBS-1 for homocysteine is lower than that of 45CBS (Table 2); however, we observed inhibition of CBS activity at 7.5 and 10 mM homocysteine and this inhibition prevented the activity from increasing to more than $\sim 1500\ \mu\text{mol} \cdot \text{h}^{-1} \cdot \text{mg}^{-1}$ (Supplementary Figure S6). Inhibition by high concentrations of homocysteine has been previously reported for yeast and human CBS enzymes [13,35]. Taken together, these data show that the nematode CBS-1 subunit is approximately 4-fold more active compared with the human 45CBS subunit as expressed by the turnover number (Table 2).

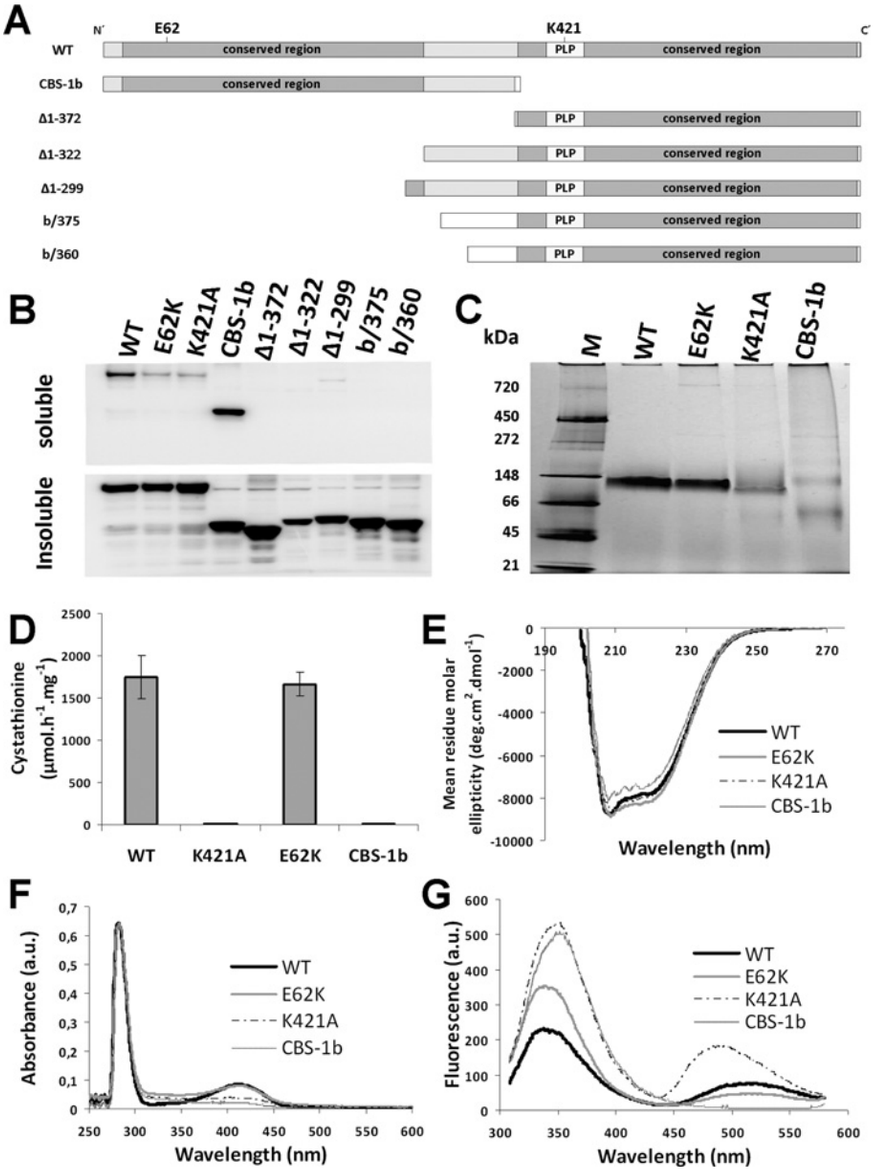


Figure 4 Structural and enzymatic analysis of recombinant CBS-1 variants

(**A**) Illustration of the CBS-1 variants expressed in *E. coli*. (**B**) Detection of CBS-1 variants in *E. coli* lysate after expression using soluble and insoluble fractions separated by centrifugation. (**C**) BN-PAGE of purified recombinant CBS-1 variants shows the monomeric status of the WT, K421A and E62K proteins. The N-terminal domain exhibits monomeric and oligomeric forms. Molecular mass markers are shown in kDa on the left-hand side. (**D**) CBS activity of the purified CBS-1 variants; K421A and CBS-1b have no CBS activity. Results are means ± S.D. (**E**) CD spectra at far-UV show a helical secondary structure for all of the purified CBS-1 variants. (**F**) The UV–visible spectrum of purified recombinant CBS-1 variants of equal concentration shows peaks in the 280 and 412 nm region, indicating light absorption by aromatic amino acids and PLP respectively. Soret peaks typical for haem are not present. (**G**) Emission spectrum after excitation of the tryptophan residues at 298 nm of purified recombinant CBS-1 variants of equal concentration.

CBS-1 mediates nematode development and maintains homocysteine homoeostasis

To explore the functional significance of *cbs-1* in *C. elegans*, we silenced the *cbs-1* gene by RNA-mediated interference and determined the phenotypic consequences of such silencing. To confirm the efficacy of *cbs-1* RNAi, we measured the amounts of CBS-1 antigen and CBS activity in worm extracts of CBS-1-inactivated and WT worms. Western blot analysis using an anti-CBS-1 antibody showed that after RNAi treatment worms exhibited a CBS-1 level that was approximately 10 % that of the control strain (Supplementary Figure S7A at http://www.BiochemJ.org/bj/443/bj4430535add.htm). Although the mean CBS activity of normal worms was 36.0 $\text{nmol} \cdot \text{h}^{-1} \cdot \text{mg}^{-1}$, *cbs-1* RNAi animals exhibited a mean activity of only 5.4 $\text{nmol} \cdot \text{h}^{-1} \cdot \text{mg}^{-1}$, approximately $\sim$15 % of the control level (Supplementary Figure S7B). The results from both Western blot analysis and CBS activity measurement consistently confirmed that the RNAi experiments efficiently reduced the amount and activity of CBS-1.

RNAi resulted in a developmental delay phenotype in 97 % of the worms (515 out of 530 individuals tested). These animals reached the egg-laying adult stage no earlier than the 9th day after embryo hatching, in contrast with control worms, which reached

Table 1 Enzymatic and structural properties of purified CBS-1 variants

ND, not detected, Trp, tryptophan.

Protein	WT	K421A	E62K	CBS-1b
Catalytic activity (μmol · h^{-1} · mg^{-1})	1742 ± 259	0.41 ± 0.03	1663 ± 144	0.27 ± 0.05
Absorption ratio 280/412 nm	7.6	16.1	7.9	29.2
PLP absorption maximum (nm)	412	403 and 418	412	ND
Trp relative fluorescence	233	534	354	505
Trp fluorescence wavelength maximum (nm)	338	350	339	352
Relative delayed fluorescence	76	181	48	ND

the same stage on the 5th day of development. The affected larvae had a shorter body length than the controls (Supplementary Figure S8 at http://www.BiochemJ.org/bj/443/bj4430535add.htm). After RNAi of larvae, the most severe abnormalities were observed in the tissues that express the highest amount of CBS-1 (i.e. gut and pharynx; see the data above on the translational cbs-1–GFP vector). The gut cells of these animals showed reduced pigment granule birefringence under Nomarski optical microscopy (Supplementary Figure 8), and the anterior bulb of the pharynx exhibited abnormal morphology, with a balloon-like appearance and enlarged diameter (Supplementary Figure 8). These data show that CBS-1 is essential for normal development in nematodes.

We determined metabolic flux through the trans-sulfuration pathway by measuring homocysteine, cystathionine and cysteine concentrations in worm homogenates. To eliminate possible differences in metabolic fluxes in worms at various stages of development, the worms were collected at the latest larval developmental stage (L4). The homocysteine and cystathionine levels in *C. elegans* extracts were ~10× and ~1.6× higher in the knockdown strain than in the controls, whereas cysteine concentrations did not differ between the two strains (Figure 6). The observation of elevated homocysteine levels in CBS-1-knockdown worms

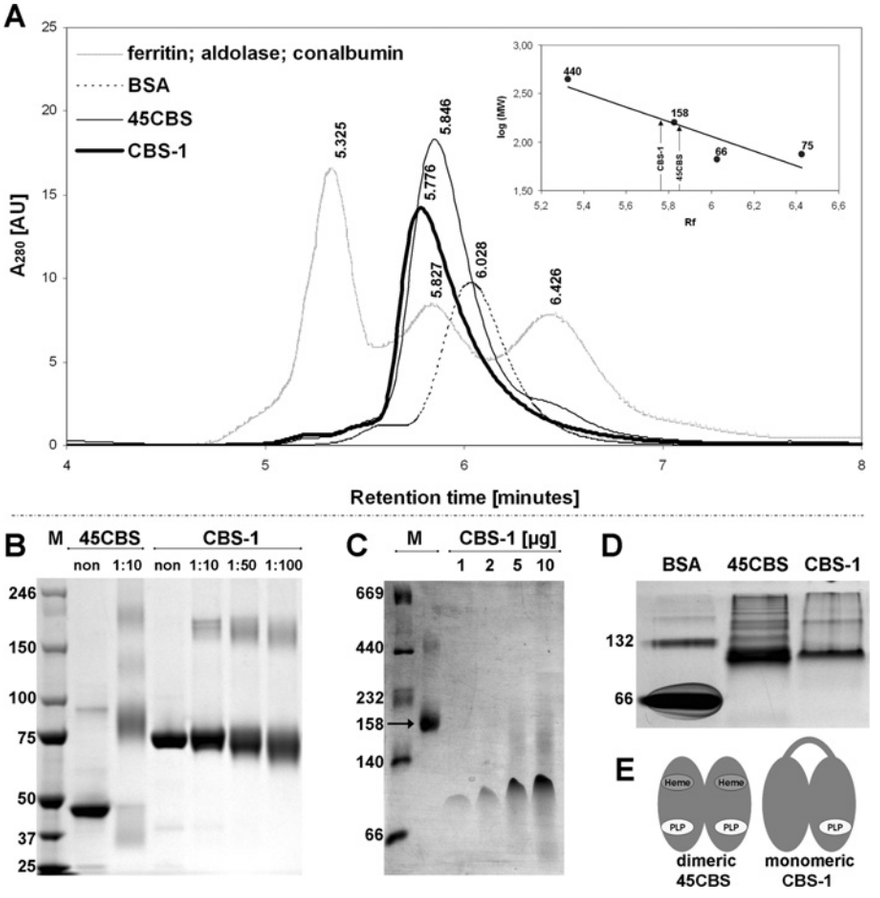


Figure 5 Determination of the quaternary structure of CBS-1

(**A**) SEC. The bold solid curve represents the elution profile of purified recombinant CBS-1, which has a retention time of 5.776 min and an estimated molecular mass of 168 kDa. The thin solid curve represents human 45CBS, which has a retention time of 5.846 min and an estimated molecular mass of 148 kDa. The dashed curve represents BSA with a retention time of 6.028. The grey (dotted) curve represents molecular standards eluted at the following retention times: ferritin (440 kDa), 5.325 min; aldolase (158 kDa), 5.827 min; and conalbumin (70 kDa), 6.426 min. AU, absorbance unit. (**B**) Cross-linking. Purified CBS-1 and human 45CBS were cross-linked with BS^3 in appropriate molar ratios of protein/modifier, as indicated in the Figure, and subjected to SDS/PAGE. In contrast with human 45CBS, which forms dimers, the mobility of CBS-1 does not change after cross-linking. (**C**) BN-PAGE. CBS-1, with a molecular mass of 78 kDa (four different amounts of loaded protein), migrates between molecular mass markers of 66 kDa and 140 kDa. The molecular protein mass markers include thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa), BSA (66 kDa) and aldolase (158 kDa). (**D**) Native PAGE. CBS-1, with a molecular mass of 78 kDa, migrates between molecular mass markers of 66 kDa and 132 kDa, similar to a ~90 kDa dimer of human 45CBS. In (**B**–**D**) the molecular mass is given in kDa on the left-hand side. M, marker. (**E**) Schematic diagram of the hypothetical quaternary structure of CBS-1 and comparison of its structure with that of human 45CBS.

Table 2 Stability and enzymatic properties of CBS-1 compared with human 45CBS

c_m, concentration of urea at which a fraction of folded proteins comprises 50 % of the entire protein population.

Protein	Nematode CBS-1	Human 45CBS
Oligomeric status	Monomer	Dimer
Michaelis constant K_m (mM)		
Serine	5.57 ± 0.68	2.20 ± 0.46†
Homocysteine	4.29 ± 0.97	0.33 ± 0.07†
Turnover number k_{cat} (s^{-1})		
Serine	48.12 ± 2.95	13.81 ± 0.88†
Homocysteine	43.31 ± 4.33	10.88 ± 0.72†
Catalytic efficiency k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)		
Serine	~8.5	~6
Homocysteine	~10	26.97 ± 5.87†
Characteristics of protein denaturation		
Midpoint of urea concentration c_m (M)	1.21 ± 0.02	4.08 ± 0.07*
Melting point T_m (°C)	41	51

*Value from Hnizda et al. [27].
†Value from Frank et al. [23].

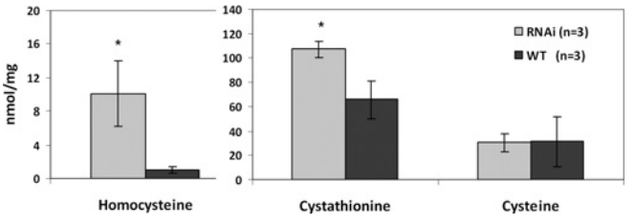


Figure 6 Metabolite levels in crude extracts of CBS-1-knockdown animals

The concentration of metabolites (in nanomoles per mg of protein) from RNAi and control experiments respectively. Homocysteine and cystathionine concentrations in CBS-1-deficient worms are significantly higher (10.1 and 107.1 nmol/mg of protein respectively) than in control worms (1.0 and 65.6 nmol/mg of protein respectively). Results are means ± S.D. from three independent experiments. *$P < 0.05$, as determined by Student's t test.

strongly supports an important role of this enzyme in maintaining homocysteine homoeostasis in *C. elegans*.

DISCUSSION

Evidence that CBS in *C. elegans* is encoded by *cbs-1*

In the present paper, we identified the *cbs-1* gene in *C. elegans*, which encodes an enzyme with cystathionine-synthesizing activity and is important for normal development of the nematode. Using a BLASTp search against a *C. elegans* protein database, we identified two nematode genes, *ZC373.1* and *F54A3.4*, that are highly homologous with CBS genes in other species. However, several lines of evidence demonstrate that only *ZC373.1* encodes a CBS, whereas *F54A3.4* is probably a pseudogene. The gene denoted *F54A3.4* has not been detected by RT–PCR, was not found by *in silico* searches in the appropriate EST and proteomic databases, and its partial deletion did not elicit biochemical or morphological phenotypic abnormalities. In contrast, *ZC373.1* mRNA has been detected by RT–PCR, its ESTs and peptides were annotated in appropriate databases, the enzyme was shown to be expressed endogenously in its entire length, and its knockdown resulted in severe biochemical and phenotypic consequences. Most importantly, the purified CBS-1 enzyme exhibited enzymatic properties consistent with previously characterized CBS enzymes from other species.

The unique domain architecture of nematode CBS enzymes

In silico analysis of the CBS-1 protein sequence showed that the CBS-1 of *C. elegans* possesses a unique multi-domain architecture that has not been reported previously for any other CBS. The unusual structure of CBS-1 includes the lack of a haem-binding region, the lack of a Bateman domain and the tandem arrangement of two conserved catalytic regions of which only the C-terminal region is catalytically active. Such a domain arrangement of predicted CBS enzymes in fully sequenced organisms has been found only in organisms from the nematode phylum, showing an evolutionarily divergent arrangement of the CBS protein in this phylum. Interestingly, the nematode *Loa loa* possesses a PLP-binding site in both CBS modules (Supplementary Figure S9 at http://www.BiochemJ.org/bj/443/bj4430535add.htm), suggesting that the unusual and unique structure of CBS enzymes in nematodes probably originates from a duplication of the conserved catalytic region in a common ancestor followed by mutations abolishing PLP binding in the N-terminal module.

To our knowledge there is no evidence, except of nematode CBS proteins, regarding fold-type II PLP-dependent proteins lacking a PLP-binding site. Thus the function, if any, of the N-terminal module in the nematode CBS-1 protein remains unclear. Several pieces of experimental evidence obtained in the present study clearly show that this module does not have canonical catalytic function. We speculate that mutations in this portion of the nematode CBS enzyme may have permitted the acquisition of novel structural and functional properties, such as changes in protein stability and folding, protein–protein interactions, or regulation of enzyme activity. Studies of truncated variants suggest that the N-terminal module may be important for proper folding and subsequent stability of CBS-1 (see above). Because CBS-1 forms a monomer, it is probable that its N-terminal and C-terminal modules interact to form a structure similar to that of the human 45CBS dimer. However, the proposed interdomain interaction cannot be sufficiently supported by the computational modelling procedures using previously solved crystal structures of CBS proteins, and thus it requires further study of the three-dimensional spatial arrangement of CBS-1 at atomic resolution. The N-terminal module of CBS-1 may also have a regulatory role. The existence of tandem duplicated conserved modules of which only one is catalytically active in a single polypeptide is similar to the well-known case of tyrosine protein kinases [JAKs (Janus kinases)]. In tandem with a catalytically active kinase domain, these kinases have a catalytically inactive pseudo-kinase domain that has been implicated in the regulation of their activity [36]. Alternatively, the N-terminal module of CBS-1 may also play a role in protein–protein interactions, such as the interactions with the SUMOylation enzyme apparatus or huntingtin that have been described for human and rodent CBS-1 orthologues respectively [37,38].

More intriguingly, expression of the spliced variant *cbs-1b* shows that the N-terminal module of CBS-1 can be produced *in vivo* without the catalytic C-terminal module. This finding suggests that the non-catalytic module may play a role in additional biological processes independent of the catalytic module. However, it should be noted that misspliced variants with premature stop codons are commonly targeted by a cellular RNA nonsense-mediated decay mechanism [39]; therefore, the existence of a separate nematode N-terminal domain *in vivo* should be investigated in future studies.

Possible biological roles of CBS-1 enzymatic activity in *C. elegans*

Because *cbs-1* is expressed in a limited number of tissues, it is tempting to speculate on the role of this enzyme in the organs

in which it is expressed. High expression of *cbs-1* was observed throughout post-embryonic development in the intestine, which is characterized by secretion of digestive enzymes and high metabolic activity in *C. elegans*, such as the synthesis and storage of macromolecules and detoxification of xenobiotics [40]. Thus the expression of *cbs-1* in *C. elegans* intestine may mirror the high expression of CBS in the mammalian liver, pancreas and small intestine, in which CBS plays an important role in homocysteine homoeostasis and/or in the provision of cysteine for glutathione production [41]. We hypothesize that the intestinal expression of *cbs-1* in worms may serve similar purposes, namely removal of homocysteine or cysteine biosynthesis.

Expression of *cbs-1* has also been observed in pharyngeal muscles and hypodermis. Because neither of these tissues shows high metabolic activity compared with the intestine, there are other possible explanations for CBS-1 activity in these tissues. Because both hypodermal cells and pharyngeal muscle cells secrete cuticle (http://www.wormatlas.org/), we propose that CBS-1 may provide cysteine, which is important for cuticle formation and its stabilization by disulfide bonds [42]. Another possible role for CBS-1 in muscle and hypodermis is the production of the neuromodulator and smooth muscle relaxant hydrogen sulfide [43]. The endogenous biosynthesis of H_2S via CBS might serve for smooth muscle relaxation in the strongly innervated nematode pharynx or in regulating the expression of HIF-1 (hypoxia-inducible factor 1) target genes in the hypodermis [44]. Interestingly, although CBS is thought to be the main enzyme that produces hydrogen sulfide in the mammalian brain [45], we did not observe a GFP signal in neurons. This finding suggests that the endogenous production of hydrogen sulfide in *C. elegans* neurons is mediated by different enzymes than in other species or that the role of hydrogen sulfide in *C. elegans* neurons is negligible.

C. elegans as a model of CBS deficiency

Because many genes implicated in human diseases are well-conserved across phyla [46], *C. elegans* is considered by many investigators to be a suitable model for studying cellular and metabolic mechanisms in selected genetic disorders [47–49]. In addition to its low cost of maintenance and short generation time, other advantages of the *C. elegans* model include the possibility of observing cellular processes *in vivo* and of easily screening for the effects of novel therapies [50,51]. In the present study, we examined the morphological and biochemical effects of nematode CBS-1 deficiency. These effects may in part recapitulate the human disease homocystinuria, which results from CBS deficiency. Homocystinuria is characterized by increased tissue, plasma and urinary concentrations of homocysteine, and by decreased concentrations of cystathionine and cysteine [52,53]. Its clinical features include liver steatosis, connective tissue disorder, thromboembolism and various degrees of central nervous system involvement [53]. In our CBS-1–GFP localization study, CBS-1-knockdown worms exhibited abnormal morphology of several tissues that express the *cbs-1* gene. Using light microscopy, we observed a reduced birefringent signal from pigment gut granules, which are considered to be lysosome-related organelles [54]. Although the function and composition of these granules has not been fully elucidated, the abnormal pattern of gut granules in CBS-1-knockdown animals may in part correspond to the liver steatosis observed in murine and human CBS deficiency. Furthermore, the observed abnormal pharyngeal morphology of CBS-1-deficient worms may possibly correspond to some of the neurological sequelae of human CBS deficiency. It appears that the CBS-knockdown nematodes produced in the present study may in part recapitulate some of the features of human homocystinuria due to CBS deficiency.

In the CBS-1-deficient nematodes produced in the present study, the amounts of CBS-1 antigen and enzyme activity decreased to ~10–15 % those of WT worms. This degree of enzyme deficiency resulted in an approximately 10-fold increase in homocysteine concentrations in worm extracts compared with the WT strain, demonstrating an essential role for CBS-1 in maintaining homocysteine homoeostasis in *C. elegans*. Because exposure of worms to homocysteine in medium [55] leads to a similar developmental delay as the *cbs-1* RNAi in the present study, it is conceivable that high tissue levels of homocysteine may be directly responsible for the developmental delay phenotype that we observed. Surprisingly, and in contrast with human patients with CBS deficiency [53], cystathionine levels in CBS-1-deficient worms were only slightly increased. However, a similar elevation in plasma cystathionine was reported for one murine model of CBS deficiency [56]. We hypothesize that elevated cystathionine in CBS-1-deficient worms may be caused by three possible mechanisms: (i) elevated homocysteine may inactivate CGL, as proposed previously for a murine model of CBS deficiency [56]; (ii) elevated homocysteine may lead to formation of cystathionine via condensation of cysteine and homocysteine by CGL [57]; and (iii) cystathionine may be synthesized by a hypothetical cystathionine γ-synthase in the reverse trans-sulfuration pathway using cysteine and *O*-succinylhomoserine. Moreover, the CBS-1-deficient worms observed in the present study did not exhibit cysteine depletion, which is a common feature of human CBS deficiency. We hypothesize that cysteine levels in deficient worms are maintained by sufficient cysteine intake from *E. coli* or by biosynthesis of cysteine via a hypothetical sulfur assimilation pathway because *C. elegans* possesses several bacterial and plant cysteine synthase homologues (see above, [58]).

AUTHOR CONTRIBUTION

Roman Vozdek designed and performed most of the experiments and wrote the first draft of the paper; Aleš Hnízda purified and further characterized recombinant CBS-1; Jakub Krijt measured aminothiols and cystathionine by HPLC and LC–MS/MS in the appropriate studies; Marta Kostrouchová co-ordinated the experiments with *C. elegans*; and Viktor Kožich co-ordinated the whole project. All authors have extensively revised various versions of the paper and approved its final version prior to submission.

ACKNOWLEDGEMENTS

We thank Dr A. Fire for vector pPD95.75, and Eva Zouharová, Mrs Kateřina Raková, Hana Prouzová, Jitka Honzíková and Professor Milan Kodíček for technical assistance and advice.

FUNDING

This work was supported by a Wellcome Trust International Senior Research Fellowship in Biomedical Science in Central Europe [number 070255/Z/03/Z], the Grant Agency of the Charles University in Prague [grant numbers 21709 and SVV262502], the Ministry of Education of the Czech Republic [grant number MSM0021620806] and the Czech Science Foundation [grant number 304/08/0970].

