

SCIENTIFIC REPORTS

OPEN

Heme interaction of the intrinsically disordered N-terminal peptide segment of human cystathionine- β -synthase

Amit Kumar¹, Amelie Wißbrock², Nishit Goradia¹, Peter Bellstedt³, Ramadurai Ramachandran¹, Diana Imhof² & Oliver Ohlenschläger¹

Cystathionine- β -synthase (CBS) belongs to a large family of pyridoxal 5'-phosphate (PLP)-dependent enzymes, responsible for the sulfur metabolism. The heme-dependent protein CBS is part of regulatory pathways also involving the gasotransmitter hydrogen sulfide. Malfunction of CBS can lead to pathologic conditions like cancer, cardiovascular and neurodegenerative disorders. Truncation of residues 1–40, absent in X-ray structures of CBS, reduces but does not abolish the activity of the enzyme. Here we report the NMR resonance assignment and heme interaction studies for the N-terminal peptide stretch of CBS. We present NMR-spectral evidence that residues 1–40 constitute an intrinsically disordered region in CBS and interact with heme *via* a cysteine-proline based motif.

Cystathionine- β -synthase (CBS) is a modular enzymatic protein of 551 amino acids. The lyase CBS acts in the transsulfuration pathway and has a central role in the mammalian sulfur metabolism by catalysing the condensation of serine and homocysteine to a cystathionine intermediate^{1,2}. Mutations in CBS are responsible for an inborn autosomal recessive inherited deviation in the amino acid metabolism, homocystinuria. This rare disease results in a clinical phenotype with elevated levels of homocysteine in blood plasma affecting four central organ systems, the cardiovascular, ocular, skeletal, and central nervous system and finally manifesting e.g. in neural tube defects, cardiovascular diseases and Alzheimer's disease^{3–5}. Currently, more than 150 pathogenic CBS mutations have been reported which are spread across the coding sequence^{6,7}. CBS is a heme-binding protein and represents an example for a human hemoprotein that can be regulated by gases *via* their heme component⁸. Especially, besides its binding CO and NO⁹, CBS is involved in cellular production of major amounts of the gasotransmitter H₂S¹⁰. It was shown that aberrant function of CBS may thus lead to cancer, cardiovascular and neurodegenerative diseases making it an interesting drug target^{11–13}.

The homotetrameric CBS, a member of the fold type II family of the pyridoxal 5'-phosphate (PLP)-dependent enzymes, is in large parts structurally well characterised^{14,15}. Members of this enzyme family are of multiple evolutionary origins. CBS contains the N-terminal heme-binding domain followed by catalytic core, which is conserved in most proteins of this class of organisms, and a C-terminal regulatory domain. The N-terminal domain (residues 1–70) is sequentially diverse in many classes of organisms and absent in lower eukaryotes like yeast and mycobacterium. The canonical heme in CBS is bound *via* an axial coordination to cysteine-52 and histidine-65 in the N-terminal region of the molecule. Interestingly, this binding site is missing in the homologous yeast enzyme^{14,16}. In contrast to covalent binding of heme, recent studies have shown that transient interactions of the heme moiety with proteins can act as functional triggers¹⁷. Here, heme-binding motifs (HBM) or heme regulatory motifs (HRM) with special amino acid combinations –among them the cysteine-proline (CP) motif– are suggested to be responsible for heme association^{17–22}. However, due to moderate binding constants these motifs at the same time allow also a fast dissociation of the heme complex. Interestingly, also in the newest X-ray structures of the human CBS protein^{23,24} (e.g. PDB code 4COO) the stretch from residue 516–526 and the first 42 N-terminal