REFERENCES

1 Finkelstein, J. D. (1998) The metabolism of homocysteine: pathways and regulation. Eur. J. Pediatr. **157**, S40–S44

1a Taoka, S., Ohja, S., Shan, X., Kruger, W. D. and Banerjee, R. (1998) Evidence for heme-mediated redox regulation of human cystathionine beta-synthase activity. J. Biol. Chem. **273**, 25179–25184

2 Banerjee, R. and Zou, C. G. (2005) Redox regulation and reaction mechanism of human cystathionine-beta-synthase: a PLP-dependent hemesensor protein. Arch. Biochem. Biophys. **433**, 144–156

3 Cherney, M. M., Pazicni, S., Frank, N., Marvin, K. A., Kraus, J. P. and Burstyn, J. N. (2007) Ferrous human cystathionine beta-synthase loses activity during enzyme assay due to a ligand switch process. Biochemistry **46**, 13199–13210

4 Smith, A. T., Majtan, T., Freeman, K. M., Su, Y., Kraus, J. P. and Burstyn, J. N. (2011) Cobalt cystathionine beta-synthase: a cobalt-substituted heme protein with a unique thiolate ligation motif. Inorg. Chem. **50**, 4417–4427

5 Mehta, P. K. and Christen, P. (2000) The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes. Adv. Enzymol. Relat. Areas Mol. Biol. **74**, 129–184

6 Kery, V., Poneleit, L. and Kraus, J. P. (1998) Trypsin cleavage of human cystathionine beta-synthase into an evolutionarily conserved active core: structural and functional consequences. Arch. Biochem. Biophys. **355**, 222–232

7 Taoka, S., Widjaja, L. and Banerjee, R. (1999) Assignment of enzymatic functions to specific regions of the PLP-dependent heme protein cystathionine beta-synthase. Biochemistry **38**, 13155–13161

8 Skovby, F., Kraus, J. P. and Rosenberg, L. E. (1984) Biosynthesis and proteolytic activation of cystathionine beta-synthase in rat liver. J. Biol. Chem. **259**, 588–593

9 Zou, C. G. and Banerjee, R. (2003) Tumor necrosis factor-α-induced targeted proteolysis of cystathionine beta-synthase modulates redox homeostasis. J. Biol. Chem. **278**, 16802–16808

10 Jhee, K. H., McPhie, P. and Miles, E. W. (2000) Yeast cystathionine beta-synthase is a pyridoxal phosphate enzyme but, unlike the human enzyme, is not a heme protein. J. Biol. Chem. **275**, 11541–11544

11 Nozaki, T., Shigeta, Y., Saito-Nakano, Y., Imada, M. and Kruger, W. D. (2001) Characterization of transsulfuration and cysteine biosynthetic pathways in the protozoan hemoflagellate, *Trypanosoma cruzi*. Isolation and molecular characterization of cystathionine beta-synthase and serine acetyltransferase from Trypanosoma. J. Biol. Chem. **276**, 6516–6523

12 Koutmos, M., Kabil, O., Smith, J. L. and Banerjee, R. (2010) Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine beta-synthase. Proc. Natl. Acad. Sci. U.S.A. **107**, 20958–20963

13 Jhee, K. H., McPhie, P. and Miles, E. W. (2000) Domain architecture of the heme-independent yeast cystathionine beta-synthase provides insights into mechanisms of catalysis and regulation. Biochemistry **39**, 10548–10556

14 Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. Genetics **77**, 71–94

15 Meier, M., Janosik, M., Kery, V., Kraus, J. P. and Burkhard, P. (2001) Structure of human cystathionine beta-synthase: a unique pyridoxal 5′-phosphate-dependent heme protein. EMBO J. **20**, 3910–3916

16 Wiederstein, M. and Sippl, M. J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. **35**, W407–W410

17 Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis **18**, 2714–2723

18 Neron, B., Menager, H., Maufrais, C., Joly, N., Maupetit, J., Letort, S., Carrere, S., Tuffery, P. and Letondal, C. (2009) Mobyle: a new full web bioinformatics framework. Bioinformatics **25**, 3005–3011

19 Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. **31**, 3497–3500

20 Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**, 696–704

21 Huson, D. H., Richter, D. C., Rausch, C., Dezulian, T., Franz, M. and Rupp, R. (2007) Dendroscope: an interactive viewer for large phylogenetic trees. BMC Bioinf. **8**, 460

22 Boulin, T., Etchberger, J. F. and Hobert, O. (2006) Reporter gene fusions. In WormBook (The *C. elegans* Research Community, ed.), pp. 1–23, doi/10.1895/wormbook.1.106.1

23 Frank, N., Kent, J. O., Meier, M. and Kraus, J. P. (2008) Purification and characterization of the wild type and truncated human cystathionine beta-synthase enzymes expressed in *E. coli*. Arch. Biochem. Biophys. **470**, 64–72

24 Janosik, M., Meier, M., Kery, V., Oliveriusova, J., Burkhard, P. and Kraus, J. P. (2001) Crystallization and preliminary X-ray diffraction analysis of the active core of human recombinant cystathionine beta-synthase: an enzyme involved in vascular disease. Acta Crystallogr. Sect. D Biol. Crystallogr. **57**, 289–291

25 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**, 680–685

26 Wittig, I., Braun, H. P. and Schagger, H. (2006) Blue native PAGE. Nat. Protoc. **1**, 418–428

27 Hnizda, A., Spiwok, V., Jurga, V., Kozich, V., Kodicek, M. and Kraus, J. P. (2010) Cross-talk between the catalytic core and the regulatory domain in cystathionine beta-synthase: study by differential covalent labeling and computational modeling. Biochemistry **49**, 10526–10534

28 Maclean, K. N., Sikora, J., Kozich, V., Jiang, H., Greiner, L. S., Kraus, E., Krijt, J., Crnic, L. S., Allen, R. H., Stabler, S. P. et al. (2010) Cystathionine beta-synthase null homocystinuric mice fail to exhibit altered hemostasis or lowering of plasma homocysteine in response to betaine treatment. Mol. Genet. Metab. **101**, 163–171

29 Krijt, J., Kopecka, J., Hnizda, A., Moat, S., Kluijtmans, L. A., Mayne, P. and Kozich, V. (2011) Determination of cystathionine beta-synthase activity in human plasma by LC–MS/MS: potential use in diagnosis of CBS deficiency. J. Inherited Metab. Dis. **34**, 49–55

30 Mello, C. and Fire, A. (1995) DNA transformation. Methods Cell Biol. **48**, 451–482

31 Janosik, M., Oliveriusova, J., Janosikova, B., Sokolova, J., Kraus, E., Kraus, J. P. and Kozich, V. (2001) Impaired heme binding and aggregation of mutant cystathionine beta-synthase subunits in homocystinuria. Am. J. Hum. Genet. **68**, 1506–1513

32 Schrimpf, S. P., Weiss, M., Reiter, L., Ahrens, C. H., Jovanovic, M., Malmstrom, J., Brunner, E., Mohanty, S., Lercher, M. J., Hunziker, P. E. et al. (2009) Comparative functional analysis of the *Caenorhabditis elegans* and *Drosophila melanogaster* proteomes. PLoS Biol. **7**, e48

33 McKay, S. J., Johnsen, R., Khattra, J., Asano, J., Baillie, D. L., Chan, S., Dube, N., Fang, L., Goszczynski, B., Ha, E. et al. (2003) Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans*. Cold Spring Harb. Symp. Quant. Biol. **68**, 159–169

34 Rege, V. D., Kredich, N. M., Tai, C. H., Karsten, W. E., Schnackerz, K. D. and Cook, P. F. (1996) A change in the internal aldimine lysine (K42) in *O*-acetylserine sulfhydrylase to alanine indicates its importance in transimination and as a general base catalyst. Biochemistry **35**, 13485–13493

35 Belew, M. S., Quazi, F. I., Willmore, W. G. and Aitken, S. M. (2009) Kinetic characterization of recombinant human cystathionine beta-synthase purified from *E. coli*. Protein Expression Purif. **64**, 139–145

36 Aringer, M., Cheng, A., Nelson, J. W., Chen, M., Sudarshan, C., Zhou, Y. J. and O'Shea, J. J. (1999) Janus kinases and their role in growth and disease. Life Sci. **64**, 2173–2186

37 Kabil, O., Zhou, Y. and Banerjee, R. (2006) Human cystathionine beta-synthase is a target for sumoylation. Biochemistry **45**, 13528–13536

38 Boutell, J. M., Wood, J. D., Harper, P. S. and Jones, A. L. (1998) Huntingtin interacts with cystathionine beta-synthase. Hum. Mol. Genet. **7**, 371–378

39 Zahler, A. M. (2005) Alternative splicing in *C. elegans*. In WormBook (The *C. elegans* Research Community, ed.), pp. 1–13, doi/10.1895/wormbook.1.31.1

40 McGhee, J. D. (2007) The *C. elegans* intestine. In WormBook (The *C. elegans* Research Community, ed.). pp. 1–36, doi/10.1895/wormbook.1.133.1

41 Finkelstein, J. D. (2000) Pathways and regulation of homocysteine metabolism in mammals. Semin. Thromb. Hemostasis **26**, 219–225

42 Page, A. P. and Johnstone, I. L. (2007) The cuticle. In WormBook (The *C. elegans* Research Community, ed.). pp. 1–15, doi/10.1895/wormbook.1.138.1

43 Kimura, H. (2011) Hydrogen sulfide: its production, release and functions. Amino Acids **41**, 113–121

44 Budde, M. W. and Roth, M. B. (2010) Hydrogen sulfide increases hypoxia-inducible factor-1 activity independently of von Hippel–Lindau tumor suppressor-1 in *C. elegans*. Mol. Biol. Cell **21**, 212–217

45 Abe, K. and Kimura, H. (1996) The possible role of hydrogen sulfide as an endogenous neuromodulator. J. Neurosci. **16**, 1066–1071

46 Kuwabara, P. E. and O'Neil, N. (2001) The use of functional genomics in *C. elegans* for studying human development and disease. J. Inherited Metab. Dis. **24**, 127–138

47 Chandler, R. J., Aswani, V., Tsai, M. S., Falk, M., Wehrli, N., Stabler, S., Allen, R., Sedensky, M., Kazazian, H. H. and Venditti, C. P. (2006) Propionyl-CoA and adenosylcobalamin metabolism in *Caenorhabditis elegans*: evidence for a role of methylmalonyl-CoA epimerase in intermediary metabolism. Mol. Genet. Metab. **89**, 64–73

48 Calvo, A. C., Pey, A. L., Ying, M., Loer, C. M. and Martinez, A. (2008) Anabolic function of phenylalanine hydroxylase in *Caenorhabditis elegans*. FASEB J. **22**, 3046–3058

49 Fisher, A. L., Page, K. E., Lithgow, G. J. and Nash, L. (2008) The *Caenorhabditis elegans K10C2.4* gene encodes a member of the fumarylacetoacetate hydrolase family: a *Caenorhabditis elegans* model of type I tyrosinemia. J. Biol. Chem. **283**, 9127–9135

50 Link, E. M., Hardiman, G., Sluder, A. E., Johnson, C. D. and Liu, L. X. (2000) Therapeutic target discovery using *Caenorhabditis elegans*. Pharmacogenomics **1**, 203–217

51 Kaletta, T. and Hengartner, M. O. (2006) Finding function in novel targets: *C. elegans* as a model organism. Nat. Rev. **5**, 387–398

52 Kraus, J. P., Janosik, M., Kozich, V., Mandell, R., Shih, V., Sperandeo, M. P., Sebastio, G., de Franchis, R., Andria, G., Kluijtmans, L. A. et al. (1999) Cystathionine beta-synthase mutations in homocystinuria. Hum. Mutat. **13**, 362–375

53 Kraus, J. P. and Kožich, V. (2001) Cystathionine-β-synthase and its deficiency. In Homocysteine in Health and Disease (Carmel, R. and Jacobsen, D. W., eds), pp. 223–243, Cambridge University Press, Cambridge

54 Hermann, G. J., Schroeder, L. K., Hieb, C. A., Kershner, A. M., Rabbitts, B. M., Fonarev, P., Grant, B. D. and Priess, J. R. (2005) Genetic analysis of lysosomal trafficking in *Caenorhabditis elegans*. Mol. Biol. Cell **16**, 3273–3288

55 Khare, S., Gomez, T., Linster, C. L. and Clarke, S. G. (2009) Defective responses to oxidative stress in protein l-isoaspartyl repair-deficient *Caenorhabditis elegans*. Mech. Ageing Dev. **130**, 670–680

56 Maclean, K. N., Sikora, J., Kozich, V., Jiang, H., Greiner, L. S., Kraus, E., Krijt, J., Overdier, K. H., Collard, R., Brodsky, G. L. et al. (2010) A novel transgenic mouse model of CBS-deficient homocystinuria does not incur hepatic steatosis or fibrosis and exhibits a hypercoagulative phenotype that is ameliorated by betaine treatment. Mol. Genet. Metab. **101**, 153–162

57 Singh, S., Padovani, D., Leslie, R. A., Chiku, T. and Banerjee, R. (2009) Relative contributions of cystathionine beta-synthase and gamma-cystathionase to H2S biogenesis via alternative trans-sulfuration reactions. J. Biol. Chem. **284**, 22457–22466

58 Budde, M. W. and Roth, M. B. (2011) The response of *Caenorhabditis elegans* to hydrogen sulfide and hydrogen cyanide. Genetics **189**, 521–532

Received 15 August 2011/13 December 2011; accepted 13 January 2012
Published as BJ Immediate Publication 13 January 2012, doi:10.1042/BJ20111478

Biochem. J. (2012) **443**, 535–547 (Printed in Great Britain) doi:10.1042/BJ20111478

online data

SUPPLEMENTARY ONLINE DATA

Novel structural arrangement of nematode cystathionine β-synthases: characterization of *Caenorhabditis elegans* CBS-1

Roman VOZDEK, Aleš HNÍZDA, Jakub KRIJT, Marta KOSTROUCHOVÁ and Viktor KOŽICH[1]

Institute of Inherited Metabolic Disorders, Charles University in Prague, First Faculty of Medicine and General University Hospital, Ke Karlovu 2, 128 08, Praha 2, Czech Republic

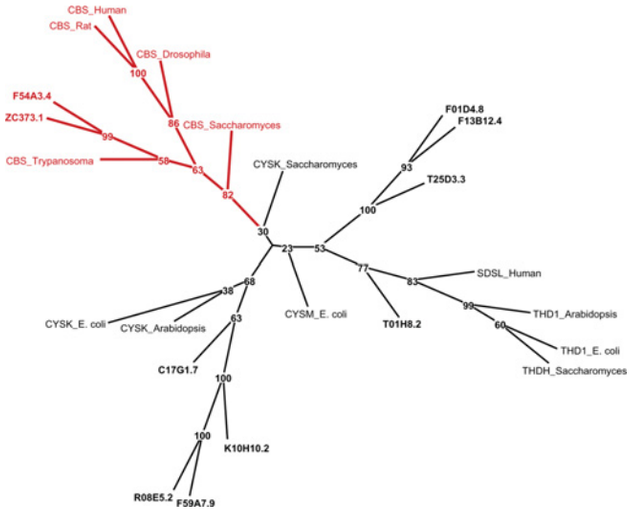


Figure S1 Unrooted tree of fold-type II PLP-dependent proteins with ten CBS homologues in *C. elegans*

For phylogenetic analysis, we used various proteins from the family of fold-type II PLP-dependent proteins: CBS_Human (UniProt entry P35520), CBS_Rat (UniProt entry P32232), CBS_Drosophila (UniProt entry Q9VRD9), CBS_Trypanosoma (UniProt entry Q9BH24), CBS_Saccharomyces (UniProt entry P32582), CYSK_Saccharomyces (UniProt entry P53206) CYSK_Arabidopsis (UniProt entry P47998), CYSM_E. coli (UniProt entry P16703), CYSK_E. coli (UniProt entry P0ABK5), SDSL_Human (UniProt entry Q96GA7), THDH_Saccharomyces (UniProt entry P00927), THD1_E. coli (UniProt entry P04968) and THD1_Arabidopsis (UniProt entry Q9ZSS6). Ten CBS homologues (bold font) are presented in Table S1. The numbers at the internal nodes represent bootstrapped values (maximum 100). The upper left-hand edge in red denotes the CBS branch. The tree topology demonstrates three separated groups for ten CBS homologues in *C. elegans*: ZC373.1 and F54A3.4 belong to the CBS branch, C17G1.7, R08E5.2, F59A7.9 and K10H10.2 belong to the cysteine synthase A branch, and the remaining homologues belong to other fold-type II PLP-dependent protein families. CYSK, cysteine synthase A; CYSM, cysteine synthase B; SDSL/THD, serine/threonine dehydratase family.

[1] To whom correspondence should be addressed (email Viktor.Kozich@LF1.cuni.cz).

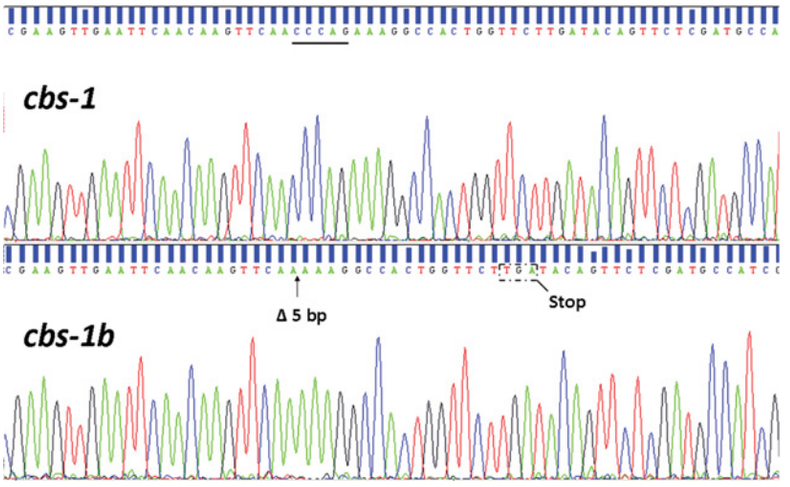


Figure S2 Sequences of *cbs*-1 RT–PCR products covering the exon 6 and 7 junction

Two splice variants of *cbs-1* were found. The novel *cbs-1b* transcript leads to a frameshift and a subsequent stop codon that allows for translation of the separated N-terminal module of CBS-1.

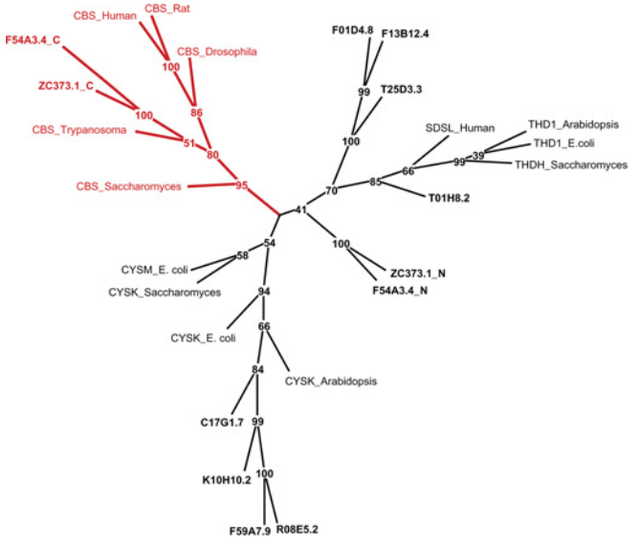


Figure S3 Unrooted tree of separated conserved regions from CBS-1 with various fold-type II PLP-dependent proteins

This tree is taken from the same phylogenetic study as that presented in Figure S1, with the exception that the two conserved regions from CBS orthologues ZC373.1 and F54A3.4, were separated for the analysis. The topology of the unrooted tree demonstrates that the N-terminal regions of CBS homologues (ZC373.1_N and F54A3.4_N) do not belong to any branch containing the analysed proteins, whereas the C-terminal regions (ZC373.1_C and F54A3.4_C) belong to the CBS branch. Alignment of the N-terminal module of *C. elegans* CBS-1 (the first of the two tandemly arranged regions, i.e. residues 14–322) revealed 29 % identity and an e-value of 5e − 27 compared with the catalytic core of human CBS (residues 72–397), whereas the C-terminal module (i.e. residues 374–702 of CBS-1) revealed much higher 54 % identity and an e-value of 2e − 99.

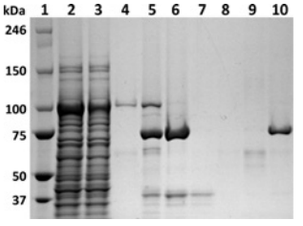


Figure S4 CBS-1 purification procedure

The purification procedure for CBS-1 is illustrated by a 3–8 % SDS-containing polyacrylamide gel stained with Coomassie EZ Blue. Lane 1, molecular mass markers; lane 2, bacterial extract after centrifugation; lane 3, flow-through fraction from glutathione–Sepharose column; 4, wash fraction of glutathione–Sepharose column; lane 5, fusion protein that was cleaved by PreScission protease; lane 6, elution of CBS-1 after on-column cleavage; lane 7, flow-through fraction from the Ni-Sepharose column; lanes 8 and 9, wash fractions of the Ni-Sepharose column by IMAC buffer containing 20 mM and 50 mM imidazole respectively; lane 10, elution of CBS-1 by IMAC buffer containing 75 mM imidazole. The molecular mass is given in kDa on the left-hand side.

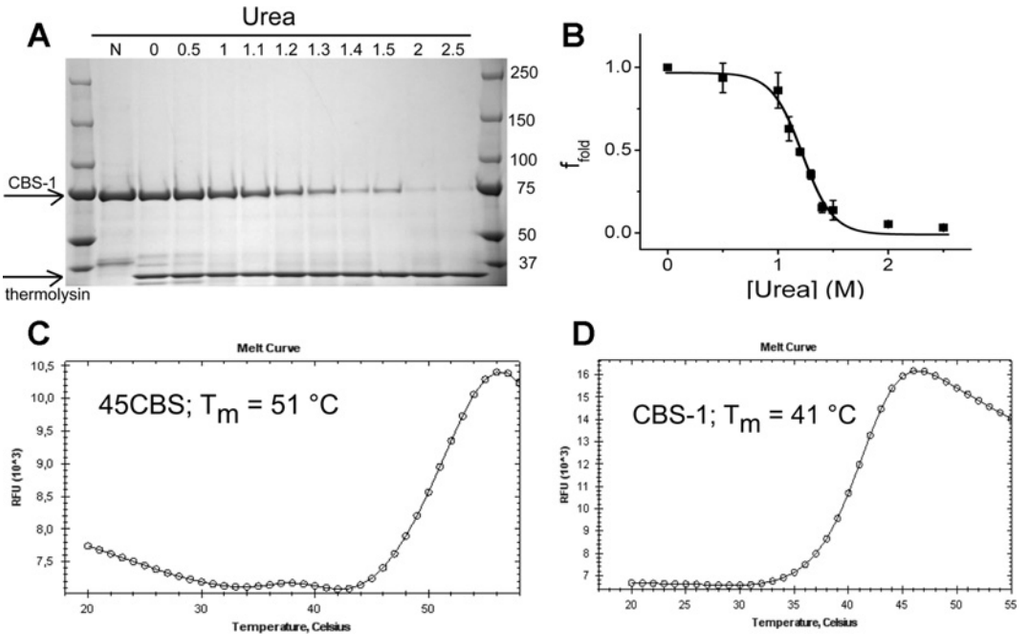


Figure S5 Pulse proteolysis and fluorescence-based thermal-shift assay

Pulse proteolysis in a urea gradient employing thermolysin and thermal-based assays were used to determine possible differences in enzyme stability between CBS-1 and human 45CBS. (**A**) Representative SDS/PAGE gel. The molar concentration of urea for the proteolytic pulse is indicated at the top of each lane and the molecular mass is given in kDa on the right-hand side. (**B**) F-fold values, which represent the fraction of the remaining intact protein after the proteolytic pulse, are plotted against the urea concentration. Results are means ± S.D. from three measurements and the curves were fitted by non-linear regression. (**C** and **D**) Melting curves in fluorescence-based thermal shift assays reveal melting points (T_m) of 51 °C and 41 °C for human 45CBS and CBS-1 respectively.

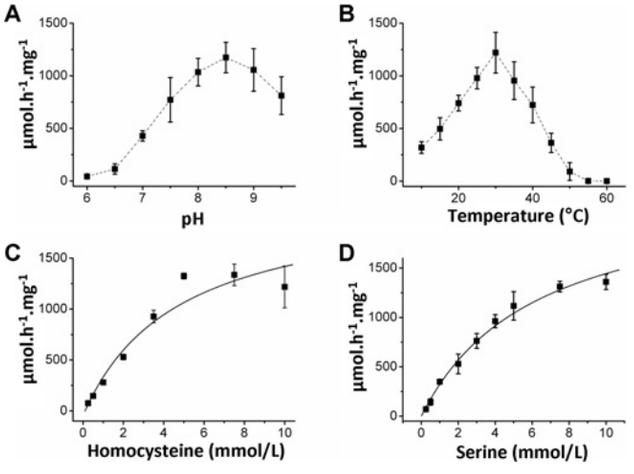


Figure S6 Enzymatic properties of recombinant CBS-1

(**A** and **B**) The dependence of CBS activity on pH and temperature respectively. The kinetic properties of CBS-1 for 1–10 mM homocysteine in a mixture with 10 mM serine are shown in (**C**), whereas the properties for 1–10 mM serine in a mixture with 10 mM homocysteine are shown in (**D**). Results are means ± S.D. from four measurements.

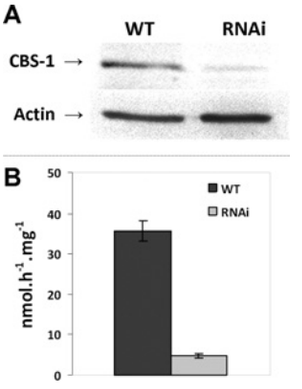


Figure S7 Western blot analysis and CBS assay of crude nematode extracts

(**A**) Western blot analysis showing a decreased level of CBS-1 protein in nematodes after RNAi. Actin was used as a reference protein. (**B**) CBS activity is significantly decreased in nematodes after RNAi. Results are means ± S.D. from two independent measurements.

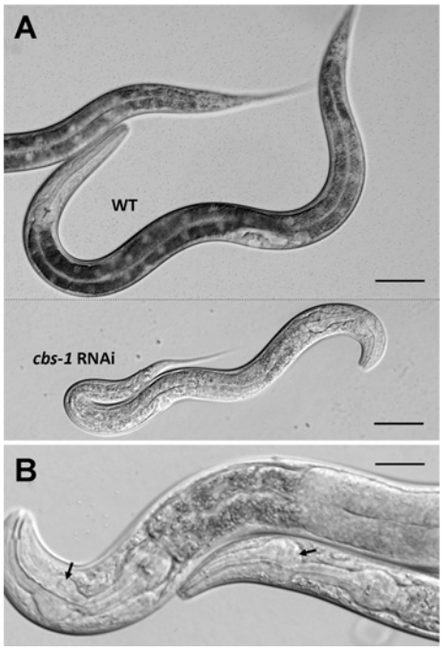


Figure S8 Inhibition of *cbs-1* by RNAi

Images show the body morphologies of worms in Nomarski optics. (**A**) L4 stage of *cbs-1* RNAi and WT worms. The affected worms exhibit decreased body mass and partial lack of pigment granules in the intestine. Scale bars, 50 μm. (**B**) Higher magnification of an affected adult nematode pharynx. The pharynx shows abnormal morphogenesis of the metacorpus with a balloon-like appearance indicated by an arrow. Scale bar, 25 μm.

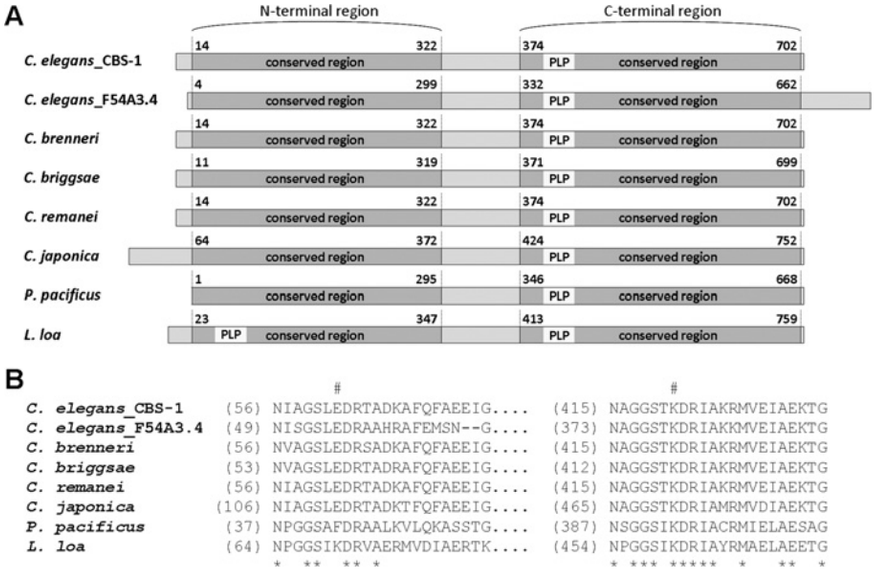


Figure S9 Domain architecture of CBS enzymes in nematodes

(**A**) Domain organization of nematode CBSs. The predicted amino acid sequences of hypothetical nematode CBSs were aligned with the sequence of *C. elegans* CBS-1, and the domain architecture of the proteins was inferred from the degree of homology. Primary structures are aligned by the PLP-binding lysine residue. The numbers indicate the first and the last amino acid residues in each of the conserved domains. Hypothetical proteins included in the alignment are as follows: *C. elegans*_ZC373.1 (CBS-1, UniProt entry Q23264); *C. brenneri* (WormBase accession number CN28558); *C. briggsae* (UniProt entry A8WRM3); *C. remanei* (UniProt entry E3MMP8); *C. japonica* (WormBase accession number JA15528); *C. elegans* F54A3.4 (UniProt entry Q9N4K2); *Pristionchus pacificus* (WormBase accession number PP12619); *Loa loa* (UniProt entry E1FTU4). (**B**) Amino acid alignment of hypothetical PLP-binding site of separated N-terminal and C-terminal conserved regions of various nematode CBSs. #, site of the putative PLP-binding lysine residues. $*$, conserved residue. Only *Loa loa* CBS contains a lysine residue in both PLP-binding sites.