¹Leibniz Institute on Aging – Fritz Lipmann Institute, Beutenbergstr. 11, D-07745, Jena, Germany. ²Pharmaceutical Biochemistry and Bioanalytics, Pharmaceutical Institute, University of Bonn, An der Immenburg 4, D-53121, Bonn, Germany. ³Friedrich Schiller University, Faculty of Chemistry and Earth Sciences, Humboldtstr. 10, D-07743, Jena, Germany. Correspondence and requests for materials should be addressed to D.I. (email: dimhof@uni-bonn.de) or O.O. (email: oliver.ohlenschlaeger@leibniz-fli.de)

<i>Ophiophagus_hannah</i>	MLCDVRGTMPTMP---SIPAETEMDTMTCTHVSVDCTI
<i>Xenopus_laevis</i>	-----MPAVPPTTDSAACPHIKENHVV
<i>Gallus_gallus</i>	----MEKLMSEKPVVSKAKPDTNSCPHASGKYF
<i>Bos_taurus</i>	-----MP-FETSPADTLCRSAGCPHLSGAHL
<i>Felis_catus</i>	-----MP-SETP--QAETGSAGCPHLSGAHL
<i>Macaca_fascicularis</i>	-----MP-SETP--QAEVGPTGCPHLSGPHS
<i>Homo_sapiens</i>	-----MP-SETP--QAEVGPTGCPHRSGPHS
<i>Pan_troglodytes</i>	-----MP-SETP--QAEVGPTGCPHRSGPHS
<i>Cricetulus_griseus</i>	-----MP-SGTS--QCEDGSARCPQPLEVHS
<i>Heterocephalus_glaber</i>	-----MP-SETP--QEESGASGCPH-----

Figure 1. Sequence alignment of the N-terminal stretches of CBS proteins from different organisms.

residues are missing. Exactly in the latter region a CP-motif is located which is followed by histidines at positions $P + 2$ and $P + 7$ relative to the cysteine. This region is structurally close to the canonical heme-binding site at Cys⁵²/His⁶⁵ which allows to speculate about a scavenger-type function of the CP-motif located in this intrinsically disordered protein region to support the localisation of the heme molecule during the binding process or to act as a second independent heme-binding site. In addition, it might be of interest to note that a sequence homology analysis reveals this CPH-motif to be conserved in higher eukaryotes with secondary interacting amino acid (H, C, Y) present in vicinity (Fig. 1).

Intrinsically disordered protein regions (IDPRs) are increasingly recognised to have functional relevance for biological regulatory processes in e.g. cellular signal transduction, molecular recognition and transcription^{21,25,26} making them an interesting target for drug intervention^{27,28}. Bioinformatic predictions (e.g. with DISOPRED²⁹) estimate that disordered regions or segments longer than 30 residues are present in approx. 2.0% of archaean, 4.2% of eubacterial and 33.0% of eukaryotic proteins³⁰ and that IDPRs are highly abundant in disease-related proteins³¹. Meanwhile the aim to structurally characterise IDPRs has fostered the development of various approaches e.g. in the areas of computational prediction^{30,32,33} as well as NMR spectroscopy^{34–43}. Interestingly, CP-motifs are also found in intrinsically disordered regions of functional relevance in other proteins that are associated with heme binding²¹.

In this context, we have expressed the CBS protein using established approaches and investigated the heme-binding domain separately from the catalytic core. Functional studies on CBS(1–413) and on mutations of the non-canonical heme-binding site (C15S) were performed to compare and analyse the heme-binding process in CBS. NMR spectroscopy was also employed to characterise the heme interaction exploiting the advantage of this technique to describe also intrinsically disordered regions at atomic level.

Results

Protein expression and purification. The recombinant CBS(1–413) and the CBS(1–413)C15S mutant were expressed and purified as per literature⁴⁴. The N-terminal CBS peptide (_{GB1}CBS(1–40)), consisting of the residues 1–40 of CBS (MPSETPQAEVGPTGCPHRSGPHSAKGSLEKGSPEDEKEAKE), was expressed as fusion to the B1 domain of streptococcal protein G (GB1)⁴⁵.