Table S1 Primer sequences

Primers A–F were used to generate the cbs-1–GFP vector (see the GFP reporter assay section of the main text), G and H were used for the amplification of the shortened *cbs-1* coding sequence (see the RNA-mediated interference section of the main text), I–K (sense) and L–O (antisense) were used in combination to amplify the *F54A3.4* open reading frame (see the PCR amplification and sequencing section of the main text), P–U (sense) and V (antisense) were used for the amplification of the *cbs-1* coding sequence with specific cloning overhangs (see the Bacterial expression and purification section of the main text). Primers W–Z were used for site-directed mutagenesis.

Primer	Sequence
A, cbs-1_S	5′-ACTTGACGGAAAAGCTGGCAGA-3′
B, cbs-1_A	5′-AGTCGACCTGCAGGCATGCAAGCTGGCGTCTAGGAAATGACGCTCATTC-3′
C, GFP_S	5′-AGCTTGCATGCCTGCAGGTCG-3′
D, GFP_AS	5′-AAGGGCCCGTACGGCCGACTA-3′
E, cbs-1_S*	5′-GAGGAATGACCATCAATTTGA-3′
F, GFP_AS*	5′-GGAAACAGTTATGTTTGGTATA-3′
G , RNAi_S	5′-GACCCTCATGGATCTATTC-3′
H, RNAi_AS	5′-GACGCTCATTCATCCAATC-3′
I, F54_S1	5′-GACGAATTCATGTGCCTGCCTACCATTAAA-3′
J, F54_S2	5′-GGCAAGACGCCACTGGTGAA-3′
K, F54_S3	5′-AGAAGACAACAGTGGTCGGAGTGAGAT-3′
L, F54_AS1	5′-CAAGCGGCCGCTCAATAGAAAATGCGAGAGCG-3′
M, F54_AS2	5′-AGAGATTCCGGTGATGGTAC-3′
N, F54_AS3	5′-CAACGGCACCAGTTGAGTTG-3′
O, F54_AS4	5′-TGGCTTCCAGCACTGCCGC-3′
P, CBS-1_S	5′-CCTGGGATCCATGATCCAAAACGAAGTTTCC-3′
Q, Δ1–372	5′-CCTGGGATCCCCAGAAAGGCCACTGGTTCTT-3′
R, Δ1–322	5′-CCTGGGATCCGTGGTGACCAGAAAAGATGGA-3′
S, Δ1–299	5′-CCTGGGATCCATGGAATTAGAAATTATC-3′
T, b/375	5′-CCTGGGATCCATGGACCACAACCAAACAGCA-3′
U, b/360	5′-CCTGGGATCCATGATCCAACTAAACTTGCTG-3′
V, CBS-1_AS	5′-GCCGCTCGAGTTAGTGGTGATGGTGATGATGGGCGTCTAGGAAATGACG-3′
W, K421A_S	5′-TGAACGCTGGGGGATCAACAGCGGATCGTATTG-3′
X, K421A_AS	5′-CATTCTTTTGGCAATACGATCCGCTGTTGATCCC-3′
Y, E62K_S	5′-TCAATATTGCGGGATCTTTGAAAGACCGTACCG-3′
Z, E62K_AS	5′-GCTTTGTCAGCGGTACGGTCTTTCAAAGATCCC-3′

Table S2 Homologous level of CBS-related proteins in *C. elegans*

C. elegans genes are arranged by the level of homology with human CBS. The query sequences are as follows: human CBS (UniProt entry P35520), trypanosomal CBS (UniProt entry Q9BH24) and bacterial CS (UniProt entry P0ABK5). In each group of comparisons, the left-hand column lists the e-value, and the middle and right-hand column list identical and positive matches in percentages respectively. The Table shows three groups of ten CBS homologues in *C. elegans*: ZC373.1 and F54A3.4 have the highest homology with CBS; C17G1.7, R08E5.2, F59A7.9 and K10H10.2 are the most homologous with cysteine synthase; and the remaining homologues belong to other unspecified fold-type II PLP-dependent protein families. AA, number of amino acids in the hypothetical protein; COG, clusters of orthologous groups of proteins; CBS RE, CBS and related enzymes; NA, not assigned; SR, serine racemase.

			Homology [e-value, identities (%), positives (%)]								
Name	AA	COG	human CBS (551 AA)			*T. cruzi* CBS (384 AA)			*E. coli* CS (323 AA)		
ZC373.1	704	CBS RE	8e – 94	54	71	2e – 84	50	66	2e – 30	30	44
F54A3.4	755	CBS RE	4e – 92	54	68	8e – 87	51	66	6e – 27	32	46
C17G1.7	341	NA	2e – 54	44	60	4e – 41	36	51	5e – 62	44	59
K10H10.2	337	CBS RE	4e – 54	38	60	9e – 43	35	56	1e – 60	45	58
R08E5.2	337	CBS RE	1e – 51	37	58	2e – 42	36	55	1e – 57	41	57
F59A7.9	337	CBS RE	7e – 42	36	58	3e – 38	35	54	4e – 56	41	55
F01D4.8	430	CBS RE	2e – 10	24	43	0.012	25	49	0.043	26	46
T25D3.3	427	CBS RE	2e – 09	25	40	0.003	24	44	0.001	28	56
T01H8.2	317	SR	1e – 07	28	43	2e – 04	23	39	0.015	29	48
F13B12.4	435	CBS mRE	5e – 07	21	40	0.037 23	47	5e-04	26	45	

Received 15 August 2011/13 December 2011; accepted 13 January 2012
Published as BJ Immediate Publication 13 January 2012, doi:10.1042/BJ20111478

Publication 5.1.4

A. Hnízda et al.

Cystathionine beta-synthase mutants exhibit changes in protein unfolding: conformational analysis of misfolded variants in crude cell extracts

Journal of Inherited Metabolic Disease, On-line article

J Inherit Metab Dis
DOI 10.1007/s10545-011-9407-4

ORIGINAL ARTICLE

Cystathionine beta-synthase mutants exhibit changes in protein unfolding: conformational analysis of misfolded variants in crude cell extracts

Aleš Hnízda · Vojtěch Jurga · Kateřina Raková · Viktor Kožich

Received: 20 July 2011 / Revised: 16 September 2011 / Accepted: 21 September 2011

Abstract Protein misfolding has been proposed to be a common pathogenic mechanism in many inborn errors of metabolism including cystathionine β-synthase (CBS) deficiency. In this work, we describe the structural properties of nine CBS mutants that represent a common molecular pathology in the *CBS* gene. Using thermolysin in two proteolytic techniques, we examined conformation of these mutants directly in crude cell extracts after expression in *E. coli*. Proteolysis with thermolysin under native conditions appeared to be a useful technique even for very unstable mutant proteins, whereas pulse proteolysis in a urea gradient had limited values for the study of the majority of CBS mutants due to their instability. Mutants in the active core had either slightly increased unfolding (p.A114V, p.E302K and p.G307S) or extensive unfolding with decreased stability (p.H65R, p.T191M, p.I278T and p.R369C). The extent of the unfolding inversely correlated with the previously determined degree of tetrameric assembly and with the catalytic activity. In contrast, mutants bearing aminoacid substitutions in the C-terminal regulatory domain (p.R439Q and p.D444N) had increased global stability with decreased flexibility. This study shows that proteolytic techniques can reveal conformational abnormalities even for CBS mutants that have activity and/or a degree of assembly similar to the wild-type enzyme. We present here a methodological strategy that may be used in cell lysates to evaluate properties of proteins that tend to misfold and aggregate and that may be important for conformational studies of disease-causing mutations in the field of inborn errors of metabolism.

Communicated by: Niels Gregersen

Competing interests: None declared.

Electronic supplementary material The online version of this article (doi:10.1007/s10545-011-9407-4) contains supplementary material, which is available to authorized users.

A. Hnízda · K. Raková · V. Kožich (✉)
Institute of Inherited Metabolic Disorders,
First Faculty of Medicine, Charles University in Prague
and General University Hospital in Prague,
Ke Karlovu 2,
128 08, Prague 2, Czech Republic
e-mail: viktor.kozich@LF1.cuni.cz

V. Jurga
Department of Biochemistry and Microbiology,
Institute of Chemical Technology in Prague,
Technická 5,
166 28, Prague 6, Czech Republic

Introduction

Protein misfolding plays a role in the pathogenesis of many diseases, including inborn errors of metabolism, such as phenylketonuria due to phenylalanine hydroxylase deficiency, lysosomal storage disorders and various acyl-CoA deficiencies (Muntau and Gersting 2010). Aberrant folding of mutant proteins may be repaired by treatment of patients with chaperones, and the possible rescue of the mutants represents an interesting option for successful therapy of conformational disorders (Leandro and Gomes 2008). The efficacy and specificity of the search for novel compounds with chaperone activity may be significantly increased if detailed knowledge of the structural mechanisms underlying the misfolding of the mutants is available. This has been demonstrated by the implementation of a novel therapy for phenylketonuria using sapropterin (Muntau et al. 2002), which possibly acts a pharmacological chaperone (Erlandsen et al. 2004; Pey et al. 2004; Perez et al. 2005).

Homocystinuria due to cystathionine beta-synthase (CBS) deficiency is a common enzymopathy that affects

Published online: 09 November 2011

the metabolism of sulfur amino acids and manifests clinically as a multisystemic disease (Mudd et al. 2001; Skovby et al. 2010). It has been proposed that many missense mutations in the *CBS* gene disrupt correct protein folding. This notion was supported by several lines of evidence, including the abnormal migration of mutant proteins in native electrophoresis (Janosik et al. 2001b; Kozich et al. 2010), the decreased solubility of mutants in SDS (Kozich et al. 2010) and the increased rate of proteasome-dependent proteolysis (Singh and Kruger 2009). The misfolding of CBS mutants was also inferred from the rescue of mutant enzymes by chemical chaperones (Singh et al. 2007; Majtan et al. 2010; Kopecka et al. 2011), CBS ligands (Kopecka et al. 2011), proteasome inhibitors (Singh et al. 2010), or by expression under temperature permissive to folding (Kozich et al. 2010). Taken together, all these experiments strongly indicate that misfolding plays an important role in CBS deficiency; however, the conformational changes underlying the abnormal folding have not been studied in detail.

Biophysical techniques used for conformational studies require relatively large amounts of purified proteins. However, only a few CBS mutants have been successfully produced as recombinant proteins in *E. coli*, purified to homogeneity in sufficient yields and structurally characterized, including the C-terminal mutants p.I435T, p.S466L (Janosik et al. 2001a) and p.D444N (Evande et al. 2002; Sen et al. 2005) and a double-linked mutant p.P78R/K102N (Sen and Banerjee 2007). Six additional CBS mutants (p.P49L, p.P78R, p.A114V, p.R125Q, p.E176K, p.P422L) were successfully purified in the presence of chemical chaperones in the culture media (Majtan et al. 2010). Because chaperones facilitate correct folding, it is not surprising that the chaperoned mutants did not exhibit gross structural abnormalities. It should be also noted that during the purification procedure, an unknown portion of misfolded and aggregated proteins may be lost and, consequently, the structural information about this fraction may not be obtained from studying purified proteins.

To study mutants that are not amenable to purification, proteolytic techniques in crude cell extracts may be used. It has been demonstrated that protein folding may be studied quantitatively by employing thermolysin, even in crude cell extracts (Chang and Park 2009; Kim et al. 2009). The methodology comprises two techniques: proteolysis under native conditions and pulse proteolysis in a urea gradient. The rate of proteolysis under native conditions reveals the extent of unfolding in the conformational ensemble of studied proteins because proteases can only cleave flexible regions and partially or globally unfolded structures (Young et al. 2007). Pulse proteolysis is a technique that uses thermolysin to monitore urea-induced unfolding. After a short proteolytic pulse, the fraction of folded proteins remains uncleaved and the unfolded species are digested. Using this premise, the c_m value, a measure of protein thermodynamic stability in the urea gradient, may be determined (Park and Marqusee 2005).

In the present work, we applied these recently developed techniques, whose feasibility for the conformational study of CBS was shown using purified wild-type proteins (Prudova et al. 2006; Hnizda et al. 2010), to study CBS mutants expressed in *E. coli*. For these analyses, we selected the three most prevalent mutants and six additional mutants at different locations within the CBS molecule, and we examined if the conformation of the CBS mutant differs from the wild-type enzyme.

Results

Selection of CBS mutants The mutations selected for this study are located in different domains of the CBS molecule (Meier et al. 2003; Yamanishi et al. 2006; Hnizda et al. 2010) (see Table 1). Advanced molecular modelling of the impact of mutations on enzyme structure could not be performed with sufficient reliability because the 3-D structure of human full-length wild-type CBS has not been experimentally solved at atomic resolution. As reported previously, these mutants vary in the degree of correct assembly and subsequent tetramer formation as well as in the levels of their catalytic activity (Kozich et al. 2010).

Determination of proteolytic kinetics under native conditions We optimized the previously described procedure (Chang and Park 2009) and final experimental conditions designed to evaluate the proteolytic cleavage of wild-type CBS in crude extracts over a time course of 20 min (for details, see Methods). The analysis of proteolytic cleavage using SDS-PAGE (Fig. 1) and native electrophoresis (Fig. S1 available in Supplementary_Info.pdf) did not reveal the presence of any partially cleaved fragments or unfolded intermediates, respectively. Therefore, the assumption for a model to describe protein unfolding and consequent proteolysis of CBS by thermolysin in which there is an equilibrium between folded tetramers and unfolded monomers (Park and Marqusee 2004a) was satisfied. Because the opening of the folded structure is necessary for sufficient proteolysis (Park and Marqusee 2004b; Chang and Park 2009), the determination of proteolytic kinetics provides data on the extent of the unfolding of the conformational ensemble.

The quantitative analysis of proteolysis under native conditions revealed that the CBS protein in crude cell extract comprises two different fractions. The major fraction of wild type CBS (≈80%) was cleaved via the first-order kinetics while the minor portion (≈20%)

Table 1 List of CBS mutants analyzed in this study.

Protein	Location of the mutated residue	The number of tetramers (% of wild-type)	Relative enzyme activity (% of wild-type activity)	Reference for the location of the mutation
p.H65R	heme-binding pocket	23.1	3.9	(Meier et al. 2003)
p.A114V	dimer interface of the active core	73.7	76.9	(Meier et al. 2003)
p.T191M	solvent-exposed at the periphery of the active core	1.7	0.3	(Kozich et al. 2010)
p.I278T	buried in the active core	1.0	0.3	(Meier et al. 2003)
p.E302K	regulatory interface	90.7	95.4	(Meier et al. 2003; Hnizda et al. 2010)
p.G307S	catalytic site	111.7	0.2	(Meier et al. 2003)
p.R369C	regulatory interface	7.8	1.8	(Yamanishi et al. 2006)
p.R439Q	regulatory domain	117.6	117.2	(Kozich et al. 2010)
p.D444N	regulatory domain - AdoMet binding site	120.2	163.8	(Hnizda et al. 2010; Kozich et al. 2010)

Relative enzyme activities and the number of tetramers are expressed as a mean value from data published elsewhere (Kozich et al. 2010).

remained uncleaved at the end of proteolytic experiment. The data from the proteolytic kinetics of wild-type CBS were consistent with Eq. 1 (see Methods; r^2=0.9066; see Fig. S2A available in Supplementary_Info.pdf) and they show that the proteolysis with thermolysin under native conditions provides two parameters about the CBS proteins: the proteolytic rate of the major fraction and the relative abundance of the fraction that resists proteolysis during experiments.

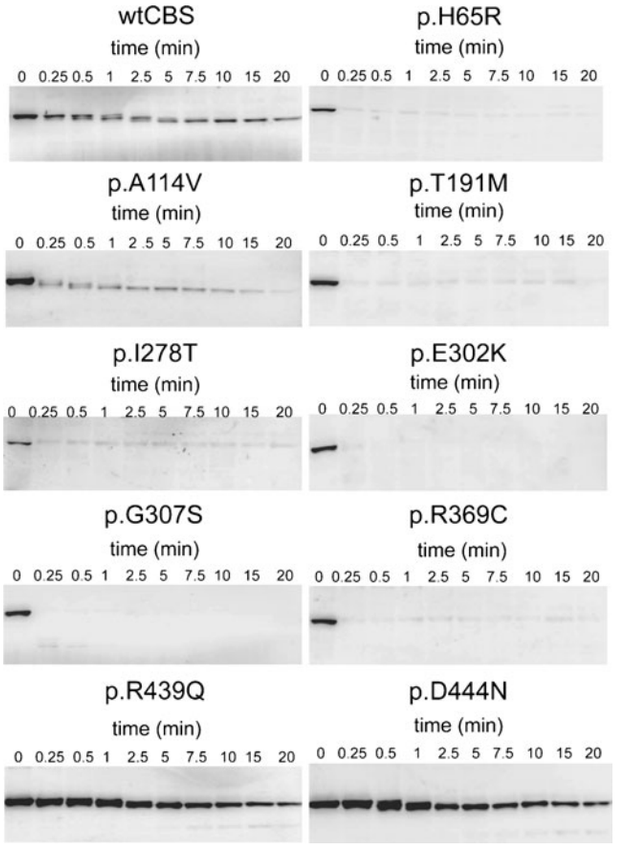


Fig. 1 Proteolytic kinetics under native conditions for wild-type and mutant CBS proteins. Cell lysates (1 mg/ml protein) were cleaved by thermolysin (0.2 mg/ml). The amount of uncleaved protein was determined by Western blot followed by immunodetection with CBS antibody. Time points are depicted in minutes

Analysis of CBS mutants by native proteolysis The kinetics of proteolysis under native conditions was successfully determined for all mutants. We show here that this method is robust and that it can be used for the analysis of even extremely unstable proteins.

All mutant proteins possessing an aminoacid substitution in the active core were cleaved too rapidly using the same default conditions for the capture of their proteolytic courses (Fig. 1). Therefore, the unstable mutants were exposed to lower and varying concentrations of thermolysin, which enabled the observation of the proteolytic time course and subsequent determination of its kinetics (Fig. S3, available in Supporting_Info.pdf). Because the dependence of the kinetic constant k_{obs} on the thermolysin concentration appeared to be linear (as shown for wild-type CBS in Fig. S2B, available in Supplementary_Info.pdf), the proteolytic susceptibility of CBS mutants was assessed by comparing the k_p values expressed in M^{-1} min^{-1}; that is, k_{obs} divided by the molar concentration of thermolysin in the proteolytic reaction (Park and Marqusee 2004b).

A significant increase in the proteolytic rates was observed for all mutants located in the active core, Fig. 1 illustrates the proteolysis of the mutants under the same conditions, and Table 2 provides the quantitative comparisons of the kinetic constant k_p for the analyzed proteins. Unfortunately, faster cleavage of the active core mutant p.I278T could not be reliably quantified due to poor reproducibility of the proteolytic cleavage of this mutant

at lower thermolysin concentration. In contrast, a decreased value of k_p was observed for mutations located in the regulatory domain, suggesting a rigidification of their conformation.

In the next step, we compared the *B* values, which give the relative abundance of proteolytically resistant fractions of CBS proteins revealed from the determination of proteolytic kinetics. It should be noted that cleavage with lower concentrations of thermolysin might lead to an overestimation of the uncleaved protein fraction, as shown for wild-type CBS in Fig. S2C (available in the Supplementary_Info.pdf). However, an increased *B* value was not observed for any of the mutants analyzed in this study (Table 2). In contrast to wild-type CBS, several mutants (p.H65R, p.I278T, p.E302K and p.G307S) exhibited a decreased *B* value an increased k_p. These results indicate that the conformer distribution in the structural ensemble of these mutants may possibly be shifted towards the unfolded intermediates.

In summary, proteolysis under native conditions revealed that all mutants had a different degree of structural flexibility. Mutations in the active core had an increased extent of protein unfolding, whereas mutations in the regulatory domain tended to form more rigid protein structures.

Table 2 Proteolytic kinetics of the CBS mutants under native conditions and pulse proteolysis in a urea gradient

	Proteolysis under native conditions		Pulse proteolysis
Protein	k_p (10^6. M^{-1}.min^{-1})	B (%)	C_m [M]
Wild-type	0.30±0.04	24.4±1.7	3.05±0.03
p.H65R	38.4±5.1	8.2±2.1	N.D
p.A114V	2.2±0.4	24.1±2.5	2.84±0.04*
p.T191M	44.3±8.8	22.9±3.0	2.4**
p.I278T	<2, 64.5>***	19.5±1.2	2.42±0.01
p.E302K	5.0±0.6	0±3.3	N.D
p.G307S	4.6±0.4	0±0.1	N.D
p.R369C	24.0±3.5	23.4±1.9	2.59±0.09
p.R439Q	0.02±0.01	25.7±9.6	3.49±0.04
p.D444N	0.03±0.01	28.4±6.2	3.32±0.04

k_p, *B* and c_m values were determined by nonlinear data fitting from four different measurements. Representative gels with corresponding plots are shown in the Supplementary materials.

*The determination of c_m for p.A114V neglected the cleavage of partially folded forms.

**Values were determined by visual inspection because these data could not be evaluated by nonlinear data fitting.

***k_p could not be exactly determined due to low reproducibility at lower concentrations of thermolysin.

N.D. not determined due to the rapid proteolysis in the absence of urea.

Pulse proteolysis A previously developed procedure for pulse proteolysis of purified CBS (Hnizda et al. 2010) was successfully adopted for crude cell extracts and the data for wild-type CBS followed the two-state unfolding model (r^2=0.9562; Fig. S2D available in Supplementary_Info.pdf). Pulse proteolysis therefore describes an equilibrium between folded tetramers and unfolded monomers in a urea gradient; the validity of this model is further supported by a previous study (Kery et al. 1998) reporting that higher concentrations of urea lead to the dissociation of CBS tetramers.

Because a proportion of folded wild-type CBS was cleaved during the proteolytic pulse even in the absence of urea (see proteolysis under native conditions), a possible systematic error in the determination of c_m was assessed according to a previously described procedure (Park and Marqusee 2005, 2006). The kinetic constant for proteolysis in 2.5 M urea was determined to be 0.13±0.01 min^{-1} (see Fig. S4, available in Supplementary_Info.pdf), indicating negligible cleavage of folded protein during the proteolytic pulse at the beginning of the transition zone and, consequently, the absence of a systematic error (Park and Marqusee 2006).

It should also be noted that the c_m calculated from the pulse proteolysis of multimeric protein does not simply represent global protein stability ($\Delta G_{unf}°$) and that the structural stability of the protein is dependent on its concentration (Park and Marqusee 2004a). However, the amount of CBS mutants in cell extracts was similar to or slightly less than the wild-type protein (Fig. S5, available in Supplementary_Info.pdf). A significant decrease in protein level to ~20% of wild-type enzyme was observed only for p.I278T. Although the decreased amount of p.I278T might lead to an underestimation of its stability, the experiments with a decreased amount of wild-type CBS (described in Materials and methods) did not reveal any changes in the apparent c_m values. It was also proposed that, in spite of some limitations, c_m values are still valuable for comparing protein stability under the same conditions (Schlebach et al. 2010). Consequently, we employed pulse proteolysis in a urea gradient for conformational studies of CBS mutants in crude cell extracts.

Analysis of CBS mutants using pulse proteolysis in a urea gradient Pulse proteolysis in a urea gradient was used for six mutant proteins, whereas three mutants (p.H65R, p.E302K and p.G307S) were cleaved rapidly during the proteolytic pulse in the absence of urea, and less than 10% of the CBS protein remained uncleaved. It has been shown that pulse proteolysis in a urea gradient cannot be used for rapidly cleaved proteins (Chang and Park 2009).

Analysis of the mutantions in the regulatory-domain revealed higher c_m values as a measure of increased protein

stability (Table 2, Figs. S4 and S6, available in Supplementary_Info.pdf). However, we experienced several difficulties during the evaluation of mutants with an increased rate of proteolysis observed under native conditions (p.A114V, p.T191M, p.I278T and p.R369C). In the case of the slightly unfolded mutant p.A114V, we observed a significant proteolytic susceptibility at the beginning of the apparent transition zone (Fig. S4 in Supplemetary_Info.pdf), and the requirement for the kinetic constant to be less than 0.2 min^{-1} for an accurate determination of c_m was therefore not satisfied. Despite this, pulse proteolysis can yield at least partial knowledge about the urea-induced transition between the partially and globally unfolded fractions. An analysis accounting for a systematic error (see Fig. S7, available in Supplementary_Info.pdf) showed that the correct value of c_m could in fact be lower that the apparent value. Taken together, both the corrected and uncorrected value of c_m indicated that the protein stability of the mutant p.A114V was at least slightly decreased compared to wild-type CBS.

Another type of obstacle arose during the pulse proteolysis of severely unfolded proteins, p.T191M, p.I278T and p.R369C. The amount of CBS mutants in cell lysates decreased after the overnight incubation even in the absence of urea in the uncleaved control samples, perhaps due to the formation of precipitates that did not enter into the electrophoretic gel (a representative gel together with a quantitative comparison of the amount of soluble CBS antigen before and after overnight incubation is shown in Fig. S8, available in Supplementary_Info.pdf).Therefore, pulse proteolysis could only be assessed for the soluble fraction of these mutants. Interestingly, the proteins were not significantly cleaved at the beginning of their apparent transition zone, indicating the reliable determination of these values (Fig. S4 in Supplementary_Info.pdf). The analyses showed a significant decrease in the stability of these mutant proteins, which is consistent with the higher extent of unfolding revealed by native proteolysis. However, these results should be interpreted with great care because we analyzed only the limited fraction of the conformational ensemble that did not precipitate.

It should be noted here that the reported difficulties in pulse proteolysis are most likely caused by the instability of the mutants. Various modifications of experimental procedures, such as a lower amount of thermolysin in the assay or a shorter equilibration time of cell lysates in urea, did not improve the outcome of the experiments (data not shown). This suggests that pulse proteolysis in a urea gradient for the conformational analysis of highly unstable mutant proteins may be of limited value. In spite of these limitations, pulse proteolysis showed that the different extent of the unfolding of CBS mutants is associated with corresponding changes in protein stability: unfolded mutant proteins had lower stability, whereas rigidified mutants were more stable than wild-type CBS.

Correlation of protein unfolding with the degree of tetramer assembly and with enzyme activity of CBS mutants The data from proteolysis under native conditions were correlated with the number of correctly assembled tetramers and their catalytic activity, which was reported in a previous study (Kozich et al. 2010). The relative positions of the mutations are similar in both plots, with the exception of the catalytic site mutant p.G307S (compare Fig. 2a and b). The plots show that hypoactive and incorrectly assembled mutants had a higher extent of protein unfolding, whereas hyperactive and properly assembled mutants were more rigidified that wild-type CBS. This analysis indicates a trend between the proteolytic resistance, k_p, and the number of tetramers (R^2=0.6646; p=0.0074) and enzyme activity (R^2=0.7566; p=0.0015) (Fig. 2). These data may imply that an optimal degree of conformational flexibility is a

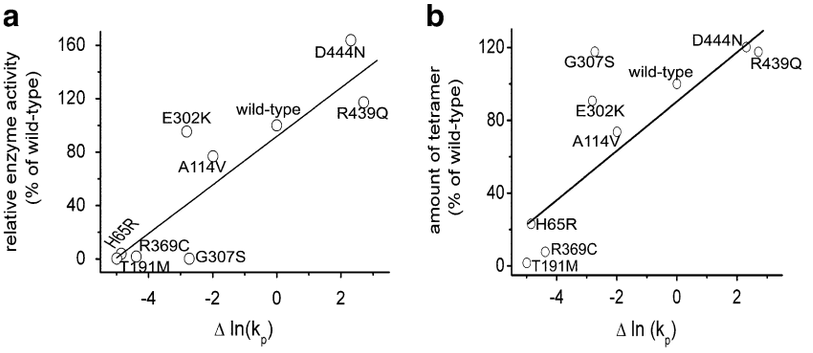


Fig. 2 Conformational properties of CBS mutants: proteolytic data correlated with enzyme activity and with the degree of tetrametric assembly. The dependence of the proteolytic kinetics with thermolysin under native conditions (ln (k_p)) on the enzyme activity of the CBS mutants (A) and on the number of assembled tetramers (B) – the correlation analysis indicates a possible trend between the proteolytic resistance and the number of tetramers (R^2=0.6646; p=0.0074) and enzyme activity (R^2=0.7566; p=0.0015). All values for the mutants are normalized to the number of tetramers or activity of the wild-type as appropriate

necessary condition for correct tetrameric assembly which further determines the catalytic efficiency of CBS mutants.

Discussion

In this study, we describe conformational properties of nine disease-causing CBS mutants that previously resisted purification and further characterization. By using proteolysis of mutants in crude cell extracts under native conditions and pulse proteolysis in a urea gradient, we demonstrated that the conformational properties of all analyzed mutants differ from the wild-type protein. Based on these data, we propose a possible molecular mechanism for the misfolding and misassembly of CBS mutants. Using a modified mathematical model for the folded tetramers-unfolded monomers equilibrium (Park and Marqusee 2004a), mutants may be divided into three major groups:

i) Severely misassembled mutants with low residual activity and extensive unfolding (p.H65R, p.T191M, p.I278T and p.R369C). These mutants predominantly form aggregates that are likely composed of weakly interacting subunits with unfolded regions of high sensitivity to proteolytic cleavage.
ii) Correctly assembled mutants that are only slightly unfolded (catalytically active p.A114V and p.E302K and the inactive p.G307S located in the active site). These proteins are characterized by increased unfolding through kinetically and/or thermodynamically more favourable formation of unfolded intermediates.
iii) Correctly assembled hyperactive mutants that are more stable and more tightly folded (p.R439Q and p.D444N) due to a decreased formation of unfolded species.

Altered conformational properties of CBS mutants may affect protein turnover in vivo The higher proteolytic susceptibility of the majority of CBS mutants in this study indicates their inclination to unfolding. This observation is consistent with previous work by Singh and Kruger demonstrating that the most common mutant, p.I278T, exhibited accelerated ubiquitin-dependent proteolysis in *S. cerevisiae* and that turnover of the mutated protein could have been altered by a manipulation of the levels of heat shock proteins (Singh and Kruger 2009). Furthermore, several mutants were rescued by treatment with the proteasome inhibitor bortezomib in yeast or by MG132 in patient-derived fibroblasts (Singh et al. 2010), indicating that more rapid degradation plays an important role in the pathophysiology of CBS deficiency.

All these data are consistent with a hypothesis that the unfolded intermediates of CBS mutants may be rapidly degraded by proteasome-dependent proteolysis. Increased protein unfolding was reported for oligomeric enzymes that are deficient in other inherited metabolic diseases such as S-adenosylhomocysteine hydrolase (Beluzic et al. 2006), medium chain acyl-CoA dehydrogenase (Maier et al. 2009) and electron transfer flavoprotein (Schiff et al. 2006). Furthermore, the relationship between increased proteolytic susceptibility in vitro and accelerated proteasome-dependent turnover in vivo was also demonstrated for misfolded mutants of alanine:glyoxylate aminotransferase (Coulter-Mackie and Lian 2006, 2008) and phenylalanine hydroxylase (Waters et al. 1999; Gersting et al. 2008).

Conversely, mutations in the regulatory domain of CBS led to an increase in proteolytic stability. This indicates that the conformation of these mutants is less flexible than of the wild-type protein. It remains to be addressed how the increased conformational rigidity and hyperactivity of the C-terminal mutants could result in their pathogenicity in vivo. In previous studies, the representative C-terminal mutants, p.I435T, p.D444N and p.S466L, had higher activity than the wild-type enzyme together and a lower affinity towards an allosteric activator of CBS, *S*-adenosylmethionine (AdoMet), and they failed to change conformation upon AdoMet binding (Janosik et al. 2001a; Evande et al. 2002; Sen et al. 2005). Based on these observations, it was proposed that the altered response to AdoMet may play a role in the pathogenicity of these C-terminal mutations (Yamanishi et al. 2006; Kozich et al. 2010). Interestingly, other work showed that the expression of the C-terminal mutant p.S466L in a transgenic mouse led to a lower steady-state protein level and to decreased catalytic efficiency in vivo (Gupta et al. 2008). Decreased amounts of hyperactive mutants were also observed in crude cell extracts for two patient-derived fibroblast cell lines carrying the mutations p.I435T/del ex8 (Maclean et al. 2002) or p.D444N/p.D444N (Evande et al. 2002). Based on these observations, we speculate that the more tightly folded mutants may be recognized by the cellular quality control machinery, thereby leading to their accelerated degradation.

It should be noted that our study revealed several limitations of proteolytic experiments for conformational studies in crude cell extracts. Pulse proteolysis could not be applied for three of the nine mutant proteins (p.H65R, p.E302K and p.G307S) because these mutants were substantially cleaved during the proteolytic pulse, even in the absence of urea. In addition, we experienced difficulties with the evaluation of the other unfolded mutants (p.A114V, p.T191M, p.I278T and p.R369C) due to significant cleavage at the beginning of the apparent transition zone or precipitation events during the overnight equilibration in the urea gradient. In this case, the determination of c_m values and subsequent data interpretation must be performed carefully and pulse proteolysis in the urea gradient

may indicate changes in thermodynamic stability of unfolded mutant proteins.

On the other hand, the determination of the rate of proteolysis with thermolysin under native conditions appeared to be robust and provided conformational data about all mutants studied. Despite this, there were some obstacles while performing this technique: a) the relatively demanding optimalization of the procedure in order to find an appropriate ratio of the total protein in lysates to the thermolysin concentration (w/w) to accurately determine proteolysis kinetics; b) the presence of a small protein fraction that was not cleaved according to first-order kinetics and that remained intact and catalytically active at the end of experiments (data not shown). Because the small uncleaved fraction was also observed in mammalian HepG and HeK cell-lines with endogenous CBS expression (our unpublished data) but not with purified CBS proteins (Hnizda et al. 2010), it is possible that the minor fraction of CBS in cell lysates may result from intermolecular interactions with small ligands or high molecular weight molecules such as other proteins or nucleic acids. This uncleaved fraction is most likely not caused by binding of a small compound because the dialysis of the bacterial lysate containing wild-type CBS prior to the assay did not significantly change the proteolytic time-course (data not shown). In spite of these limitations, the proteolytic techniques used in crude cell extracts enabled the expansion of our knowledge about the conformational properties of the CBS mutants that are not amenable for purification and consequently not available for detailed studies by more conventional methods.

We present here a general methodological strategy that may be used directly in crude cell extracts to evaluate properties of proteins that tend to misfold and aggregate and that may be important for conformational studies of disease-causing mutations in the field of inborn errors of metabolism. This approach may benefit the conformational analysis of mutants after expression without affinity tags or after endogenous expression in patient-derived fibroblasts, other available cell lines or tissues. Proteolytic techniques are able to reveal conformational abnormalities even in mutant enzymes with normal catalytic activity and may be applied to the mutational analysis of non-enzymatic proteins whose functional assays are lacking or are difficult to employ.

Materials and methods

Materials If not specified otherwise, all chemicals were purchased from Sigma-Aldrich.

Expression of CBS enzymes and preparation of crude cell extract The CBS proteins were expressed at 37°C in *E. coli* transformed with pHCS3 plasmids that express untagged full-length wild-type or mutant CBS proteins (Kozich and Kraus 1992). Crude cell-extracts were prepared as described previously (Kozich et al. 2010). The protein concentration was determined using Bradford reagent with bovine serum albumin as a standard.

Proteolysis with thermolysin under native conditions Crude cell extracts (final protein concentration 1 mg/ml; total volume 25 μl) were diluted in 20 mM Tris-HCl (pH 8.0) containing 10 mM $CaCl_2$, equilibrated at 25°C for 10 minutes and then digested with thermolysin. The concentration of thermolysin ranged from 0.4 to 200 μg/ml as described in the Results. At the selected time points, proteolysis was quenched in 20 mM EDTA.

Pulse proteolysis Pulse proteolysis was performed according to a previously reported procedure (Hnizda et al. 2010) with some modifications. Cell lysates (final protein concentration 1 mg/ml; total volume 25 μl) were incubated overnight at 4°C in 20 mM Tris-HCl (pH 8.0) containing 10 mM $CaCl_2$ and urea (0–7 M). After subsequent equilibration at 25°C for 10 minutes, the samples were digested by thermolysin from *Bacillus thermoproteolyticus* (0.2 mg/ml) at 25°C. The proteolytic pulse (60 s) was quenched in 20 mM EDTA.

Because the thermodynamic stability of multimers may be dependent on protein concentration, we assessed the effect of different amounts of wild-type CBS in the assay on the determined c_m values. Pulse proteolysis of the wild-type CBS was performed for lower concentration of total protein (0.5 or 0.25 mg/ml) that was digested by the corresponding amount of thermolysin (100 or 50 μg/ml). The lower protein concentration in pulse proteolysis did not change the apparent c_m values (data not shown). These experiments showed that lower amounts of CBS proteins in the assay did not have an effect on the accurate determination of their thermodynamic stability implying that this approach is also useful for mutant proteins with decreased steady-state levels.

Western blotting and data fitting The amount of uncleaved CBS was analyzed under denaturing or native conditions by the Western blot procedure described previously (Kopecka et al. 2011). The CBS antigen was detected using purified chicken anti-hCBS serum H19 (HenA, Czech Republic) diluted 1:500 in 3% non-fat dry milk in 10 mM Tris-HCl (pH 7.4) containing 150 mM NaCl. Secondary rabbit anti-chicken horseradish-peroxidase-cojugated antibody (ThermoFischer Scientific) was diluted 1:15,000 in PBS containing 3% non-fat dry milk. The signal was quantified, and the data were fitted to non-linear curves as reported previously (Hnizda et al. 2010; Kozich et al. 2010). Each experiment was carried out twice, each in duplicate; data

shown in figures and tables are therefore the means and standard deviations of four determinations.

Proteolysis under native conditions was evaluated using the first-order kinetics (Park and Marqusee 2004b). Data were fitted into the equation as follows:

$$N = N_0 \bullet e^{(-k_{obs} \bullet t)} + B, \quad (1)$$

where N and N_0 are the amounts of intact protein, k_{obs} is the first-rate kinetics constant, t is time and B is the fraction of protein that was not cleaved via the first-order kinetics and that remained uncleaved at the end of proteolytic experiment. Data from the pulse proteolysis were fitted into the sigmoidal equation as follows:

$$f_{fold} = \frac{1}{1 + e^{p \cdot (c_m - c)}}, \quad (2)$$

where f_{fold} is a fraction of folded proteins remaining intact after the proteolytic pulse and is related to the amount of uncleaved protein after the pulse in absence of urea, c_m is the urea concentration at which f_{fold} is 0.5 and c is the urea concentration. p is a slope of the curve at c_m and reflects unfolding cooperativity.

Acknowledgements We thank Dr. Jakub Krijt for determination of CBS activity in the proteolytically resistant fraction.

Funding This work was supported by the Welcome Trust International Senior Research Fellowship in Biomedical Science in Central Europe (reg. No 070255/Z/03/Z), by the grant No. 7310 from Grant Agency of Charles University in Prague and by the grant SVV262502 from the Charles University in Prague. Institutional support was provided by the Research Project from the Ministry of Education of the Czech Republic No. MSM0021620806.

Authors confirm independence from the sponsors; the content of the article has not been influenced by the sponsors.

References

Beluzic R, Cuk M, Pavkov T et al (2006) A single mutation at Tyr143 of human S-adenosylhomocysteine hydrolase renders the enzyme thermosensitive and affects the oxidation state of bound cofactor nicotinamide-adenine dinucleotide. Biochem J 400:245–253

Chang Y, Park C (2009) Mapping transient partial unfolding by protein engineering and native-state proteolysis. J Mol Biol 393: 543–556

Coulter-Mackie MB, Lian Q (2006) Consequences of missense mutations for dimerization and turnover of alanine:glyoxylate aminotransferase: study of a spectrum of mutations. Mol Genet Metab 89:349–359

Coulter-Mackie MB, Lian Q (2008) Partial trypsin digestion as an indicator of mis-folding of mutant alanine:glyoxylate aminotransferase and chaperone effects of specific ligands. Study of a spectrum of missense mutants. Mol Genet Metab 94:368–374

Erlandsen H, Pey AL, Gamez A et al (2004) Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. Proc Natl Acad Sci USA 101:16903–16908

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 41:11832–11837

Gersting SW, Kemter KF, Staudigl M et al (2008) Loss of function in phenylketonuria is caused by impaired molecular motions and conformational instability. Am J Hum Genet 83:5–17

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 29:1048–1054

Hnizda A, Spiwok V, Jurga V, Kozich V, Kodicek M, Kraus JP (2010) Cross-talk between the catalytic core and the regulatory domain in cystathionine beta-synthase: Study by differential covalent labeling and computational modelling. Biochemistry 49:10526–10534

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 40:10625–10633

Janosik M, Oliveriusova J, Janosikova B et al (2001b) Impaired heme binding and aggregation of mutant cystathionine beta-synthase subunits in homocystinuria. Am J Hum Genet 68:1506–1513

Kery V, Poneleit L, Kraus JP (1998) Trypsin cleavage of human cystathionine beta-synthase into an evolutionarily conserved active core: structural and functional consequences. Arch Biochem Biophys 355:222–232

Kim MS, Song J, Park C (2009) Determining protein stability in cell lysates by pulse proteolysis and Western blotting. Protein Sci 18:1051–1059

Kopecka J, Krijt J, Rakova K, Kozich V (2011) Restoring assembly and activity of cystathionine beta-synthase mutants by ligands and chemical chaperones. J Inherit Metab Dis 34:39–48

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 1:113–123

Kozich V, Sokolova J, Klatovska V et al (2010) Cystathionine beta-synthase mutations: effect of mutation topology on folding and activity. Hum Mutat 31:809–819

Leandro P, Gomes CM (2008) Protein misfolding in conformational disorders: rescue of folding defects and chemical chaperoning. Mini Rev Med Chem 8:901–911

Maclean KN, Gaustadnes M, Oliveriusova J et al (2002) High homocysteine and thrombosis without connective tissue disorders are associated with a novel class of cystathionine beta-synthase (CBS) mutations. Hum Mutat 19:641–655

Maier EM, Gersting SW, Kemter KF et al (2009) Protein misfolding is the molecular mechanism underlying MCADD identified in newborn screening. Hum Mol Genet 18:1612–1623

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 285:15866–15873

Meier M, Oliveriusova J, Kraus JP, Burkhard P (2003) Structural insights into mutations of cystathionine beta-synthase. Biochim Biophys Acta 1647:206–213

Mudd SH, Levy LH, Kraus JP (2001) Disorders of transsulfuration. In: Scriver CR, Beaudet AL, Sly WS, Vale D, Childs B, Vogelstein B

(eds) The metabolic and molecular bases of inherited disease, 7th edn. McGraw-Hill, 1007–1056

Muntau AC, Gersting SW (2010) Phenylketonuria as a model for protein misfolding diseases and for the development of next generation orphan drugs for patients with inborn errors of metabolism. J Inherit Metab Dis 33:649–658

Muntau AC, Roschinger W, Habich M et al (2002) Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria. N Engl J Med 347:2122–2132

Park C, Marqusee S (2004a) Analysis of the stability of multimeric proteins by effective DeltaG and effective m-values. Protein Sci 13:2553–2558

Park C, Marqusee S (2004b) Probing the high energy states in proteins by proteolysis. J Mol Biol 343:1467–1476

Park C, Marqusee S (2005) Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. Nat Methods 2:207–212

Park C, Marqusee S (2006) Quantitative determination of protein stability and ligand binding by pulse proteolysis. In: Coligan JE, Dunn BM, Speicher DW, Wingfield PT, Ploegh HL (eds) Current protocols in protein sciences. Wiley, Hoboken, pp 20.11.1–20.11.14

Perez B, Desviat LR, Gomez-Puertas P, Martinez A, Stevens RC, Ugarte M (2005) Kinetic and stability analysis of PKU mutations identified in BH4-responsive patients. Mol Genet Metab 86 (Suppl 1):S11–S16

Pey AL, Perez B, Desviat LR et al (2004) Mechanisms underlying responsiveness to tetrahydrobiopterin in mild phenylketonuria mutations. Hum Mutat 24:388–399

Prudova A, Bauman Z, Braun A, Vitvitsky V, Lu SC, Banerjee R (2006) S-adenosylmethionine stabilizes cystathionine beta-synthase and modulates redox capacity. Proc Natl Acad Sci USA 103:6489–6494

Schiff M, Froissart R, Olsen RK, Acquaviva C, Vianey-Saban C (2006) Electron transfer flavoprotein deficiency: functional and molecular aspects. Mol Genet Metab 88:153–158

Schlebach JP, Kim MS, Joh NH, Bowie JU, Park C (2010) Probing membrane protein unfolding with pulse proteolysis. J Mol Biol 406:545–551

Sen S, Banerjee R (2007) A pathogenic linked mutation in the catalytic core of human cystathionine beta-synthase disrupts allosteric regulation and allows kinetic characterization of a full-length dimer. Biochemistry 46:4110–4116

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 44:14210–14216

Singh LR, Kruger WD (2009) Functional rescue of mutant human cystathionine beta-synthase by manipulation of Hsp26 and Hsp70 levels in Saccharomyces cerevisiae. J Biol Chem 284:4238–4245

Singh LR, Chen X, Kozich V, Kruger WD (2007) Chemical chaperone rescue of mutant human cystathionine beta-synthase. Mol Genet Metab 91:335–342

Singh LR, Gupta S, Honig NH, Kraus JP, Kruger WD (2010) Activation of mutant enzyme function in vivo by proteasome inhibitors and treatments that induce Hsp70. PLoS Genet 6:e1000807

Skovby F, Gaustadnes M, Mudd SH (2010) A revisit to the natural history of homocystinuria due to cystathionine beta-synthase deficiency. Mol Genet Metab 99:1–3

Waters PJ, Parniak MA, Akerman BR, Jones AO, Scriver CR (1999) Missense mutations in the phenylalanine hydroxylase gene (PAH) can cause accelerated proteolytic turnover of PAH enzyme: a mechanism underlying phenylketonuria. J Inherit Metab Dis 22:208–212

Yamanishi M, Kabil O, Sen S, Banerjee R (2006) Structural insights into pathogenic mutations in heme-dependent cystathionine-beta-synthase. J Inorg Biochem 100:1988–1995

Young TA, Skordalakes E, Marqusee S (2007) Comparison of proteolytic susceptibility in phosphoglycerate kinases from yeast and E. coli: modulation of conformational ensembles without altering structure or stability. J Mol Biol 368:1438–1447

SUPPLEMENTARY INFORMATION

A. Hnízda et al.: Cystathionine beta-synthase mutants exhibit changes in protein unfolding: conformational analysis of misfolded variants in crude cell extracts

SUPPLEMENTARY INFORMATION

Figure S1: Proteolysis of wtCBS under native conditions: analysis by native electrophoresis.

"N" refers to the uncleaved control, and the time points are depicted in minutes. The expected migraton of oligomers is shown by arrows in accordance with the previous report on the behavior of CBS in native electrophoretic gels *(Kozich et al. 2010)*. Shifts in the migration of CBS during the proteolytic time course exhibit a dose-response pattern that was reported previously *(Kopecka et al. 2011)*. The analysis shows the absence any detectable monomeric or dimeric intermediate during the proteolytic time course.

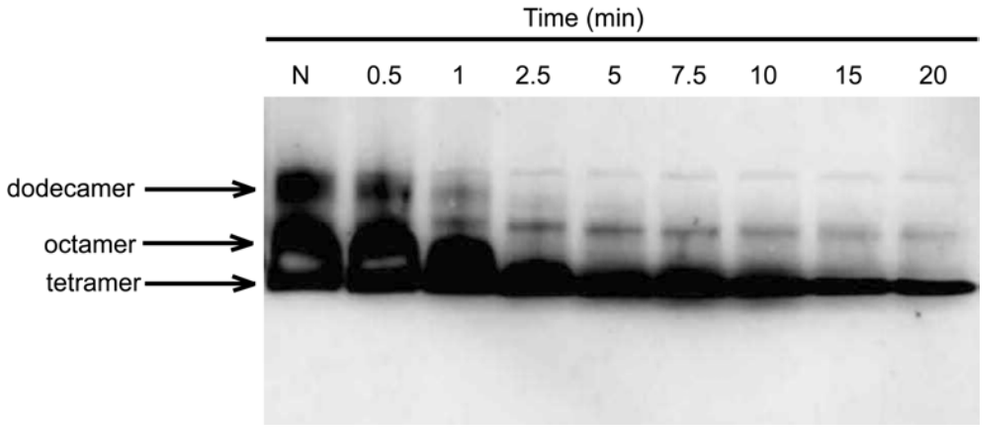


Figure S2: Nonlinear data fitting for wild-type CBS in bacterial crude extract. Points are the mean with standard deviations from four measurements. A. Proteolytic kinetics under native conditions; proteins in cell lysate (1 mg/ml) were cleaved by thermolysin (0.2 mg/ml). B and C. Dependence of k_{obs} (B) and the fraction of uncleaved protein B (C) on the concentration of thermolysin during proteolysis of wild-type CBS. D. Pulse proteolysis of wild-type CBS in bacterial crude extract. Proteins in cell lysate (1 mg/ml protein) were cleaved by thermolysin (0.2 mg/ml) for 60 seconds. The concentration of urea is plotted against f_{fold}, the fraction of folded proteins that remained uncleaved after the proteolytic pulse. The f_{fold} values at each point were calculated as a ratio of the remaining protein to the amount of uncleaved protein in the absence of urea.

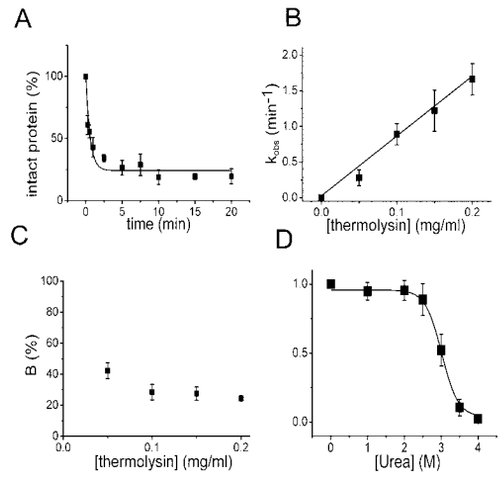


Figure S3: Proteolytic kinetics under native conditions of the wild-type and mutant CBS proteins at the thermolysin concentration that enabled the observation of the proteolytic time-course. Points are the mean with standard deviations from four measurements. Below each plot, the corresponding representative gels are shown. The values of k_p depicted in Table 2 were determined from these experiments. For cleavage of p.E302K and p.G307S, partially cleaved fragments were observed; this limited proteolysis was caused by the extremely low amounts of thermolysin in these experiments.

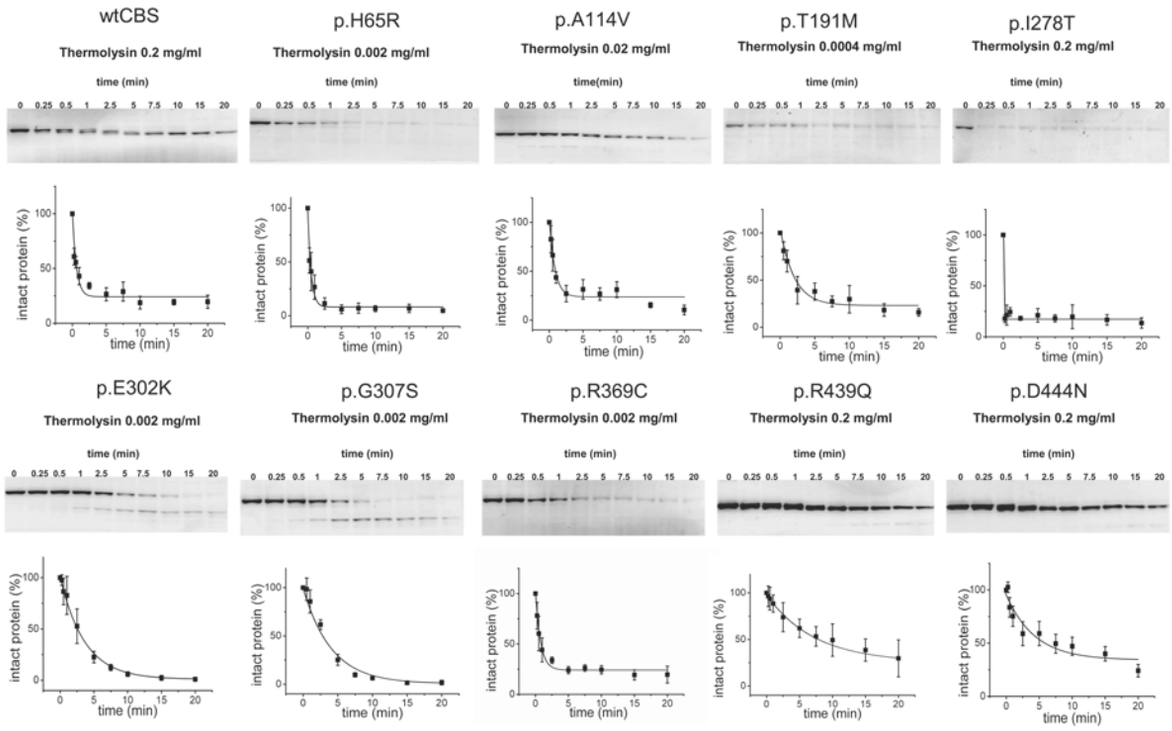


Figure S4: **Proteolytic cleavage of wild-type and mutant CBS in the concentration of urea that correspond to the beginning of the unfolding transition zone for each protein - plots with corresponding representative gels.** This experiment assessed the possible systematic error in the determination of the c_m value. Systematic error in determination of c_m value is negligible if kinetic constant of the cleavage is lower than 0.2 min^{-1} *(Park and Marqusee 2006)*.