UV/Vis study. UV/Vis spectroscopy was employed to monitor the interaction of the proteins with heme. UV/Vis data shows the heme binding by a Soret band shift to ~421 nm (Fig. 2a). Previous studies on heme binding to peptides using UV/Vis, Raman, EPR and NMR spectroscopy revealed that a UV shift to ~420 nm may be related to a hexacoordinated complex^{46–50}. Thus, this bathochromic effect is characteristic of heme binding and suggests formation of a hexacoordinated state which in CBS involves residues cysteine-15 and histidine-22. The K_D value was determined to be $2.18 \pm 0.64 \mu\text{M}$. According to the best fit possible _{GB1}CBS(1–40) binds to heme with a 1:1 stoichiometry (Fig. 2b).

Enzyme activity assay. The relative activities of CBS(1–413) and of the CBS(1–413)C15S mutant were compared in an enzyme activity assay using 7-azido-4-methylcoumarin as probe as described earlier in the literature¹⁰. To assure reproducibility six independent measurements were performed. During expression of the two enzymes in *E. coli* heme is endogenously produced and incorporated in the CBS proteins. This is reflected in yellowish colours of the protein samples with a darker colour of the CBS(1–413) with respect to the CBS(1–413)C15S mutant (data not shown) as shown for CBS and other recombinant heme-binding molecules^{51,52}. Hence, exogenous addition of heme during the assay does not lead to activity differences off to the mean from the independent measurements of both CBS(1–413) and CBS(1–413)C15S mutant. This indicates the saturation of the compounds with endogenous heme already during expression in the *E. coli* cells. Figure 3 shows that CBS(1–413)C15S was on average 32% (between 20% to 40% in the individual measurements) less active than CBS(1–413). The assay indicates that beside of its canonical binding site at Cys⁵²/His⁶⁵ heme is also binding to the CP-motif at position 15 and involving H22. This result of the assay is consistent with an observation of a previous study⁴⁴ where truncation of amino acids M1-K39 in CBS led to a similar decrease in enzyme activity.

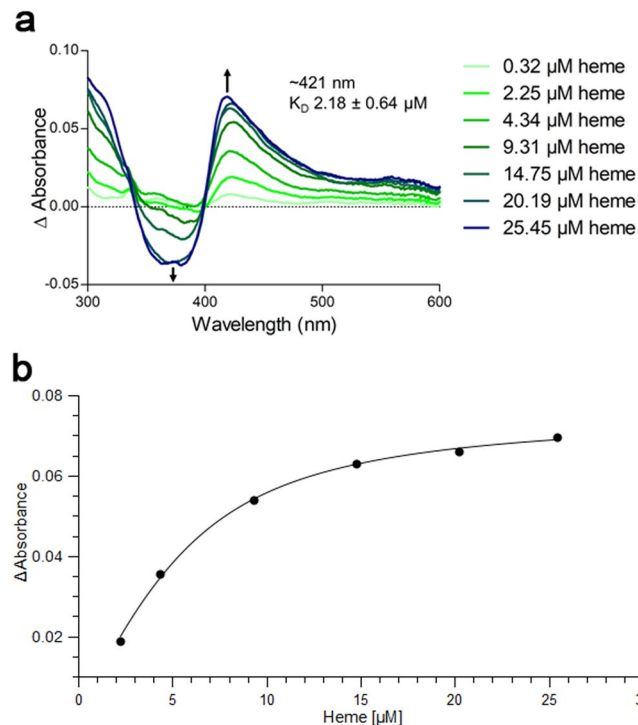


Figure 2. Heme binding to the $_{GB1}$ CBS(1–40) fusion protein determined by UV/Vis spectroscopy. **(a)** The $_{GB1}$ CBS(1–40) fusion protein was incubated with hemin at the concentrations indicated (5 μ M protein, 0.32–25.45 μ M hemin). **(b)** Titration curve observed at 421 nm.

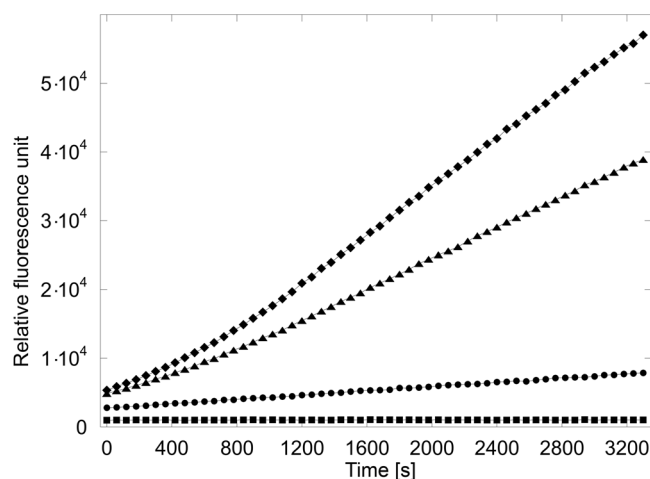


Figure 3. Fluorescence assay of CBS(1–413) and CBS(1–413)C15S. The average curve of six independent measurements is displayed. Symbols used: \blacklozenge = CBS(1–413), \blacktriangle = CBS(1–413)C15S, \bullet = blank, no protein, \blacksquare = blank, no probe.

NMR resonance assignment of the CBS N-terminal peptide. Sequence specific resonance assignments for $_{GB1}$ CBS(1–40) (Supplementary Table 1) were derived by a set of triple resonance 3D NMR spectra e.g. including well established experiments such as HNCA, HNCOCA, HNCO, HNCACO, HNCACB and HNN. Figure 4 shows the [1 H, 15 N]-HSQC spectrum of the $_{GB1}$ CBS(1–40) fusion protein including the assignments for the 40 N-terminal residues of CBS. The INEPT-based HNN experiment employed in this study essentially leads to the filtering out of signals from the structured part of the GB1 fusion protein due to significant relaxation losses, as shown in Supplementary Figure S1 which gives the [1 H, 15 N]-projections from the 3D HNCA and HNN experiments. This considerably simplified the data analysis of the $_{GB1}$ CBS(1–40) fusion protein. Representative strips indicating the sequential connectivities established are given in the Supplementary Material (Supplementary Figure S2). The calculation of the chemical shift index based on the obtained resonance assignment indicates an intrinsically disordered peptide which is supported by the prediction of the web server (IUPRED⁵³) using the primary sequence as sole input.

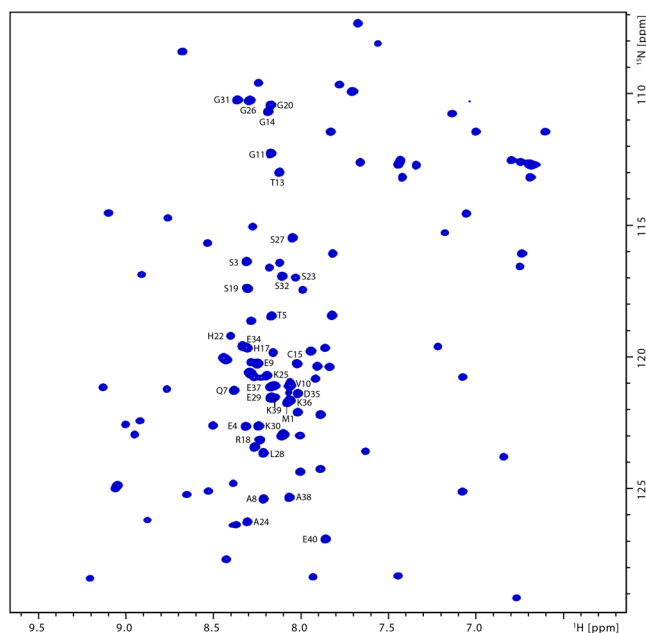


Figure 4. $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectrum of the $\text{GB1 CBS}(1-40)$ fusion protein. Cross peak assignments for the forty $\text{GB1 CBS}(1-40)$ peptide residues are indicated.