A. The kinetic constant of wild-type CBS in 2.5M urea was determined to be 0.13 ± 0.01 min^{-1}.Points are the mean with standard deviations from four measurements. B. The kinetic constant of p.A114V in 2.5M urea was determined to be 1.14 ± 0.30 min^{-1}. Points are the mean with standard deviations from four measurements. Rapid cleavage indicates that c_m value for p.A114V may be determined with systematic error.

The cleavage of the unfolded mutants p.T191M, p.I278T and p.R369C (C) and the rigidified mutant and p.D444N (D) was negligible during the time intervals chosen. The mutant p.R439Q (D) was cleaved too slowly for reliable nonlinear data fitting during time interval chosen.

For wild-type CBS, p.R439Q and p.D444N, proteolytic fragments were observed; limited proteolysis was likely caused a by low catalytic activity of thermolysin in higher concentrations of urea.

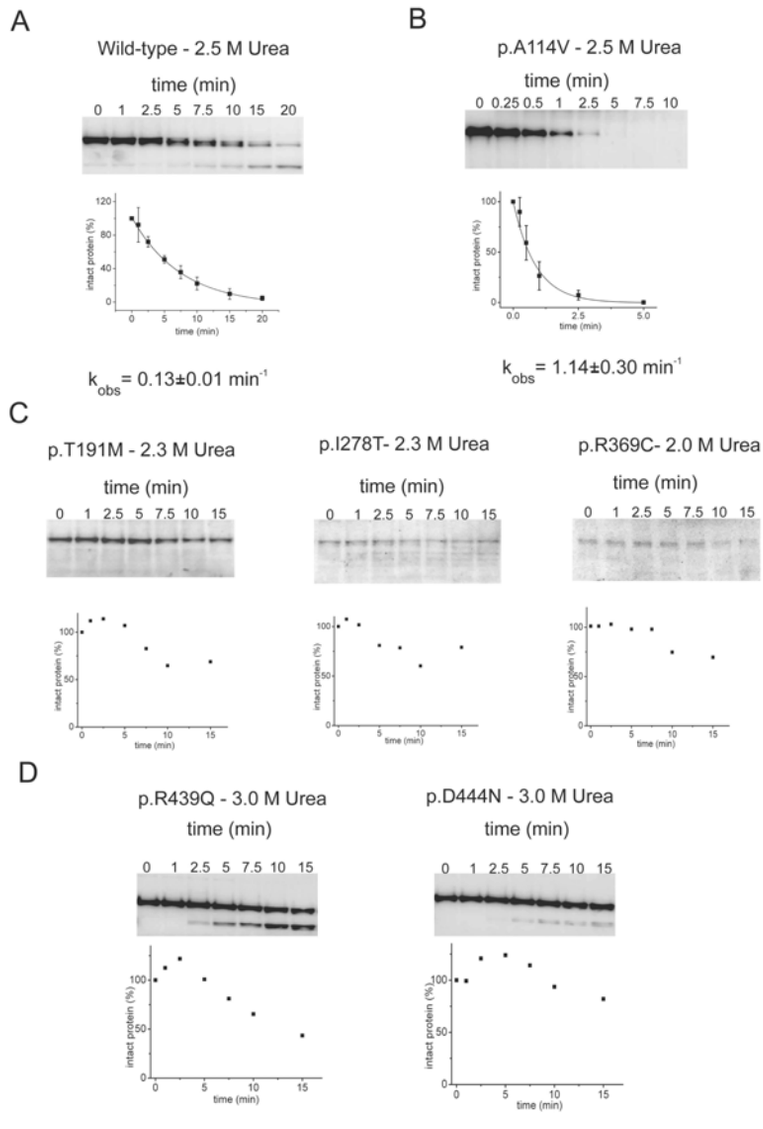


Figure S5: The amount of CBS proteins in crude cell extracts after expression in E. coli at 37 °C.

Bacterial cell lysates (20 μg total protein) were loaded to each lane after expression of CBS proteins using corresponding pHCS3 plasmids.

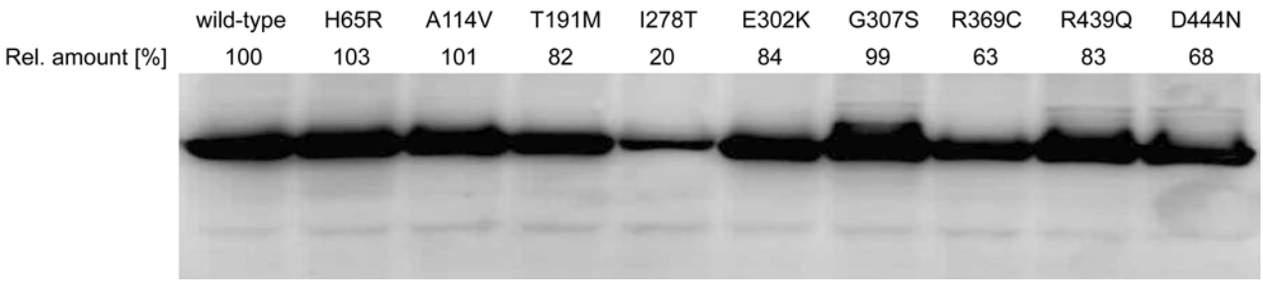


Figure S6: Pulse proteolysis of wild-type and mutant CBS proteins – plots with corresponding representative gels. The concentration of urea is plotted against f_{fold}, the fraction of folded proteins that remained uncleaved after the proteolytic pulse. The f_{fold} values at each point were calculated as a ratio of remaining protein to the amount of uncleaved protein in the absence of urea. Points are the mean with standard deviations from four measurements. Dashed lines in the plot of each mutant represent the curve determined for wild-type CBS. The presence of proteolytic fragments was observed only in several concentrations of urea for mutants p.A114V and p.R439Q; limited proteolysis was likely caused by a low catalytic activity of thermolysin in higher concentrations of urea.

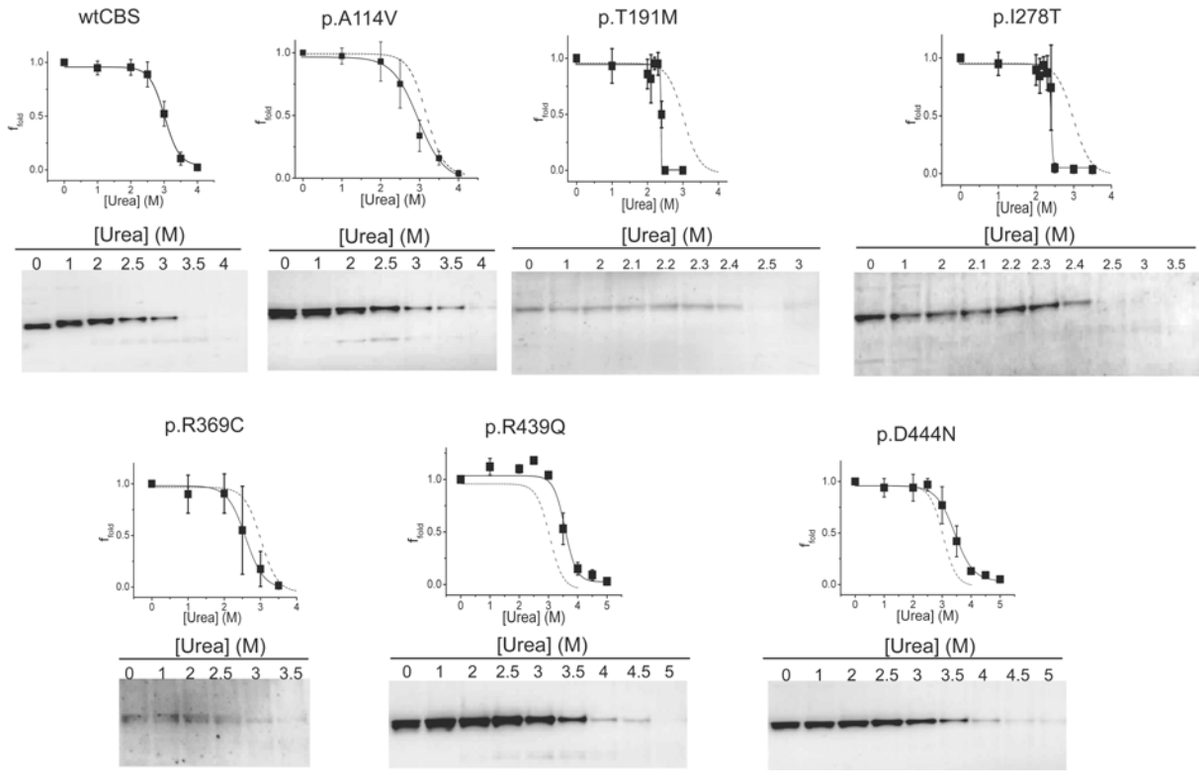


Figure S7: Evaluation of the pulse proteolysis of p.A114V , which calculates the correction on cleavage of partially unfolded proteins.

Table. For each concentration of urea, the following values were calculated: i, "*estimated* k_{obs}" is a kinetic constant for the proteolysis of p.A114V in urea concentrations if urea-induced unfolding did not occur. The values are calculated according to observed k_{obs} in the absence of urea together with theoretical decrease in k_{cat} of thermolysin in the urea gradient that was previously published *(Park and Marqusee 2004b)*. ii, "*estimated amount of uncleaved protein in the absence of urea-induced unfolding*" represents the amount of p.A114V that will be uncleaved after the pulse if urea-induced unfolding did not occur. It is calculated according to values of "*estimated* k_{obs}." iii, "*fraction of globally unfolded protein*" was calculated as the difference between the estimated and observed amount of uncleaved protein. iv, the fraction of partially unfolded protein is *100 -fraction of globally unfolded protein.* Using this estimation, we assessed the degree of urea-induced unfolding of the protein.

In the plots below the table, dashed lines represent the curve of wild-type CBS. Points are the mean with standard deviation from four measurements. *Left*. the default analysis neglects the cleavage of the partially folded mutant at the beginning of the apparent transition zone. *Right*. the plot shows the corrections due to the rapid cleavage of the partially unfolded mutant protein based on the data from the table.

Urea (M)	estimated k_{obs} (min^{-1})	estimated amount of uncleaved protein in absence of urea-induced unfolding (%)	observed amount of uncleaved protein (%)	fraction of globally unfolded proteins (%)	fraction of partially unfolded proteins (%)
0	0.80	44.9	44.9	0	100
1	0.38	68.1	52.4	15.7	84.3
2	0.18	83.2	44.6	38.6	61.4
2.5	0.13	88.0	40.0	48.0	52.0
3	0.09	91.6	16.2	75.4	24.6
3.5	0.06	93.8	7.7	86.1	13.9
4	0.03	97.2	2.5	94.7	5.3

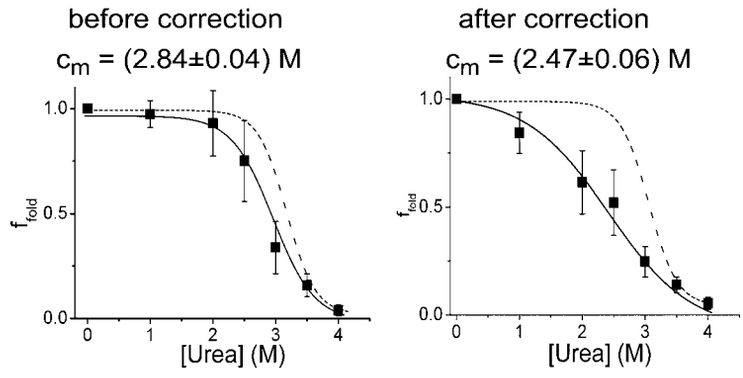
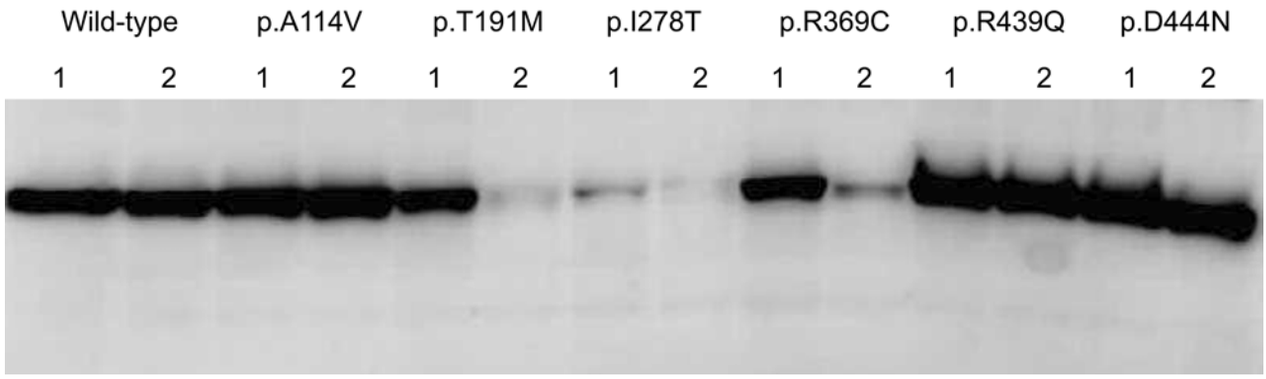


Figure S8: Differences in CBS signal after overnight incubation in the absence of urea at 4 °C. Bacterial lysates containing CBS proteins (20 µg total protein per lane) are shown with identical numbering: 1 – freshly thawed sample; 2 – freshly thawed sample after overnight incubation at 4 °C. Extensively unfolded mutants (p. T191M, p.I278T and p.R369C) showed a significant decrease in the amount of CBS antigen after the overnight incubation, indicating the formation of aggregates.

Table. The residual amount of the CBS antigen was determined using densitometry. The values are related to the amount of freshly thawed samples containing the corresponding CBS protein.



Protein	Residual amount of the CBS antigen after overnight incubation [%]
Wild-type	94
p.A114V	120
p.T191M	10
p.I278T	66
p.R369C	20
p.R439Q	118
p.D444N	113

Manuscript 5.1.5

A. Hnízda et al.

Conformational properties of nine purified cystathionine beta-synthase mutants

Conformational properties of nine purified cystathionine beta-synthase mutants

*Aleš Hnízda[1†], Tomáš Majtan[2,3†], Lu Liu[4], Angel L. Pey[5], John F. Carpenter[4], Milan Kodíček[6], Viktor Kožich[1] and Jan P. Kraus[2]**

[1] Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague, Ke Karlovu 2, 128 08 Prague 2, Czech Republic

[2] Department of Pediatrics and the Colorado Intellectual and Developmental Disabilities Research Center (IDDRC), University of Colorado at Denver, 12800 E 19th Ave, Aurora, Colorado 80045, USA

[3] Department of Genomics & Biotechnology, Institute of Molecular Biology, Slovak Academy of Sciences, Bratislava, 84551, Slovakia

[4] Department of Pharmaceutical Sciences, School of Pharmacy, University of Colorado at Denver, Aurora, Colorado 80045, USA

[5] Department of Physical Chemistry, Faculty of Sciences, University of Granada, Granada 18071, Spain

[6] Department of Biochemistry and Microbiology, Institute of Chemical Technology in Prague, Technická 5, 166 28, Prague 6, Czech Republic

[†] Authors contributed equally

* Corresponding author; tel: (303) 724-3812; fax: (303) 724-3838; email: jan.kraus@ucdenver.edu

Running title: Conformational Analysis of Purified CBS Mutants

FUNDING SOURCE STATEMENT: This work was supported by the research programs of the Charles University in Prague PRVOUK-P24/LF1/3, UNCE 204011 and by the grant No. 7310 from the Grant Agency of the Charles University in Prague (to AH and VK), by Postdoctoral Fellowship 0920079G from the American Heart Association (to TM), by National Institutes of Health Grant HL065217, by American Heart Association Grant In-Aid 09GRNT2110159 and by a grant from the Jerome Lejeune Foundation (to JPK), by Ramon y Cajal research contract RYC-2009-04147 from the Spanish Ministry of Science and Innovation (to ALP) and grant P11-CTS-7187 from the Junta de Andalucia (to ALP).

ABBREVIATIONS: AdoMet, S-adenosylmethionine; CBS, cystathionine beta-synthase; CBSDH, cystathionine beta-synthase deficient homocystinuria; CD, circular dichroism; c_m, midpoint of urea-induced unfolding determined by pulse proteolysis; k_p, rate constant of proteolysis under native conditions; PLP, pyridoxal-5´-phosphate

ABSTRACT

Protein misfolding due to missense mutations is a common pathogenic mechanism in cystathionine beta-synthase (CBS) deficiency. In our previous studies, we have successfully expressed, purified and characterized nine CBS mutant enzymes containing the following patient mutations: P49L, P78R, A114V, R125Q, E176K, R266K, P422L, I435T and S466L. These purified mutants exhibited full heme saturation, normal tetrameric assembly and high catalytic activity. In this work, we used several spectroscopic and proteolytic techniques to provide a more thorough insight into the conformation of these mutant enzymes. Far-UV circular dichroism, fluorescence and second derivative-UV spectroscopy revealed that the spatial arrangement of these CBS mutants is similar to the wild-type although the microenvironment of the chromophores may be slightly altered. Using proteolysis with thermolysin under native conditions, we found that the majority of the studied mutants is more susceptible towards cleavage suggesting their increased local flexibility or propensity to local unfolding. Interestingly, the presence of the CBS allosteric activator, S-adenosylmethionine (AdoMet), increased the cleavage rate of wild-type and the AdoMet-responsive mutants, while the proteolytic rate of the AdoMet-unresponsive mutants was not significantly changed. Pulse proteolysis analysis suggested that the protein structure of the R125Q and E176K mutants is significantly less stable than that of wild-type and the other mutants. Taken together, the proteolytic data show that the conformation of pathogenic mutants is altered despite retained catalytic activity and normal tetrameric assembly. This study demonstrates that the proteolytic techniques are a useful tool for the assessment of the biochemical penalty of missense mutations in CBS.

INTRODUCTION

Cystathionine beta-synthase (CBS) deficient homocystinuria (CBSDH; OMIM# 236200) is the most common inherited defect in sulfur amino acids metabolism characterized by severely elevated levels of plasma homocysteine, methionine and S-adenosyl-L-homocysteine (*1*). The worldwide frequency is estimated to be around 1:330,000 (*1*), thus classifying CBSDH as a rare disease. Interestingly, this number may be underestimated as several molecular epidemiological studies revealed the prevalence of homozygosity or compound heterozygosity for pathogenic mutations to be around 1:10,000 (*2-5*). If untreated, CBSDH manifests clinically with connective tissue symptoms such as dislocated lenses and skeletal abnormalities, mental retardation and vascular complications, particularly thromboembolic episodes (*1*). About half of CBS deficient patients respond to a treatment with pharmacological doses of pyridoxine (vitamin B6) with a significant lowering of plasma homocysteine levels and an alleviation of the clinical phenotype. The treatment of pyridoxine non-responsive patients involves a low methionine diet and supplementation with betaine which lowers homocysteine by promoting its remethylation to methionine via betaine:homocysteine S-methyltransferase.

Cystathionine beta-synthase (EC# 4.2.1.22) is a pyridoxal-5'-phosphate (PLP) dependent hemeprotein which catalyzes condensation of serine and homocysteine to form cystathionine as the first committed step in transsulfuration and subsequent biosynthesis of cysteine, glutathione and taurine. Human enzyme has a modular structure and a complex regulatory behavior (reviewed in (*6,7*)). More than 160 mutant alleles have been described in CBSDH patients so far (http://cbs.lf1.cuni.cz/cbsdata/cbsmain.htm) with missense mutations as the most common variants in the CBS gene accounting for 87% of analyzed patients alleles. Our previous study proposed that misfolding of CBS mutants may be responsible for pathogenicity in CBS deficiency as majority of tested CBS mutants formed large inactive aggregates devoid of heme (*8*). This proposed mechanism was later supported by several studies exploring the beneficial effect of chemical chaperones on the recovery, activity and assembly of various CBS mutant proteins (*9-11*). The presence of chemical chaperones such as ethanol or dimethylsulfoxide during expression in bacteria permitted the purification of several CBS mutants which exhibit normal heme saturation, native tetrameric assembly and the same or higher specific activities than the wild type CBS (*12*). Although the recovery of fully active CBS mutants suggests an improved folding, the final conformation of the purified mutants most likely differs from that of wild-type protein as the purified fully active CBS mutants varied in their response to exogenous addition of PLP cofactor as well as in their response to S-adenosyl-L-methionine (AdoMet) stimulation (*12*). In addition to studying spatial arrangement of proteins by spectroscopic monitoring of their conformation, proteolytic techniques such as proteolysis under native conditions and pulse proteolysis proved to be useful and effective to examine some aspects of protein structure (*13,14*); this

approach uses thermolysin, an endoprotease that preferentially cleaves peptide bonds in regions containing hydrophobic amino acids. The rate of proteolysis with thermolysin under native conditions reveals the extent of unfolding of the studied proteins because this endoprotease can only cleave flexible regions and partially or globally unfolded structures (*15*). Further, pulse proteolysis is a technique that uses thermolysin to monitor urea-induced unfolding. After a short proteolytic pulse, the fraction of folded proteins remains intact whereas the locally and/or globally unfolded species are digested. Using this premise, the protein is analyzed in varying concentrations of urea and the c_m value, a measure of protein thermodynamic stability in the urea gradient, may be determined (*14*). The recent study of several CBS mutants in crude extracts showed that such techniques can reveal conformational abnormalities even in CBS mutants with normal activity and/or the degree of assembly similar to the wild-type enzyme (*16*).

In this study we applied several spectroscopic and proteolytic techniques to get an insight into the changes in CBS protein conformation induced by disease-causing missense mutations. We studied nine purified mutants P49L, P78R, A114V, R125Q, E176K, R266K, P422L, I435T and S466L. The purification and biochemical characterization of these mutants was reported elsewhere (*12,17*); the determination of their conformational properties adds additional new information thus permitting an understanding of pathogenicity of missense mutations in CBS.

EXPERIMENTAL PROCEDURES

Purification of CBS proteins. The wild-type and mutant CBS enzymes except for the R266K CBS were expressed as fusion proteins with the N-terminal GST in *E. coli* Rosetta2 (DE3) cells and purified essentially as described previously (*12*). The R266K mutant enzyme was expressed in *E. coli* Rosetta2 (DE3) cells with a C-terminal 6xHis tag and isolated following the previously reported procedure (*17*).

Circular dichroism (CD) and fluorescence spectroscopy. The CD spectra of CBS proteins (0.5 mg/ml; 50 mM phosphate buffer, pH 7.5) were recorded by chiroptic spectrometer Jasco J-810. The intrinsic fluorescence of CBS proteins was measured in the same buffer using Perkin Elmer LS55 fluorescence spectrometer. The excitation wavelength was 295 nm (slit width 5 nm) for tryptophans and 420 nm (slit width 10 nm) for internal aldimines with an emission signal scanned from 300 to 450 nm (slit width 5 nm) and from 430 to 700 nm (slit width 10 nm), respectively.

Second-derivative UV spectroscopy. The UV spectra were recorded between 10 and 90°C with 2.5°C increments and a 3 min equilibrium time at each temperature. Measurements were made on an Agilent diode array model 8453 UV-visible spectrophotometer equipped with a Peltier temperature controller. Two ml of protein samples (0.2 mg/ml) were prepared by diluting the stock protein with the appropriate

amount of buffer (20 mM HEPES pH 7.4, 1 mM Tris-(2-carboxyethyl)-phosphine, 0.01% Tween 20) and then placed in a quartz cuvette with a 1 cm path length. A micro stir bar (100 rpm) was put into the cuvette improve the heat-exchange in the sample. The second derivative UV spectra at each temperature were used to compare the changes of tertiary structure of proteins. The peak positions for aromatic amino acids tryptophan, tyrosine and phenylalanine were plotted as a function of temperature.

Proteolysis with thermolysin under native conditions and pulse proteolysis. Proteolytic techniques were performed and evaluated according to the previously published procedures (*18*). Each experiment was repeated at least twice.

The rate of proteolysis under native conditions was expressed as k_p, the constant in a single exponential equation.

Pulse proteolysis revealed the c_m value, i.e. concentration of urea at which fraction of folded proteins comprises 50 % of the entire protein population. To assess accuracy of the determined c_m values, we observed proteolytic cleavage of the CBS proteins in 2M urea solution. The kinetic constant for proteolysis of all proteins with determined c_m was lower than 0.2 min^{-1} indicating negligible cleavage of the protein during the proteolytic pulse at the beginning of the transition zone and thus the absence of a systematic error (*19*).

RESULTS

Mutants analyzed in this study. All studied mutants were purified into homogeneity. They retained catalytic activity and their saturation with heme and tetrameric assembly were similar to wild-type CBS as previously reported (*12*). These mutations are located a) in the proximity of the heme-binding pocket (P49L, R125Q and R266K), b) at the dimer interface (P78R, A114V and E176K), and c) in the regulatory domain (P422L, I435T and S466L) (Fig.1).

Far UV-CD spectroscopy. To assess the impact of mutations on the formation of the secondary structure of the protein, we analyzed CBS mutants by far UV-CD spectroscopy. Far-UV CD spectra (Fig. S1 in the Supporting Information) showed maxima at ~208 and ~222 nm indicating a large content of helical structures. No noticeable differences in the content of secondary structure were observed among the studied CBS mutants compared to wild-type enzyme with the single exception of the E176K mutant. This mutant exhibited a different shape of the CD spectrum indicating a partial decrease in proportion of helical structures together with formation of a different type of secondary structure; nevertheless, these changes are subtle and the overall helical content is retained. Far-UV CD spectra obtained in the presence of AdoMet did not show any change in the overall secondary structure compared to the CD spectra in the absence of AdoMet (data not shown). Taken together, the CD results

suggest that the majority of CBS mutants possessed unaltered secondary structure compared to that of wild-type enzyme and the presence of AdoMet does not alter the secondary structure in CBS enzymes.

Fluorescence spectroscopy. Using fluorescence spectroscopy, we studied conformation of the CBS mutants by analyzing the microenvironment of the tryptophans. Tryptrophan fluorescence spectra for all CBS enzymes exhibited broad maxima resulting from the fluorescence of eight tryptophan residues in each CBS subunit (Fig.2). These spectra did not show any major differences in the position of emission maxima indicating an unaltered polarity in the tryptophan environment. However, significant changes were observed for relative intensities of the emission, namely, increased intensity for P49L, R125Q and S466L and, on the contrary, decreased intensity for R266K in comparison to the wild-type. The altered intensity of emission may be caused by a different microenvironment of tryptophans. The altered relative orientation of one or more tryptophans towards possible quenching groups, such as polar amino acid residues, and/or mutual position of tryptophan and tyrosine residues, affecting resonance energy transfer from hydroxyphenyl to indolyl groups, could have caused the intensity changes observed (*20*).

Fluorescence spectroscopy was further used for analyses of PLP that is covalently bound in the active site of CBS in the form of aldimine. These measurements revealed that the majority of the mutants exhibited decreased intensity of emission (Fig.2). These changes may be caused by increased extent of quenching by surrounding amino acid residues or by the heme moiety (*21*). Alternatively, lower intensity of fluorescence may be attributed to a decreased saturation of CBS proteins with PLP that was previously observed for R125Q, E176K and partially also for the C-terminal mutants P422L, I435T and S466L (*12*). In summary, both fluorescence-based measurements showed that the majority of the studied mutants are similar to wild-type CBS with minor differences in the microenvironment of the fluorophores.

Second-derivative UV spectroscopy. To determine possible changes in the microenvironment of chromophores, we also performed a second-derivative UV spectroscopy that may reveal information about the microenvironments of tryptophan, tyrosine and phenylalanine residues (*22*). Positions of the maxima were unaltered for the majority of the studied mutants (see P49L as an example for the unaltered mutant CBS spectrum in Fig.3). A decrease in the wavelength of the maximum for tryptophan and tyrosine were recorded for A114V, R125Q and E176K (Fig.3; spectral traces are shown at Fig.S2 in the Supporting Information) suggesting possible differences in the microenvironment of these chromophores. It should be noted that these changes were subtle, typically around 0.1 nm. Nevertheless, the altered positions of maxima were consistently observed for all temperatures below the melting point, i.e. from 10 °C to 55 °C (Fig.3). These data further suggest that some of the studied mutants may differ from the wild-type in the local spatial arrangement of the CBS protein and/or in a mutual position of the catalytic core and the C-terminal AdoMet-binding regulatory domain.

Proteolysis with thermolysin under native conditions. To determine the structural flexibility and tendency towards unfolding of the CBS mutants, we assessed their proteolytic susceptibility towards thermolysin under native conditions.

Three mutants, namely P49L, P78R and R266K, were resistant towards proteolysis to an extent similar or even higher than the wild-type. The remaining six CBS mutants showed higher susceptibility to proteolysis and extremely rapid cleavage was observed for R125Q and E176K (Tab.1 and Fig.4). In case of the mutant I435T we were not able to determine proteolytic kinetics since this mutant was almost instantly, within several seconds, cleaved to a fragment having molecular weight of ≈ 40 kDa (Fig.4B). Thus one of the requirements for this technique, i.e. the cleavage without formation of major fragments, has not been met.

We also studied the effect of AdoMet on the proteolytic kinetics (Tab.1). We have demonstrated previously that allosteric activation of CBS is associated with opening of the protein conformation and that the wild-type CBS is cleaved in the presence of AdoMet more rapidly with doubling of the kinetic constant (*18*). The similar increase in cleavage rate was also observed in this study for all AdoMet-responsive CBS mutants, namely the P49L, P78R, A114V and R266K. In case of the A114V mutant, the presence of the AdoMet in the proteolytic assay promoted an extensive opening of conformation accompanied with a dramatic 9-fold increase in the rate of cleavage compared to the rate in the absence of the ligand. In contrast, virtually no change in the cleavage rate in the presence of AdoMet was observed for the mutants with the highest cleavage rates (R125Q and E176K) and for the C-terminal mutants (P422L and S466L). Interestingly, the same mutants did not show any response to AdoMet allosteric stimulation of catalytic activity in our previous study (*12*).

Data from proteolysis with thermolysin under native conditions suggest that the majority of the studied CBS mutants has higher structural flexibility and is more susceptible to proteolytic cleavage than the wild-type enzyme despite only subtle changes in protein conformation observed by spectroscopic techniques. Additionally, increased proteolytic rates of the wild-type and AdoMet-responsive mutant enzymes in the presence of AdoMet are indicative of conformational rearrangement upon AdoMet binding. On the contrary, the cleavage rates of AdoMet-unresponsive mutants were not altered suggesting their inability to change conformation upon AdoMet binding and thus explaining their inability to be allosterically stimulated by AdoMet.

Pulse proteolysis. Using pulse proteolysis, we have complemented the proteolytic data under native conditions by determining the resistance of wild-type and mutant CBS proteins against urea-induced denaturation. Interestingly, the majority of the mutants were found to be more resistant towards urea-induced denaturation than the wild-type (Tab.2; a representative gel with corresponding curve of the P49L mutant is shown at Fig.3 in the Supporting Information). On the other hand, mutants R125Q and E176K - sensitive to rapid proteolytic cleavage - exhibited abnormal behavior in pulse proteolysis

which resulted in the non-sigmoidal curves of the urea-induced unfolding (Fig.S4 in the Supplementary Information); this finding indicates very low cooperativity of the proteins tertiary structure.

The presence of AdoMet increased the stability of wild-type CBS as reported previously (18,23) whereas the c_m values for the majority of mutants were not changed (Tab.2). Taken together, the stability of the CBS mutants in the absence of AdoMet did not change after addition of AdoMet and was comparable to the stability of wild-type CBS incubated with AdoMet. Different behavior was observed only for the R266K mutant which, unlike the remaining wild-type and mutant CBS enzymes, exhibited a significantly lowered protein stability upon AdoMet binding with concomitantly decreased c_m value.

DISCUSSION

In this study, we analyzed the conformational properties of nine disease-causing CBS mutant enzymes that have normal catalytic activity and for which the structural causes of their pathogenicity were not apparent from the previous work (*12*). We compared their conformational properties with the wild-type enzyme using far-UV CD, fluorescence and second-derivative UV spectroscopy. In the next step, we determined their structural flexibility by proteolysis with thermolysin under native conditions as well as their sensitivity towards urea-induced denaturation by pulse proteolysis. Only subtle conformational differences between the wild-type and the mutants studied were detected by the spectroscopic methods used. These findings are consistent with our previous work showing that the studied mutants did not exhibit dramatic abnormalities in the specific activities, heme saturation or native tetramer formation (*12*).

The major differences in the properties of CBS mutants were observed when tested for their proteolytic susceptibility under native conditions. These data suggest that the mutant proteins adopted conformation, which differs from the wild-type CBS, being more flexible and exposing more hydrophobic residues for the thermolysin to attack. Moreover, the less compact structure of the mutants A114V, R125Q and E176K can also be assumed from their spectra of the second-derivative UV spectroscopy which revealed subtle decrease in the wavelength for the maxima assigned for tryptophan and tyrosine residues; similar blue shift was observed in unfolding of several model proteins (*24*).

Interestingly, the increased structural mobility of the mutants is accompanied by impaired protein stability in urea solution only for the extensively flexible CBS mutants, R125Q and E176K, while the other CBS mutants exhibit unaffected global protein stability as was demonstrated by pulse proteolysis. The same or even higher resistance against urea-induced unfolding compared to the wild-type is also consistent with the thermostability of these mutants previously determined by absorption spectrophotometry (*12*). Analogous increased proteolytic susceptibility associated with subtle conformational changes and with unaltered thermodynamic stability was reported for the yeast

phosphoglycerate kinase compared to its ortholog from *E. coli* (*15*). It was proposed that the discrepancies in the phosphoglycerate kinase orthologs were caused by a divergent interdomain cooperativity and consequently different mechanisms of unfolding in these modular proteins. It is tempting to speculate that the increased structural mobility of majority of the CBS mutants is not caused by their low thermodynamic stability but more likely by the lower kinetic barrier of the protein unfolding; the altered interdomain communication, despite only minor conformational changes in each particular domain, may be responsible for the increased unfolding rates of these mutants.

The findings of increased proteolytic susceptibility of CBS mutants towards thermolysin are consistent with our previous study conducted directly in bacterial lysates for a different set of CBS mutants (*16*). However, higher proteolytic susceptibility of all studied mutants in the previous study was associated with their increased sensitivity toward urea-induced denaturation. This discrepancy may be due to a different panel of studied CBS mutations and different degree of purity of CBS proteins in each study. In the present study, we analyzed CBS mutants that were successfully purified to homogeneity after expression in *E. coli* whereas the previous work was carried out mainly on mutants that were not amenable to purification due to an excessive aggregation. It indicates that decreased global protein stability may be observed only for the severely affected CBS mutants but not for the mutants exhibiting subtle conformational changes. This notion is also supported by the present study showing that only the R125Q and E176K CBS mutants exhibited impaired protein stability in a urea gradient.

Since the altered response to AdoMet was proposed as one of the possible pathogenic mechanisms in CBS deficiency, particularly for the C-terminal missense CBS mutations, we compared the kinetics of proteolytic cleavage of wild-type and mutant CBS enzymes in the presence of this allosteric activator. The most rapidly cleaved CBS mutants, namely R125Q and E176K, exhibited unaltered k_p in the presence of AdoMet compared to that obtained in the absence of AdoMet. This indicates that these mutants cannot bind AdoMet and/or are locked in a specific conformation that prevents allosteric change upon AdoMet binding. This suggestion is also supported by the previously reported lack of stimulation of the catalytic activity by AdoMet and by heating of mutants (*12*). More surprisingly, extremely increased proteolytic susceptibility of the A114V mutant in the presence of AdoMet suggests that its allosteric activation is likely associated with an extensive opening of the folded structure exposing naturally buried hydrophobic residues on the protein surface. Interestingly, the effect of AdoMet on native proteolysis of the R266K mutant was similar to wild type CBS but different behavior of this mutant was observed using pulse proteolysis. The presence of AdoMet led to a lower c_m value indicating that this CBS ligand decreases thermodynamic stability of the R266K mutant. This result correlates well with the previously observed decreased thermal stability and AdoMet activation of this mutant compared to wild-type CBS (*17*). The impaired response to AdoMet activation was also observed for the C-terminal mutants. The proteolytic cleavage of the P422L mutant was not

significantly increased and, moreover, the S466L mutant was cleaved even less rapidly in the presence of AdoMet. These findings further support the previous notion that the C-terminal mutants are locked in a specific conformation, which results in a permanently activated mutant CBS enzymes lacking proper response to AdoMet stimulation (*25,26*). Interestingly, our results indicate that these mutant proteins may be more flexible in the absence of AdoMet than the wild-type. Even though the S466L mutant does not respond to AdoMet, it is still capable of binding it as reported previously (*25*). As CBS domains in the C-terminal region very likely fold independently of the catalytic core (*27*), these mutants may still bind AdoMet, but are apparently unable to rearrange their conformation. It is tempting to speculate that this locked conformation is *in vivo* recognized as a misfolded structure by the cellular control machinery and are consequently targeted for degradation (*28*).

It should be noted that the mutants P49L and P78R did not exhibit structural abnormalities by the approaches used in this study, which correlates well with their biochemical properties very similar to that of wild type enzyme (*12*). The P49L exhibited high catalytic activity when expressed in the pKK expression vector without any additional tags or fusion partners (*29*). On the contrary, the P78R possessed a decreased enzyme activity in the same study using the pKK construct which is consistent with a study analyzing the purified mutant reported by the Banerjee group (*30*). Data on P78R indicate that pathogenicity of these mutants may be revealed by employing different expression system or using specific conditions. Nevertheless, the P49L mutation is often associated with very mild clinical manifestions (CBS Mutation Database;http://cbs.lf1.cuni.cz/cbsdata/cbsmain.htm and Sally Stabler, oral communication at the 8th Conference on Homocysteine Metabolism, Lisboa 2011).

It should be noted that the application of a bacterial expression systems that produce the target protein with an affinity tag may have an artificial effect on quality of the purified proteins. For instance, the GST-tag that was used for the majority of the mutants may increase protein stability during the expression and subsequent purification (*31*). The dramatic effect of the type and position of the employed purification affinity tag (i.e. bulky fusion partner, such as GST versus short flexible tag, such as 6xHis) and its position on the proper folding was recently demonstrated in the expression studies of the R266K mutant (*17*). Nevertheless, using of affinity tags is necessary for production of target proteins in sufficient yields for their conformational analysis.

In this study, we demonstrated that protein structures of the studied CBS mutants are more locally flexible than that of the wild-type despite their normal catalytic activity and unaffected sensitivity towards urea-induced denaturation. In conclusion, the conformational analysis of the mutants using spectroscopic and proteolytic approach proved to be a useful tool for the assessment of the biochemical penalty of the CBS mutations.

ACKNOWLEDGMENT

We thank to Ms. Kateřina Raková and to Jitka Honzíková, MSc. for technical assistance and to Dr. Jose M. Sanchez-Ruiz for useful comments on the manuscript.

REFERENCES

1. Mudd, S. H., Levy, H. L., and Kraus, J. P. (2001) Disorders of transsulfuration, in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Childs, B., Kinzler, K., and Vogelstein, B., Eds.), 8th ed., McGraw-Hill, New York.

2. Gaustadnes, M., Ingerslev, J., and Rutiger, N. (1999) Prevalence of congenital homocystinuria in Denmark, *N. Engl. J. Med. 340*, 1513.

3. Refsum, H., Fredriksen, A., Meyer, K., Ueland, P. M., and Kase, B. F. (2004) Birth prevalence of homocystinuria, *J. Pediatr. 144*, 830-832.

4. Janosik, M., Sokolova, J., Janosikova, B., Krijt, J., Klatovska, V., and 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*, 431-437.

5. Linnebank, M., Homberger, A., Kraus, J. P., Harms, E., Kozich, V., and Koch, H. G. (2001) Haplotyping of wild type and I278T alleles of the human cystathionine beta-synthase gene based on a cluster of novel SNPs in IVS12, *Hum. Mutat. 17*, 350-351.

6. Miles, E. W., and Kraus, J. P. (2004) Cystathionine beta-synthase: structure, function, regulation, and location of homocystinuria-causing mutations, *J. Biol. Chem. 279*, 29871-29874.

7. Singh, S., Madzelan, P., and Banerjee, R. (2007) Properties of an unusual heme cofactor in PLP-dependent cystathionine beta-synthase, *Nat. Prod. Rep. 24*, 631-639.

8. Janosik, M., Oliveriusova, J., Janosikova, B., Sokolova, J., Kraus, E., Kraus, J. P., and Kozich, V. (2001) Impaired heme binding and aggregation of mutant cystathionine beta-synthase subunits in homocystinuria, *Am. J. Hum. Genet. 68*, 1506-1513.

9. Singh, L. R., Chen, X., Kozich, V., and Kruger, W. D. (2007) Chemical chaperone rescue of mutant human cystathionine beta-synthase, *Mol. Genet. Metab. 91*, 335-342.

10. Singh, L. R., and Kruger, W. D. (2009) Functional rescue of mutant human cystathionine beta-synthase by manipulation of Hsp26 and Hsp70 levels in Saccharomyces cerevisiae, *J. Biol. Chem. 284*, 4238-4245.

11. Kopecka, J., Krijt, J., Rakova, K., and Kozich, V. (2011) Restoring assembly and activity of cystathionine beta-synthase mutants by ligands and chemical chaperones, *J. Inherit. Metab. Dis. 34*, 39-48.

12. Majtan, T., Liu, L., Carpenter, J. F., and Kraus, J. P. (2010) Rescue of cystathionine beta-synthase (CBS) mutants with chemical chaperones: purification and characterization of eight CBS mutant enzymes, *J. Biol. Chem. 285*, 15866-15873.

13. Park, C., and Marqusee, S. (2004) Probing the high energy states in proteins by proteolysis, *J. Mol. Biol. 343*, 1467-1476.

14. Park, C., and Marqusee, S. (2005) Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding, *Nat. Methods 2*, 207-212.

15. Young, T. A., Skordalakes, E., and Marqusee, S. (2007) Comparison of proteolytic susceptibility in phosphoglycerate kinases from yeast and E. coli: modulation of conformational ensembles without altering structure or stability, *J. Mol. Biol. 368*, 1438-1447.

16. Hnizda, A., Jurga, V., Rakova, K., and Kozich, V. Cystathionine beta-synthase mutants exhibit changes in protein unfolding: conformational analysis of misfolded variants in crude cell extracts, *J. Inherit. Metab. Dis*., in press, DOI 10.1007/s10545-010-9178-3

17. Majtan, T., and Kraus, J. P. (2012) Folding and activity of mutant cystathionine beta-synthase depends on the position and nature of the purification tag: Characterization of the R266K CBS mutant, *Protein Expr. Purif. 82*, 317-324.

18. Hnizda, A., Spiwok, V., Jurga, V., Kozich, V., Kodicek, M., and Kraus, J. P. (2010) Cross-talk between the catalytic core and the regulatory domain in cystathionine beta-synthase: study by differential covalent labeling and computational modeling, *Biochemistry 49*, 10526-10534.

19. Park, C., and Marqusee, S. (2006) Quantitative determination of protein stability and ligand binding by pulse proteolysis, in *Current protocols in protein science* (Coligan, J.E., Dunn, B. M., Speicher, D.W., Wingfield, P.T., and Ploegh, H.L., Eds.), pp 20.11.1-20.11.14, Wiley, Hoboken.

20. Willaert, K., and Engelborghs, K. W. (1991) The quenching of tryptophan fluorescence by protonated and unprotonated imidazole, *Eur. Biophys. J. 20,* 177–182.

21. Kery, V., Poneleit, L., Meyer, J. D., Manning, M. C., and Kraus, J. P. (1999) Binding of pyridoxal 5'-phosphate to the heme protein human cystathionine beta-synthase, *Biochemistry 38*, 2716-2724.

22. Mach, H., and Middaugh, C. R. (1994) Simultaneous monitoring of the environment of tryptophan, tyrosine, and phenylalanine residues in proteins by near-ultraviolet second-derivative spectroscopy, *Anal. Biochem. 222*, 323-331.

23. Prudova, A., Bauman, Z., Braun, A., Vitvitsky, V., Lu, S. C., and Banerjee, R. (2006) S-adenosylmethionine stabilizes cystathionine beta-synthase and modulates redox capacity, *Proc. Natl. Acad. Sci. U. S. A. 103*, 6489-6494.

24. Wetlaufer, D.B. (1962) Ultraviolet spectra of proteins and amino acids, *Adv. Protein. Chem. 17*, 303-390.

25. Janosik, M., Kery, V., Gaustadnes, M., Maclean, K. N., and Kraus, J. P. (2001) 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*, 10625-10633.

26. Sen, S., Yu, J., Yamanishi, M., Schellhorn, D., and Banerjee, R. (2005) Mapping peptides correlated with transmission of intrasteric inhibition and allosteric activation in human cystathionine beta-synthase, *Biochemistry 44*, 14210-14216.

27. Koutmos, M., Kabil, O., Smith, J. L., and Banerjee, R. (2010) Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine {beta}-synthase, *Proc. Natl. Acad. Sci. U. S. A. 107*, 20958-20963.

28. Gupta, S., Wang, L., Hua, X., Krijt, J., Kozich, V., and Kruger, W. D. (2008) Cystathionine beta-synthase p.S466L mutation causes hyperhomocysteinemia in mice, *Hum. Mutat. 29*, 1048-1054.

29. Kozich, V., Sokolova, J., Klatovska, V., Krijt, J., Janosik, M., Jelinek, K., and Kraus, J. P. (2010) Cystathionine beta-synthase mutations: effect of mutation topology on folding and activity, *Hum. Mutat. 31*, 809-819.

30. Sen, S., and Banerjee, R. (2007) A pathogenic linked mutation in the catalytic core of human cystathionine beta-synthase disrupts allosteric regulation and allows kinetic characterization of a full-length dimer, *Biochemistry 46*, 4110-4116.

31. Waugh, D. S. (2005) Making the most of affinity tags, *Trends Biotechnol. 23*, 316-320.

TABLES

Table 1. Rate constants of proteolytic kinetics under native conditions.

Protein	k_p (min^{-1})	k_p in presence of AdoMet (min^{-1})	$Adomet^+$ / $AdoMet^-$ ratio of k_p
wtCBS	0.026 ± 0.005	0.056 ± 0.005	2.15
P49L	0.025 ± 0.001	0.057 ± 0.005	2.28
P78R	0.0097 ± 0.001	0.031 ± 0.002	3.20
A114V	0.054 ± 0.011	0.5 ± 0.08	9.26
R125Q	0.87 ± 0.18	1.105 ± 0.008	1.27
E176K	0.97 ± 0.18	1.19 ± 0.18	1.22
R266K	0.014 ± 0.001	0.041 ± 0.02	2.93
P422L	0.036 ± 0.003	0.041 ± 0.001	1.14
I435T	N.D.	N.D.	N.D.
S466L	0.109 ± 0.008	0.063 ± 0.007	0.58

The values were determined using non-linear data fitting into a single exponential equation.

Table 2. The midpoint values of urea-induced unfolding (c_m) for CBS mutants in the absence and in the presence of AdoMet.

Protein	c_m in the absence of AdoMet (M)	c_m in the presence of AdoMet (M)
wtCBS	2.7 ± 0.08	3.26 ± 0.08
P49L	3.31 ± 0.13	3.40 ± 0.15
P78R	3.39 ± 0.06	3.42 ± 0.05
A114V	2.94 ± 0.20	3.46 ± 0.13
R125Q	N.D.	N.D.
E176K	N.D.	N.D.
R266K	3.08 ± 0.10	2.62 ± 0.11
P422L	3.85 ± 0.12	4.00 ± 0.09
I435T	N.A.	N.A.
S466L	3.39 ± 0.18	3.18 ± 0.12

The values were determined using pulse proteolysis followed by nonlinear data fitting into a sigmoidal equation. N.D. - not determined due to non-sigmoidal behavior of the mutant proteins (see Fig.S3 in the Supplementary Information). N.A. - not applicable; pulse proteolysis could not have been used due to a rapid cleavage in the absence of urea

FIGURE CAPTIONS

Figure 1. Model of hCBS. Model of the full-length human CBS based on the crystal structure of the dimeric Drosophila melanogaster CBS (PDB: 3PC2). The CBS cofactors (heme and PLP) are displayed as sticks. The arrows are pointing to the mutated residues (displayed as scaled balls and sticks) in one of the subunit of WT CBS. The analyzed pathogenic mutations are located in the proximity of the heme-binding pocket (P49L, R125Q and R266K), at the dimer interface (P78R, A114V and E176K) and in the regulatory AdoMet-binding domain (P422L, I435T and S466L).

Figure 2. Fluorescence spectroscopy of CBS proteins. A-C Emission spectra of tryptophan residues (excitation at $\lambda = 295$ nm); D-F Emission spectra of internal aldimines (excitation at $\lambda = 420$ nm).

Figure 3. Second-derivative UV spectroscopy: peak position for aromatic amino acids (Phe, Tyr and Trp) as a function of temperature. The exact peaks for a particular aromatic amino acid were determined by using 2nd derivative UV spectra recorded every 2.5°C for temperatures from 10°C to 90°C (see Supplementary Information for raw spectra) and subsequently plotted as a function of temperature. The wild type enzyme (WT, open squares) is separately compared with P49L, A114V, R125Q and E176K CBS mutants (filled circles). The P49L represent the CBS mutants that were similar to wild-type enzyme. On the contrary, A114V, R125Q and E176K exhibited blue shifts for tyrosines and tryptophans.

Figure 4. Proteolysis of CBS mutants with thermolysin under native conditions. A.Comparison of wild-type CBS with mutant proteins. Each point represents a mean from at least two independent experiments. B. Representative gels depicting proteolytic cleavage of selected mutants. P78R represents the proteolytically resistant mutants, while R125Q belongs to the more rapidly cleaved proteins. The mutant I435T is rapidly cleaved with the formation of the major fragment of molecular weight ≈ 40 kDa. “M“ refers to molecular weight marker, “N“ refers to the uncleaved control sample.

FIGURES

Figure 1

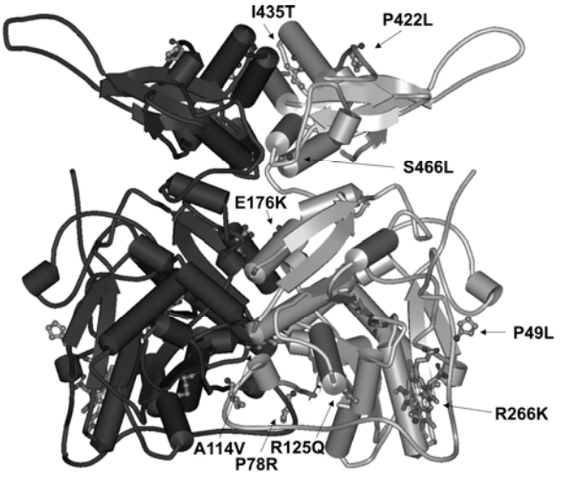


Figure 2

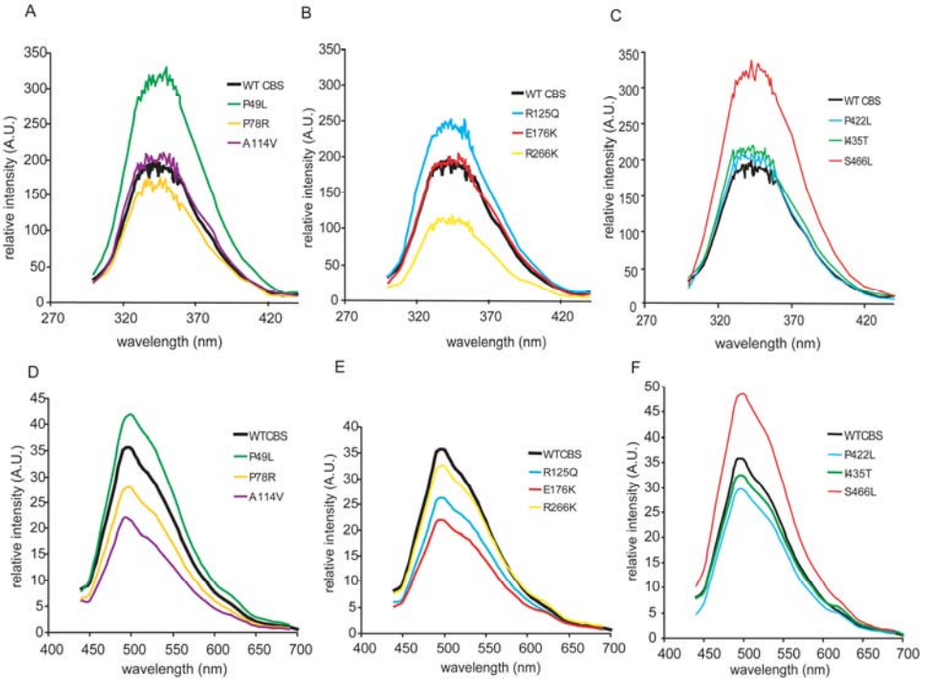


Figure 3

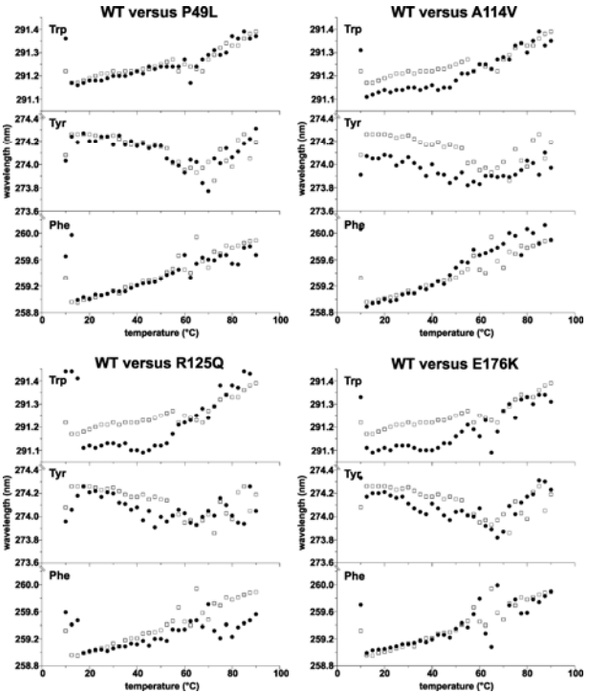
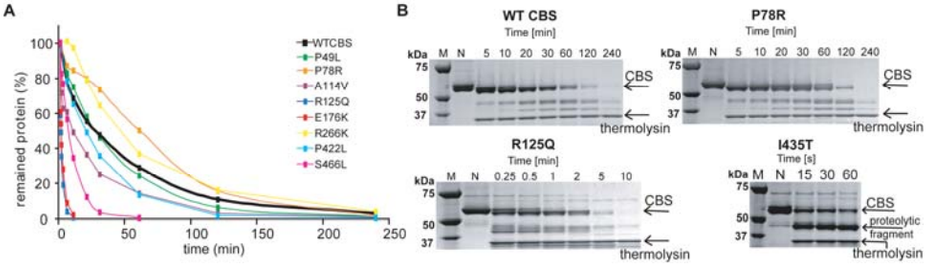
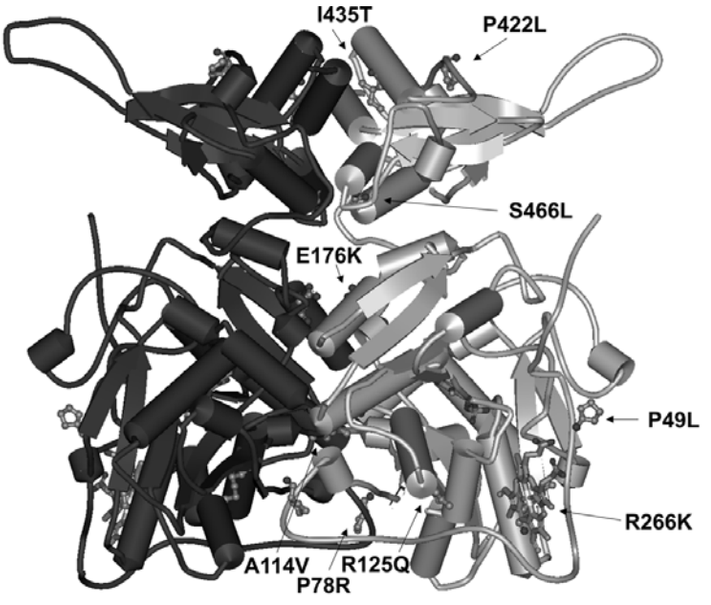


Figure 4



TOC figure



SUPPORTING INFORMATION

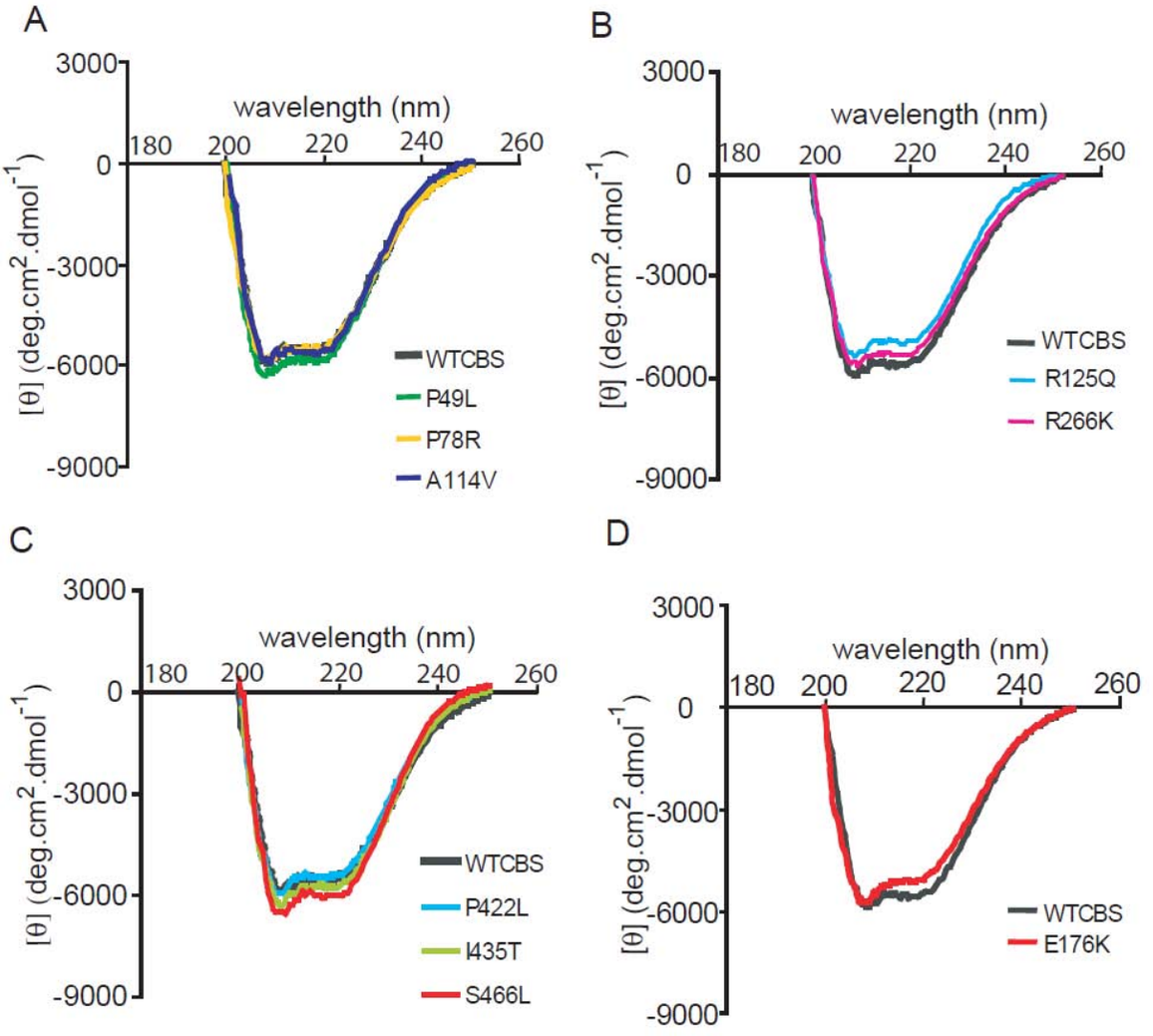


Figure S1: CD spectroscopy of CBS proteins. A-C CD spectra reveal that mutant proteins retain helical structure. D. Spectrum of E176D indicates partial decrease in the amount of helices together with the formation of a different secondary structure.

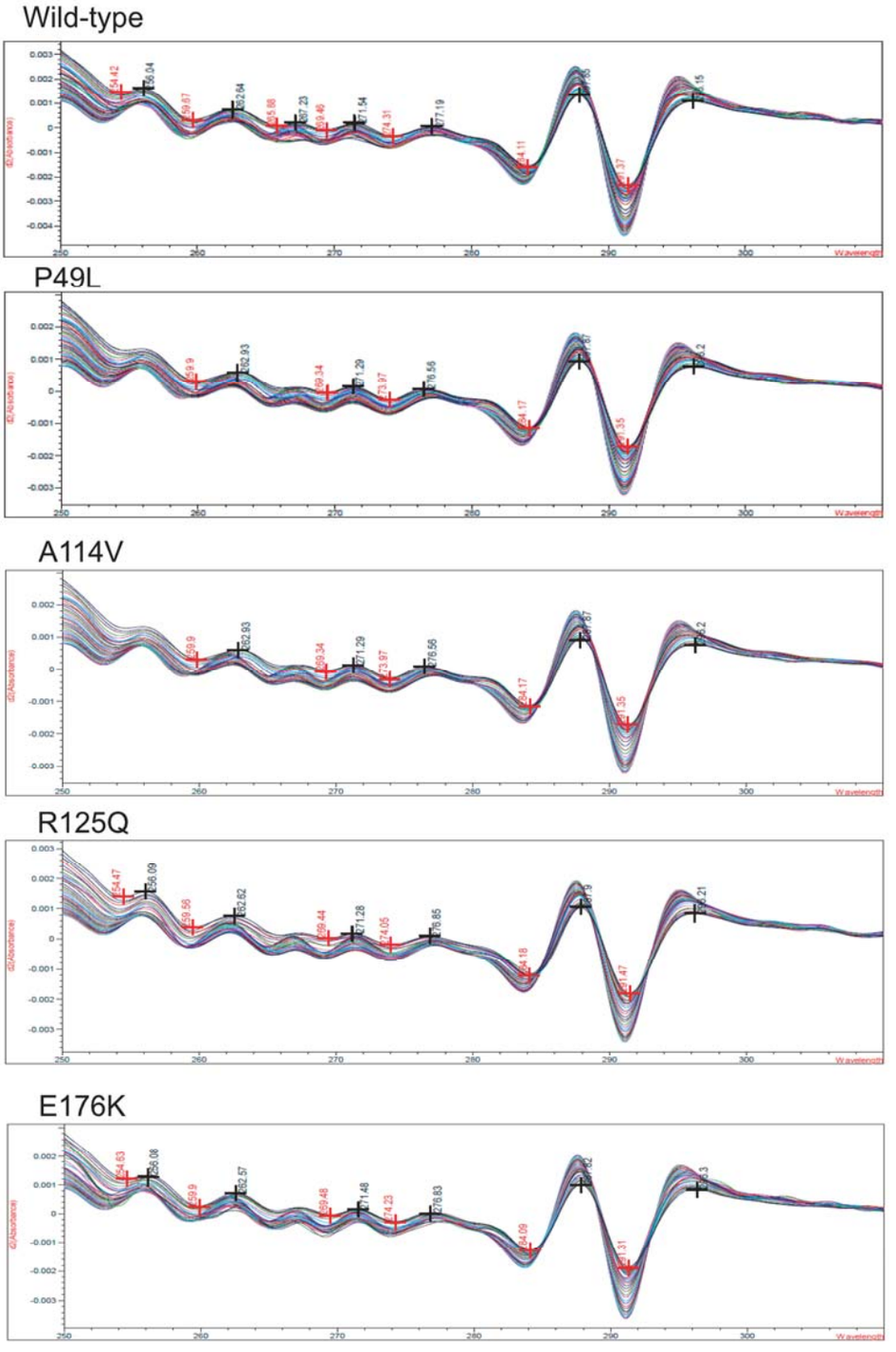


Figure S2: Second-derivative UV spectroscopy: spectral traces of wild-type CBS and representative mutant proteins for temperature gradient from 10 to 90 °C. The crosses along with wavelength values indicate the assigned maxima at spectra of proteins recorded at 90 °C.

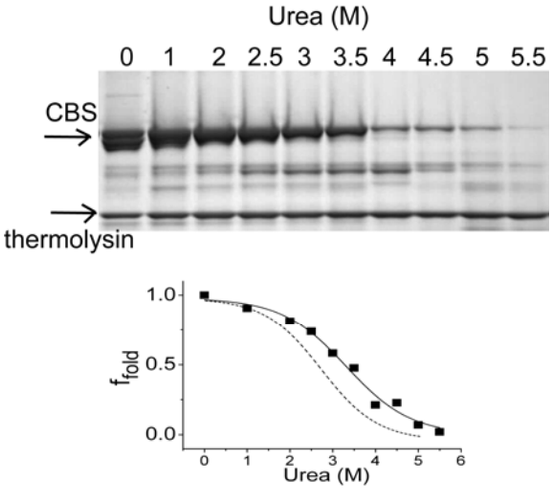


Figure S3. Representative data for pulse proteolysis in a urea gradient shows behaviour of the P49L mutant. All the mutants, similarly to P49L , exhibited higher c_m than the wild-type. Below the representative gel for P49L, corresponding plot is shown. Solid and dashed line represents curve for P49L and wild-type enzyme, respectively. F_{fold} is a fraction of folded protein that remained uncleaved after the proteolytic pulse. Points in the plot show means from at least two independent experiments.

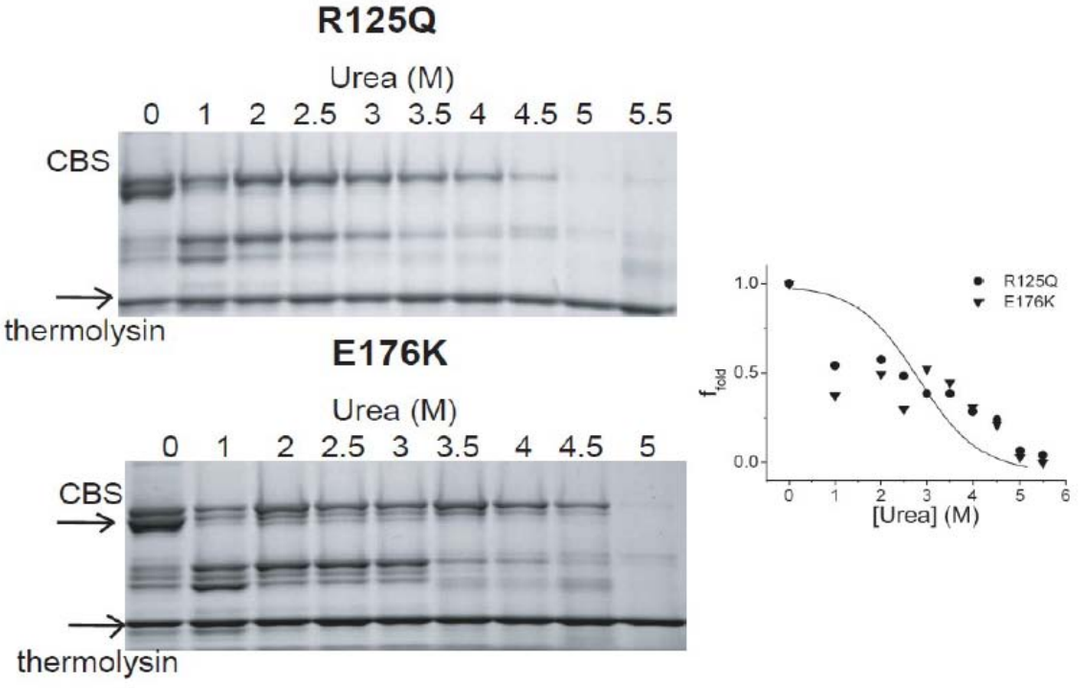


Figure S4: Pulse proteolysis of R125Q and E176K. In the plot, solid line represents the curve for wild-type CBS, circles and triangles refers to R125Q and E176K, respectively. Data points of the mutants could not be fitted to the sigmoidal equation.

Publication 5.1.6

J. Krijt et al.

Determination of cystathionine beta-synthase activity in human plasma by LC-MS/MS: potential use in diagnosis of CBS deficiency

Journal of Inherited Metabolic Diseasse, 2011

J Inherit Metab Dis (2011) 34:49–55
DOI 10.1007/s10545-010-9178-3

Determination of cystathionine beta-synthase activity in human plasma by LC-MS/MS: potential use in diagnosis of CBS deficiency

Jakub Krijt · Jana Kopecká · Aleš Hnízda · Stuart Moat · Leo A. J. Kluijtmans · Philip Mayne · Viktor Kožich

Received: 23 February 2010 / Revised: 4 June 2010 / Accepted: 19 July 2010 / Published online: 7 September 2010

Abstract Cystathionine β-synthase (CBS) deficiency is usually confirmed by assaying the enzyme activity in cultured skin fibroblasts. We investigated whether CBS is present in human plasma and whether determination of its activity in plasma could be used for diagnostic purposes. We developed an assay to measure CBS activity in 20 μL of plasma using a stable isotope substrate - 2,3,3-^{2}H serine. The activity was determined by measurement of the product of enzyme reaction, 3,3-^{2}H-cystathionine, using LC-MS/MS. The median enzyme activity in control plasma samples was 404 nmol/h/L (range 66–1,066; n=57). In pyridoxine nonresponsive CBS deficient patients, the median plasma activity was 0 nmol/ho/L (range 0–9; n=26), while in pyridoxine responsive patients the median activity was 16 nmol/hour/L (range 0–358; n=28); this overlapped with the enzyme activity from control subject. The presence of CBS in human plasma was confirmed by an in silico search of the proteome database, and was further evidenced by the activation of CBS by S-adenosyl-L-methionine and pyridoxal 5′-phosphate, and by configuration of the detected reaction product, 3,3-^{2}H-cystathionine, which was in agreement with the previously observed CBS reaction mechanism. We hypothesize that the CBS enzyme in plasma originates from liver cells, as the plasma CBS activities in patients with elevated liver aminotransferase activities were more than 30-fold increased. In this study, we have demonstrated that CBS is present in human plasma and that its catalytic activity is detectable by LC-MS/MS. CBS assay in human plasma brings new possibilities in the diagnosis of pyridoxine nonresponsive CBS deficiency.

Communicated by: Cornelis Jakobs

References to electronic databases:
http://www.chem.qmul.ac.uk/iubmb/enzyme/EC4/2/1/22.html;
http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=236200;
http://www.proteinatlas.org

Competing interest: None declared.

Presented at the 7th International Conference on Homocysteine Metabolism, Prague, 21-25 June 2009

Electronic supplementary material The online version of this article (doi:10.1007/s10545-010-9178-3) contains supplementary material, which is available to authorized users.

J. Krijt · J. Kopecká · A. Hnízda · V. Kožich (✉)
Institute of Inherited Metabolic Disorders—1st Faculty of Medicine, Charles University in Prague,
Ke Karlovu 2, 128 08,
Praha 2, Czech Republic
e-mail: Viktor.Kozich@LF1.cuni.cz

S. Moat
Department of Medical Biochemistry and Immunology,
University Hospital of Wales,
Cardiff CF14 4XW, UK

L. A. J. Kluijtmans
Department of Laboratory Medicine, Laboratory of Genetic, Endocrine and Metabolic Diseases,
Radboud University Nijmegen Medical Centre,
Nijmegen, the Netherlands

P. Mayne
National Newborn Screening Laboratory,
Children's University Hospital,
Temple St,
Dublin 1, Ireland

Abbreviations

ALT	Alanine aminotransferase
p-AMS	Pancreatic amylase
AST	Aspartate aminotransferase
CBS	Cystathionine β-synthase
IEM	Inborn errors of metabolism
GGT	γ-glutamyl transferase
PLP	Pyridoxal 5′-phosphate
SAM	S-adenosyl-L-methionine

Introduction

Cystathionine β-synthase (CBS; EC 4.2.1.22) (http://www.chem.qmul.ac.uk/iubmb/enzyme/EC4/2/1/22.html) deficiency (http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=236200) is a well-known genetic disease affecting the first step in the conversion of homocysteine to cysteine and ultimately to inorganic sulfur. Although this disorder was originally considered a severe multisystem disease that affected connective tissue and central nervous and vascular systems (Mudd et al. 1985), recent reports suggest that an unknown proportion of patients with CBS deficiency may suffer from only a mild vascular form of the disease or may be asymptomatic (Gaustadnes et al. 2000; Skovby et al. 2010). Diagnostic hallmarks of this disease are grossly elevated concentrations of plasma total homocysteine combined with decreased plasma concentrations of cysteine and cystathionine (Stabler et al. 1993) with varying elevations of plasma methionine levels. The DNA analysis may be used to confirm the diagnosis only if mutations with known pathogenicity are found at both patient CBS alleles; its utility is limited in other situations.

Demonstration of decreased enzyme activity is a common diagnostic approach in patients suffering from inborn errors of metabolism (IEM) including CBS deficiency. Since the majority of enzymes relevant for IEMs are located intracellularly, the determination of enzymatic activity requires in most cases sampling of patient tissue ranging from simple venepuncture to biopsies of skin or organs, such as the liver. In addition, laborious, time-consuming and technically demanding culturing of tissue such as skin may be required to produce sufficient amounts of cells to express the enzyme of interest. However, these intracellular enzymes can be released in small amounts into plasma reflecting turnover of the respective tissue(s). The amounts of these non-plasma-specific enzymes in body fluids may increase if cell integrity is impaired; clinical chemistry laboratories utilize this phenomenon and routinely analyze activities of liver-, pancreatic-, muscle-, myocardial-, prostatic- and other organ-specific enzymes as biomarkers of diseases (Bhagavan 2001). Determination of enzymes in extracellular fluids is only rarely used for routine diagnostic purposes in IEMs (e.g., biotinidase or total hexosaminidase).

In our study, we hypothesized that CBS may be released into plasma from organs expressing this enzyme and that its activity may be measurable by a sensitive LC-MS/MS assay employing deuterium-labeled substrates. To explore this hypothesis, we studied the CBS catalytic activity of plasma and stimulatory effect of CBS ligands, the presence of the CBS protein by immunological and by an in silico search, and its possible release to plasma from liver and other organs. Finally, we assessed plasma CBS activity in samples from patients with confirmed CBS deficiency to assess the diagnostic utility of this assay.

Materials and methods

Chemicals

All chemicals, if not stated otherwise, were obtained from Sigma–Aldrich (Prague)

Plasma samples

Archived anonymized plasma samples from controls and patients with CBS deficiency were obtained from the authors' repositories in the Netherlands, UK, Ireland and the Czech Republic. With the exception of 2 individuals, all patients were treated with pyridoxine at the time of sampling, regardless of their pyridoxine responsiveness (for details see Supplementary Table).

Anonymized samples from individuals with presumed liver, pancreatic and kidney disease (i.e., with elevated plasma levels of alanine aminotransferase, α-amylase and creatinine) were obtained as remnant samples after routine biochemical investigations from the Institute of Clinical Biochemistry and Laboratory Diagnostics of the General Faculty Hospital and The First Medical Faculty of Charles University. All experiments performed using human plasma samples were approved by the Ethics Committee of the General University Hospital, Prague.

Apparatus

The LC-MS/MS system consisted of an Agilent 1100 Series LC System (Agilent Technologies, Palo Alto, CA, USA) coupled with API 3200 triple quadrupole mass spectrometer with electron ion source and operated with Analyst software, Vision 1.4 (Applied Biosystems, Foster City, CA, USA).

Sample preparation

To 25 μL of solution containing 200 mmol/L Tris-HCl (pH 8.6), 1 mmol/L pyridoxal 5′-phosphate and 40 mmol/L 2,3,3-^{2}H-labeled serine (Cambridge Isotope Laboratories, USA), 20 μL of plasma or serum was added. The assay was initiated by addition of 5 μL starting solution containing 280 mmol/L homocysteine thiolactone in 100 mmol/L Tris-HCl (pH 8.6), 10 mmol/L DTT, 1 mol/L HCl and 1.225 mol/L NaOH. This starting solution was prepared freshly before sample processing and was preincubated for 5 min at 37°C after alkalinization with NaOH to allow cleavage of the thiolactone ring. For the measurement of

the activation of CBS by S-adenosyl-L-methionine, the final assay mixture contained 0.5 mmol/L SAM. The final assay mixture (total volume 50 μL) was incubated at 37°C for 4 h or was stopped immediately if a blank was processed. The reaction was stopped by acidification of the reaction mixture to pH 1-2 with 100 μL of Internal Standard Solution (EZ:faast; Phenomenex, Torrance, USA) containing 3.3 μmol/L of internal standard 3,3,4,4-^{2}H-labeled cystathionine (CDN Isotopes,Quebec, Canada).

LC-MS/MS analysis

The CBS activity was determined by the LC-MS/MS measurement of the product of enzyme reaction, 3,3-^{2}H-cystathionine, using a commercially available kit for amino acid analysis (EZ:faast; Phenomenex). Sample preparation according to the manufacturer's instruction involved a solid phase extraction step, derivatization with propyl chloroformate, and an extraction into an organic solvent prior to LC-MS/MS analysis.

The LC-MS/MS analysis was performed on proprietary EZ:faast AAA-MS column (250×2.0 mm) using LC and MS settings described in EZ:faast user manual.

The retention time of cystathionine isotopes was 15.5 min; total analysis time including regeneration of the column was 25 min. Detection of the analytes was carried out using positive electrospray ionization technique and selected multiple reaction monitoring. The precursor → product transitions for 3,3-^{2}H-cystathionine (enzyme reaction product; m/z 481.3→230.3), cystathionine (calibration standards; m/z 479.3→230.3) and 3,3,4,4-^{2}H-labeled cystathionine (internal standard; m/z 483.3→234.3) were monitored.

Standards, calibration curves and activity calculation

Quantitation of the product of enzyme reaction 3,3-^{2}H-cystathionine, which is not commercially available, was performed using cystathionine standard samples in the concentration range 0–5 μmol/L; the individual calibration points were 0, 0.25, 1 and 5 μmol/L. The calibration curve was obtained by linear regression; the peak area ratio (analyte/internal standard) was plotted against the analyte concentration. For the activity calculations, the measurable concentration of 3,3-^{2}H-cystathionine in the blank sample (t=0) was subtracted from the concentration produced by the enzyme reaction after incubation for 4 h.

Results

Human plasma exhibits CBS activity

With the employment of LC-MS/MS technology, we introduced a modification of an enzyme assay previously published (Janosik et al. 2009). We observed that incubation of control plasma samples according to the protocol described in "Materials and methods" resulted in the formation of 3,3-^{2}H-cystathionine. The signals of the product of the enzyme reaction, 3,3-^{2}H-cystathionine, determined in the assay mixture after incubation of control plasma and plasma from a CBS-deficient patient for 4 h are shown in Fig. 1a, b, respectively. The signals of 3,3-^{2}H-cystathionine in the blank sample (control plasma, incubation time = 0 h) and internal standard are shown in Fig. 1c, d, respectively. The figures show a clearly higher signal of 3,3-^{2}H-cystathionine determined in the assay using control plasma sample compared to the signals obtained by analysis of assay mixture with plasma from the CBS-deficient patient. The appearance of the signal of 3,3-^{2}H-cystathionine in the blank sample could be partially explained by isotope contribution of unlabelled cystathionine presented in the analyzed sample to the signal of 3,3-^{2}H-cystathionine. Nevertheless, despite the occurrence of the product signal in the blank sample, the assay performance allowed sufficient discrimination between plasma samples with normal and low or undetectable activities.

In order to confirm that the 3,3-^{2}H-cystathionine formation was due to the presence of CBS in plasma, to elucidate the source of the enzyme in circulation and to validate the enzyme assay we performed experiments described in the following section.

Demonstration of CBS protein in human plasma by immunodetection and in silico searches

To demonstrate the presence of CBS antigen in plasma, we first performed western blotting after immunoprecipitation preceded by low abundant plasma protein enrichment. Multiple fractions including bands of approximately 60 and 45 kDa which were consistent with the presence of full-length and truncated forms of CBS were observed (data not shown). To confirm the identity of these fractions, we used gel digestion followed by peptide mass fingerprinting with MALDI-TOF MS detection; only abundant serum proteins such as albumin were identified with no peptide related to CBS being detected (data not shown). These data suggested that the fractions observed on western blots were due to the presence of cross-reactive highly abundant plasma proteins. In addition, we were unable to detect the CBS antigen by ELISA as this in-house assay had poor sensitivity (see Discussion).

As an alternative approach to demonstrate the presence of CBS in human plasma, we searched the available proteomic database Peptide Atlas (www.peptideatlas.org). As of February 1, 2010, the search only identified the presence of CBS peptides 2–18 and 414–439 in an

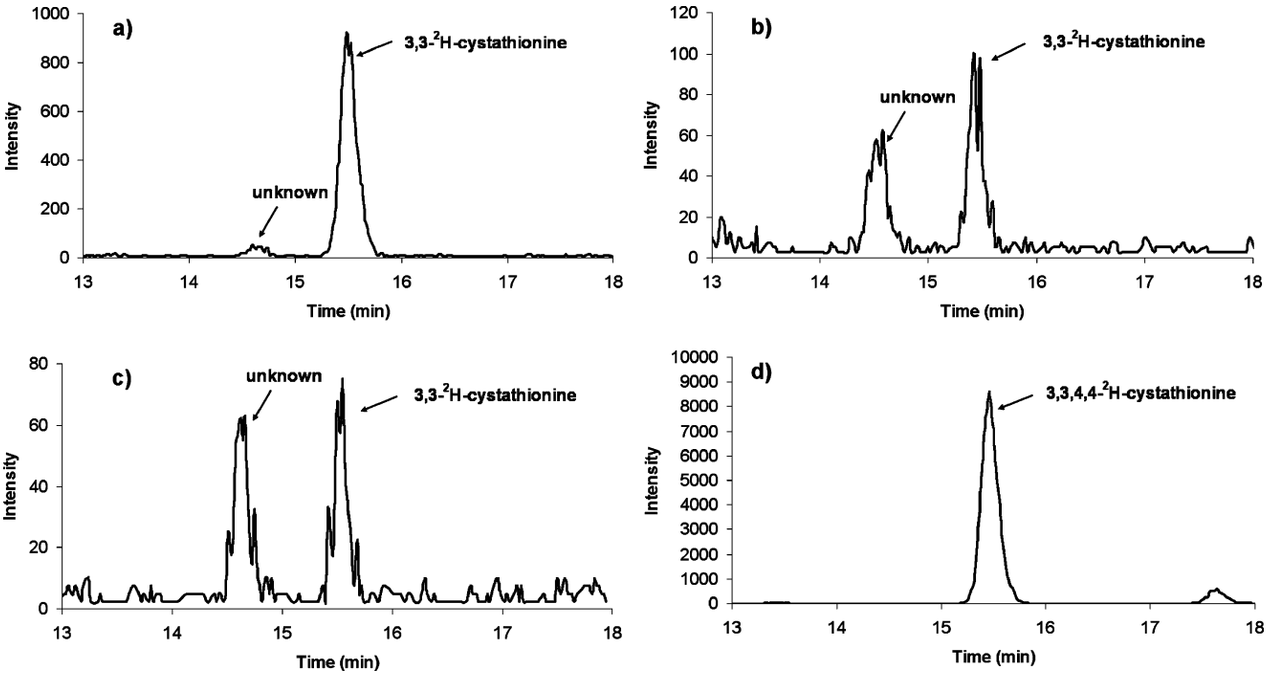


Fig. 1 The LC-MS/MS chromatograms. Signal of 3,3-^{2}H-cystathionine (retention time 15.4–15.5 min) was determined after incubation of the assay mixture with control plasma (**a** incubation time 4 h, **c** incubation time 0 h, sample blank), and a plasma from a CBS-deficient patient (**b** incubation time 4 h). **d** The signal of the internal standard 3,3,4,4-^{2}H-labeled cystathionine. The peak with retention time 14.5 min belongs to an unknown compound, which is not a product of the CBS-catalyzed reaction

experiment that employed cysteinyl-peptides enrichment (Liu et al. 2004), demonstrating that CBS is indeed present in normal human plasma but in only very low quantities.

Responsiveness of plasma CBS activity to S-adenosyl-L-methionine and pyridoxal 5′-phosphate

Since S-adenosyl-L-methionine (SAM) is an allosteric activator and pyridoxal 5′-phosphate (PLP) is a cofactor of human CBS, we tested whether the addition of these compounds into the assay mixture would increase the production of 3,3-^{2}H-cystathionine.

We performed the assays in plasma samples (n=6) according to the protocol described in "Materials and methods" with and without the addition of SAM (final concentration in the assay was 0.5 mmol/L) to the assay mixture. We observed that the addition of SAM to the assay mixture increased the production of 3,3-^{2}H-cystathionine by a factor of 2.2 ± 0.2. Similarly, the assay was performed using control plasma samples (n=6) with and without PLP added to the assay mixture. The production of 3,3-^{2}H-cystathionine in the assay with PLP was 1.6 ± 0.1 times higher than without the cofactor added to the assay. The stimulatory effect of both compounds on the production of 3,3-^{2}H-cystathionine in the assay gives further support to the hypothesis that human plasma contains CBS activity.

Origin of CBS in human plasma

To explore the hypothesis that CBS is released into plasma from organs with a high content of the enzyme, we determined the enzyme activity in plasma samples obtained from individuals with elevated biomarkers for liver (alanine aminotransferase, ALT and aspartate aminotransferase, AST), pancreatic (pancreatic amylase, p-AMS) and kidney (creatinine) disease. The results, shown in the Supplementary Table, clearly demonstrated an extremely elevated CBS activity (22463-231179 nmol/h/L) in plasma from individuals with increased ALT (2.4-15 μkat /L) and AST (2.1-14 μkat/L). Elevation of CBS activity was present also in the majority of plasma samples from individuals with elevated creatinine and p-AMS, respectively, although the plasma CBS activity increase was less striking than in patients with elevated plasma aminotransferase activities. These data are congruent with the hypothesis that CBS may be released into plasma from organs with a high content of this enzyme, mainly from liver.

Method validation

Assay linearity To determine the linearity of the assay we incubated the assay mixture for different time intervals (1, 2, 4 and 6 h). In addition, three experiments assays were performed with the addition of different plasma volumes to

the assay mixture (5, 10 and 20 μL, respectively) and with an incubation time 4 h. Product formation was measured and plotted against different time points or varying plasma volumes, respectively. In both experiments, we observed linear relationships with correlation coefficients r^2=0.997 and r^2=0.999, respectively. The linearity of the time and dose relationships supports the presumption that the monitored compound, 3,3-^{2}H-cystathionine, is a product of the enzyme reaction.

Intra- and Inter-day variation of the assay The assay reproducibility was assessed by determination of intra- and inter-day coefficients of variation of the assay. For determination of the intra-day variation, the assay was performed with 5 identical control plasma samples in 1 day. For determination of the inter-day variation, the assays were performed using the control plasma on six separate days over an 11-month period. The mean of the activities were determined and coefficients of variation (in parentheses) for intra- and inter-day experiments were 354 nmol/h/L (1.4%) and 309 nmol/h/L (17 %), respectively. These data show that reproducibility of this assay was satisfactory and was consistent with the usually observed performance of enzyme assays.

Sample stability Sample stability was assessed by assaying control plasma immediately following venepuncture and plasma separation, and after one and ten cycles of freezing and thawing. In addition, samples were also tested following incubation of plasma samples stored at 25°C for 6 h. No change in activity was seen after one freeze and thaw cycle. However, after ten cycles, a decrease of 13% in activity was observed. A similar decrease in activity (i.e., 11%) was also observed after incubation of the plasma for 6 h at room temperature. The activity changes resulting from freezing and thawing or sample storage at 25°C may have contributed to the significantly higher inter-day coefficient of variation.

Sample matrix effects The validation experiments described were carried out on plasma with EDTA used as anticoagulant because the majority of archived samples from CBS-deficient patients available for this study were prepared with EDTA. However, to explore the influence of other sample matrices on CBS activity, we performed the assay using plasma (from one person) following blood being collected into EDTA- and heparin-containing tubes and serum collection tubes. The results showed similar activities in heparinized plasma samples and serum, whereas the activity in EDTA-treated plasma was substantially lower (~22%) compared to the activity determined in the other two matrices. In addition, we evaluated the effect of haemolysis on CBS activity in plasma samples from a control subject and from two CBS-deficient patients with low (10 nmol/h/L) and undetectable enzyme activity, respectively. Haemolysis in one aliquot of EDTA blood was induced by a single cycle of freezing and thawing before centrifugation. Haemolysis had no measurable effect on plasma CBS activity.

CBS activities in plasma obtained from controls and CBS deficient patients

To assess the clinical utility of the plasma CBS assay in diagnosing patients with CBS deficiency, we measured activities in plasma samples obtained from control subjects, CBS-deficient patients and their parents. These samples were obtained from four European countries (see Supplementary Table and Fig. 2). The median plasma CBS enzyme activity from control subjects (n=57) was 404 nmol/hour/L, the median plasma enzyme activity from pyridoxine nonresponsive CBS deficient patients (n=26) was 0 nmol/hour/L. The median plasma enzyme activity from pyridoxine responsive patients (n=28) was 16 nmol/h/L, higher than in the pyridoxine nonresponsive patients, but compared to the median of control subjects, it represented just 4% of the activity The median enzyme activities in the heterozygote parents (n=9) was 547 nmol/h/L, showing no utility of the assay to recognize heterozygotes.

Because of the elevation of plasma CBS activity observed in samples from individuals with presumed liver disease (see "Origin of CBS in human plasma") the most stable biomarker for the liver disease, γ-glutamyl transferase (GGT), was determined in archived samples analyzed in this study. The samples with elevated levels of GGT (higher than 0.8 μkat/L) showed increased CBS activity, i.e. higher than 1,300 nmol/h/L, and were therefore excluded from the control dataset. One pyridoxine responsive CBS deficient patient had an elevated plasma GGT (1.5 μkat/L) and CBS activity (770 nmol/h/L) and was also excluded from the presented dataset.

Discussion

Utilization of mass spectrometry for determination of CBS activity

The introduction of novel analytical techniques into routine diagnostic laboratories in the last decade has opened up the field for new approaches in enzyme assay methodology. Tandem mass-spectrometry in enzymology allows the use of substrates that are labeled with stable isotopes instead of radioisotopes; its utility has been discussed elsewhere (Gelb et al. 2006). In this study, we employed LC-MS/MS to design a novel CBS assay, utilizing isotopically labeled

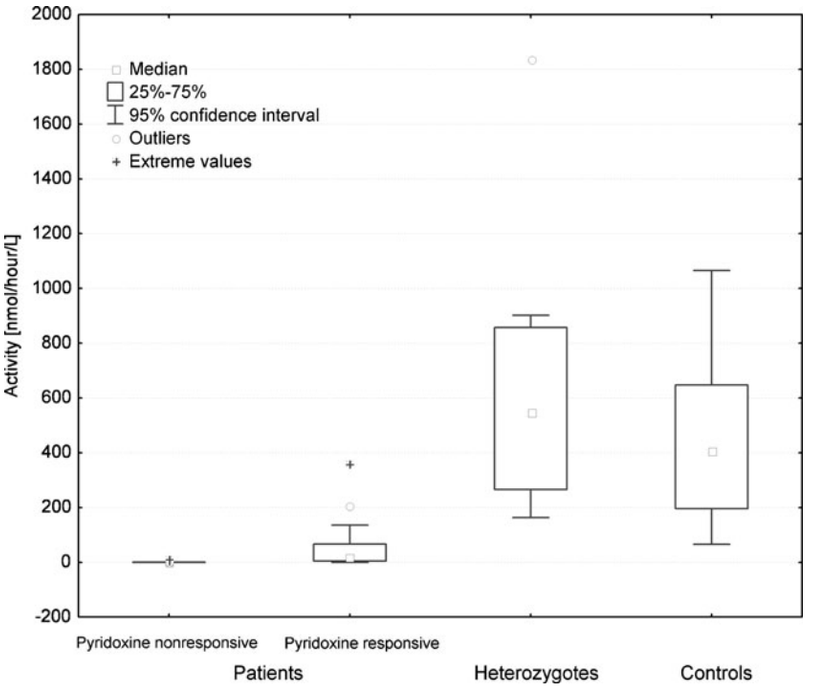


Fig. 2 CBS activities in plasma samples. The graph shows distribution of CBS activities determined in plasma samples obtained from pyridoxine nonresponsive CBS deficient patients (n=26), pyridoxine responsive CBS deficient patients (n=28), heterozygotes for CBS deficiency (n=9) and controls (n=57)

standard, 2,3,3-^{2}H- serine. This method has excellent sensitivity and can detect nanomolar concentrations of the reaction product, 3,3-^{2}H-cystathionine.

Compared to the previously described radiometric assay using ^{14}C serine with a detection limit in the low micromolar range (Kopecka et al. 2010), the LC-MS/MS method is at least 2 orders of magnitude more sensitive and enables the determination of very low catalytic activities of CBS in human plasma, which would be undetectable by the established radiometric method. Using the radiometric method, we were able to detect product of the CBS activity, radioactive cystathionine, only in the sample from a patient with the most severe liver disease (data not shown).

Is there any measurable CBS activity present in human plasma?

Although CBS is an intracellular cytosolic enzyme, the breakdown of tissue containing high amounts of this enzyme may result in leakage of CBS protein into plasma, a phenomenon commonly observed in the case of other liver enzymes that are routinely measured in human plasma or serum (e.g., ALT, AST, GGT and others). Our study has shown that control human plasma is able to catalyze the formation of cystathionine with a median activity (404 nmol/h/L, i.e. 125 pkat/L) which is three to five orders of magnitude lower than activities of serum enzymes routinely used in clinical chemistry. Several lines of evidence demonstrate that this activity can be attributed to the presence of CBS in human plasma: (1) the absence of activities in plasma samples from pyridoxine nonresponsive CBS deficient patients and its presence in controls; (2) activation of plasma CBS by SAM and PLP by a factor of 2.2 and 1.6, respectively; (3) configuration of the detected reaction product, 3,3-^{2}H-cystathionine, is in agreement with the previously observed CBS reaction mechanism (Banerjee and Zou 2005; Borcsok and Abeles 1982; Miles and Kraus 2004); this mechanism involves the formation of aminoacrylate intermediate with release of water containing one deuterium atom bound originally on the alpha-carbon of serine; and (4) the presence of CBS protein in normal human plasma was demonstrated by in silico search of a proteome database. Considering the specific activity of purified CBS protein (150 units/mg protein) and the plasma CBS activity in controls, we estimated that the quantity of CBS in human plasma is approximately in the range of nanograms per millilitre which does not allow the confirmation of the CBS presence in human plasma by western blotting.

Plasma CBS activity in CBS-deficient patients – utilization of the assay in diagnosis and study of CBS deficiency

In our study, we measured CBS activities in plasma from 54 patients with CBS deficiency and 57 control subjects. The range of activities in the plasma from pyridoxine nonresponsive CBS-deficient patients was clearly separated from the range obtained in plasma samples from control subjects. In contrast, the range of activities in plasma from pyridoxine responsive CBS-deficient patients overlaps with those from the control subjects.

To evaluate the use of the CBS assay in plasma for the diagnosis of CBS deficiency, we calculated the specificity and sensitivity of the assay. Since the lowest CBS activity in samples from control subjects was 67 nmol/h/L, this value was used as a threshold value for differentiating patients from controls. The sensitivity and specificity of the assay was excellent for pyridoxine nonresponders (i.e., sensitivity = 100% and specificity = 100%), while it was suboptimal for pyridoxine responders (i.e., sensitivity = 70% and specificity = 100%). These data suggest that the plasma CBS assay may be useful in detecting pyridoxine non-responders, while its use for diagnosing responders may be limited.

The reason for the wide range of plasma CBS activities observed in pyridoxine responders (0-358 nmol/h/L) is unknown. This variability is caused not only by the underlying genetic defect but may result from a contribution of additional and yet unknown factors (as demonstrated by a wide range of plasma CBS activities in patients homozygous for the c.833T>C mutation; see Supplementary Table). It is at present unknown whether the CBS assay in plasma will be useful for predicting the phenotypic severity in pyridoxine responsive patients. The measurement of plasma CBS activity may be useful to evaluate or monitor the effectiveness of pyridoxine therapy. The latter hypothesis is supported by an experiment in which it was observed that plasma CBS activity in one patient increased from 4.5 to 55.9 nmol/h/L after the patient started pyridoxine therapy. However, in two other samples from patients responding to pyridoxine therapy, the activity was undetectable before as well as after the start of pyridoxine treatment (data not shown). Evaluation of the effects of pyridoxine therapy on plasma CBS activity would require further studies on patients before and on therapy.

To our knowledge, this is the first report demonstrating the presence of CBS activity in human plasma. This proof-of-principle study shows that very low catalytic activities of enzymes originating from liver can be measured in cell-free extracellular fluids. In summary, this paper highlights the possibility of measuring intracellular enzymes in plasma or serum using isotopically labeled substrates and LC-MS/MS instrumentation, and opens new possibilities for diagnosing other inborn errors of metabolism.

Acknowledgements This work was supported by grant MZ0VFN2005 from the Ministry of Health of the Czech Republic. Institutional support was provided by the Research Project of the Ministry of Education of the Czech Republic (reg. no. MSM0021620806). The authors would like to express their gratitude to Ms. Alena Dutá for technical help, to Jitka Sokolová,MSc. for assistance in preparing the manuscript, to Dr. Jana Barcalová for the measurements of the liver enzymes, and to Květa Pelinková, MSc. for providing the plasma samples from individuals with presumed liver, pancreatic and kidney disease.

References

Banerjee R, Zou CG (2005) Redox regulation and reaction mechanism of human cystathionine-beta-synthase: a PLP-dependent heme-sensor protein. Arch Biochem Biophys 433(1):144–156

Bhagavan NV (2001) Medical biochemistry, 4th edn. Academic, London

Borcsok E, Abeles RH (1982) Mechanism of action of cystathionine synthase. Arch Biochem Biophys 213(2):695–707

Gaustadnes M, Rudiger N, Rasmussen K, Ingerslev J (2000) Intermediate and severe hyperhomocysteinemia with thrombosis: a study of genetic determinants. Thromb Haemost 83(4):554–558

Gelb MH, Turecek F, Scott CR, Chamoles NA (2006) Direct multiplex assay of enzymes in dried blood spots by tandem mass spectrometry for the newborn screening of lysosomal storage disorders. J Inherit Metab Dis 29(2–3):397–404

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):431–437

Kopecka J, Krijt J, Rakova K, Kozich V (2010) Restoring assembly and activity of cystathionine beta-synthase mutants by ligands and chemical chaperones. J Inherit Metab Dis (in press)

Liu T, Qian WJ, Strittmatter EF, Camp DG 2nd, Anderson GA, Thrall BD, Smith RD (2004) High-throughput comparative proteome analysis using a quantitative cysteinyl-peptide enrichment technology. Anal Chem 76(18):5345–5353

Miles EW, Kraus JP (2004) Cystathionine beta-synthase: structure, function, regulation, and location of homocystinuria-causing mutations. J Biol Chem 279(29):29871–29874

Mudd SH, Skovby F, Levy HL, Pettigrew KD, Wilcken B, Pyeritz RE, Andria G, Boers GH, Bromberg IL, Cerone R et al (1985) The natural history of homocystinuria due to cystathionine beta-synthase deficiency. Am J Hum Genet 37(1):1–31

Skovby F, Gaustadnes M, Mudd SH (2010) A revisit to the natural history of homocystinuria due to cystathionine beta-synthase deficiency. Mol Genet Metab 99(1):1–3

Stabler SP, Lindenbaum J, Savage DG, Allen RH (1993) Elevation of serum cystathionine levels in patients with cobalamin and folate deficiency. Blood 81(12):3404–3413

Supplementary Table: Overview of analyzed patients

Phenotype	Patient	CBS activity (nmol/hour/L of plasma)	Sample origin	Genotype/Comment	Pyridoxine treatment
Pyridoxine responsive patients with CBS deficiency	1	0.0	UK	c.[430G>A]+[919G>A]	+
	2	0.0	UK	N.A.	+
	3	0.0	CZ	c.[833T>C]+[430G>A, 463G>A]	+
	4	0.0	UK	c.[430G>A]+[919G>A]	+
	5	1.8	UK	c.[833T>C]+?	+
	6	2,3	NL	c.[833T>C]+[373C>T]	+
	7	5.0	CZ	c.[833T>C]+[430G>A, 463G>A]	+
	8	5.1	NL	c.[833T>C]+[833T>C]	N.A.
	9	7.4	NL	c.[833T>C]+[833T>C]	+
	10	10.0	CZ	c.[833T>C]+[1126G>A]	+
	11	11.4	UK	c.[833T>C]+[833T>C]	+
	12	12.0	NL	c.[833T>C]+[833T>C]	+
	13	13.2	NL	c.[833T>C]+[833T>C]	-
	14	15.6	NL	c.[833T>C]+[833T>C]	+
	15	15.6	CZ	c.[833T>C]+[430G>A, 463G>A]	+
	16	17.3	NL	c.[833T>C]+[833T>C]	+
	17	24.6	NL	c.[833T>C]+[833T>C]	+
	18	28.0	CZ	c.[833T>C]+[IVS 11-2A>C]	+
	19	34.4	NL	c.[833T>C]+[833T>C]	+
	20	46.0	CZ	c.[833T>C]+[833T>C]	+
	21	62.6	CZ	c.[833T>C]+[833T>C]	+
	22	70.6	NL	c.[833T>C]+[833T>C]	+
	23	75.8	NL	c.[833T>C]+[833T>C]	-
	24	103.6	NL	c.[833T>C]+[833T>C]	+
	25	131.2	NL	c.[833T>C]+[833T>C]	+
	26	136.0	NL	c.[833T>C]+[833T>C]	+
	27	358.0	CZ	c.[341C>T]+[1226A>G]	+
	28	206.0	CZ	c.[833T>C]+[?]	+
	Median Mean Range S.D.	15.6 49,8 0-358 79			
Pyridoxine non-responsive patients with CBS deficiency	1	0.0	CZ	c.[494G>A]+[494G>A]	+
	2	0.0	CZ	c.[IVS 11-2A>C]+[IVS 7+1G>A]	+
	3	0.0	CZ	c.[IVS 11-2A>C]+[430G>A, 463G>A]	+
	4	0.0	CZ	c.[IVS 11-2A>C]+[1226A>G]	+
	5	0.0	IRE	c.[919G>A]+?	+
	6	0.0	IRE	?+?	+
	7	0.0	UK	?+?	N.A.
	8	0.0	UK	?+?	N.A.
	9	0.0	UK	?+?	+
	10	0.0	UK	?+?	+
	11	0.0	IRE	c.[919G>A]+[919G>A]	+
	12	0.0	IRE	c.[919G>A]+?	+
	13	0.0	IRE	c.[919G>A]+?	+
	14	0.0	IRE	?+?	+
	15	0.0	IRE	?+?	+
	16	0.0	IRE	c.[919G>A]+[919G>A]	+
	17	0.0	IRE	c.[919G>A]+?	+
	18	0.0	IRE	?+?	+
	19	0.0	IRE	c.[919G>A]+[919G>A]	+
	20	0.9	IRE	c.[919G>A]+[919G>A]	+
	21	2.8	IRE	c.[919G>A]+[919G>A]	+
	22	3.5	CZ	c.[IVS 1-1G>C]+[28delG]	+
	23	6.8	CZ	?+ c.[430G>A, 463G>A]	+
	24	8.0	CZ	c.[IVS 1-1G>C]+[28delG]	+
	25	9.0	CZ	c.[IVS 11-2A>C]+[IVS 11-2A>C]	+
	26	9.2	IRE	c.[919G>A]+[302T>C]	+
	Median Mean Range S.D.	0 1,6 0-9.3 3.1			
Controls	Median Mean Range S.D.	403,8 433,5 65.6-1065.6 256,7	CZ (n=30) NL (n=18) IRE (n=5) UK (n=4)		
Heterozygotes for CBS deficiency	Median Mean Range S.D.	547,4 671,8 163.6-1836.9 512,4	CZ (n=6) UK (n=3)		
Individuals with elevated serum aminotransferases	1	231178.8	CZ	ALT=15.0 µkat/L; AST=14.0 µkat/L	
	2	30531.2	CZ	ALT=2.6 µkat/L; AST=5.3 µkat/L	
	3	27350.0	CZ	ALT=2.4 µkat/L; AST=2.7 µkat/L	
	4	22462.5	CZ	ALT=2.5 µkat/L; AST=2.1 µkat/L	
Individuals with elevated serum creatinine	1	33787.5	CZ	creatinine=278 µmol/L	
	2	9195.0	CZ	creatinine=292 µmol/L	
	3	570.6	CZ	creatinine=221 µmol/L	
	4	67.6	CZ	creatinine=665 µmol/L	
	5	493.1	CZ	creatinine=326 µmol/L	
Individuals with elevated serum pancreatic amylase	1	17769.8	CZ	p-AMS=4.14 µkat/L	
	2	3106.3	CZ	p-AMS=13.9 µkat/L	
	3	18062.5	CZ	p-AMS=3.5 µkat/L	
	4	336.9	CZ	AMS=4.4 µkat/L	
	5	3425.0	CZ	p-AMS=2.7 µkat/L	
	6	1750.0	CZ	p-AMS=8.9 µkat/L	
	7	893.8	CZ	p-AMS=26.1 µkat/L	