NMR study of the heme interaction with the CBS N-terminal peptide (1–40). Employing $100\ \mu\text{M}$ samples of ^{15}N , ^{15}N and ^{13}C -labelled GB1 -fusion peptides, $[^1\text{H}, ^{15}\text{N}]$ -HSQC ligand titration experiments were first carried out using paramagnetic heme samples. As the ligand concentration was varied over a range of 0 to $100\ \mu\text{M}$, neither chemical shift changes were observed nor new additional peaks with increasing intensities were seen as the concentration of the heme ligand was increased (Supplementary Figure S3). Only a reduction in the intensities of few peaks (T13, G14, C15, H17, R18, S19, H22, S23) was observed, with the reduction in the intensity being maximal near C15 (Fig. 5a). This indicates the spatial proximity of these residues to the paramagnetic center and is consistent with the observation of the Soret band near 420 nm in the UV/Vis spectra (Fig. 2a) of the heme complex which resembles the situation found for transient heme binding by histidine-containing peptides^{46,49}. A Soret band near 420 nm is typically associated with the occurrence of a hexacoordinated heme^{21,22}. It is worth mentioning, that as with the paramagnetic Fe-PPIX, similar intensity reductions in the $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectrum were also seen with diamagnetic Ga-PPIX (Fig. 5b). This suggests that in addition to the paramagnetic relaxation enhancement (PRE) effect^{54,55} there is also an exchange contribution to the reduction in the intensities observed with the paramagnetic Fe-PPIX. We also observed that the single mutation C15S (Fig. 5c) essentially leads to no effect on the relative intensities of the cross peaks in the HSQC spectra upon heme addition. This indicates that mutation of this residue leads to no heme binding and is consistent with the UV/Vis data of this mutant which shows no sharp Soret peak (Supplementary Figure S5) either at ~ 420 nm (hexacoordination) or at ~ 370 nm (pentacoordination). It is also apparent that C15 is involved in the transient heme interaction forming a hexacoordinated complex with H22 forming the other axial ligand (H17 as second coordination site can be excluded due to steric reasons). The role of both amino acids as axial ligands is further supported by the $[^1\text{H}, ^{13}\text{C}]$ -HSQC spectral data recorded with and without heme (Fig. 6, Supplementary Figure S4). Spectral cross sections taken at the C^β positions corresponding to the C15 and H17/H22 residues show a similar variation of the intensity indicating the proximity of these nuclei to the heme central ions. Consistent with this, similar intensity variations in the aromatic HSQC spectral cross sections taken at the histidine $^{13}\text{C}^\epsilon$ and $^{13}\text{C}^\delta$ positions (Supplementary Figure S6) were also observed.

In addition to the C15 mutation we have also mutated the H22 site. Although the UV/Vis spectra of this H22L mutant (Supplementary Figure S5) show a Soret peak at ~ 369 nm, possibly corresponding to pentacoordination at C15, the $[^1\text{H}, ^{15}\text{N}]$ -HSQC spectra carried out without and in presence of heme display no significant intensity variations of the cross peaks (Fig. 5d). This implies that either the heme-peptide interaction is much weaker and/or the rate of exchange between the heme-bound and the free peptide is not in the favorable range on the PRE timescale^{54,55}.

Viewed in totality, the results reported here suggest that the wild-type N-terminal CBS peptide upon heme binding undergoes a conformational change to a hexacoordinated complex with cysteine-15 and histidine-22 as ligands that is sparsely populated and is in exchange with a highly populated free peptide.

Discussion

Heme binding to the N-terminal stretch of CBS was confirmed by a combined UV/Vis and NMR study. In this study we decided for a fusion approach with the streptococcal protein GB1 since it is known to improve the solubility of the target protein without interfering with its fusion partner because of its inert nature. In addition,

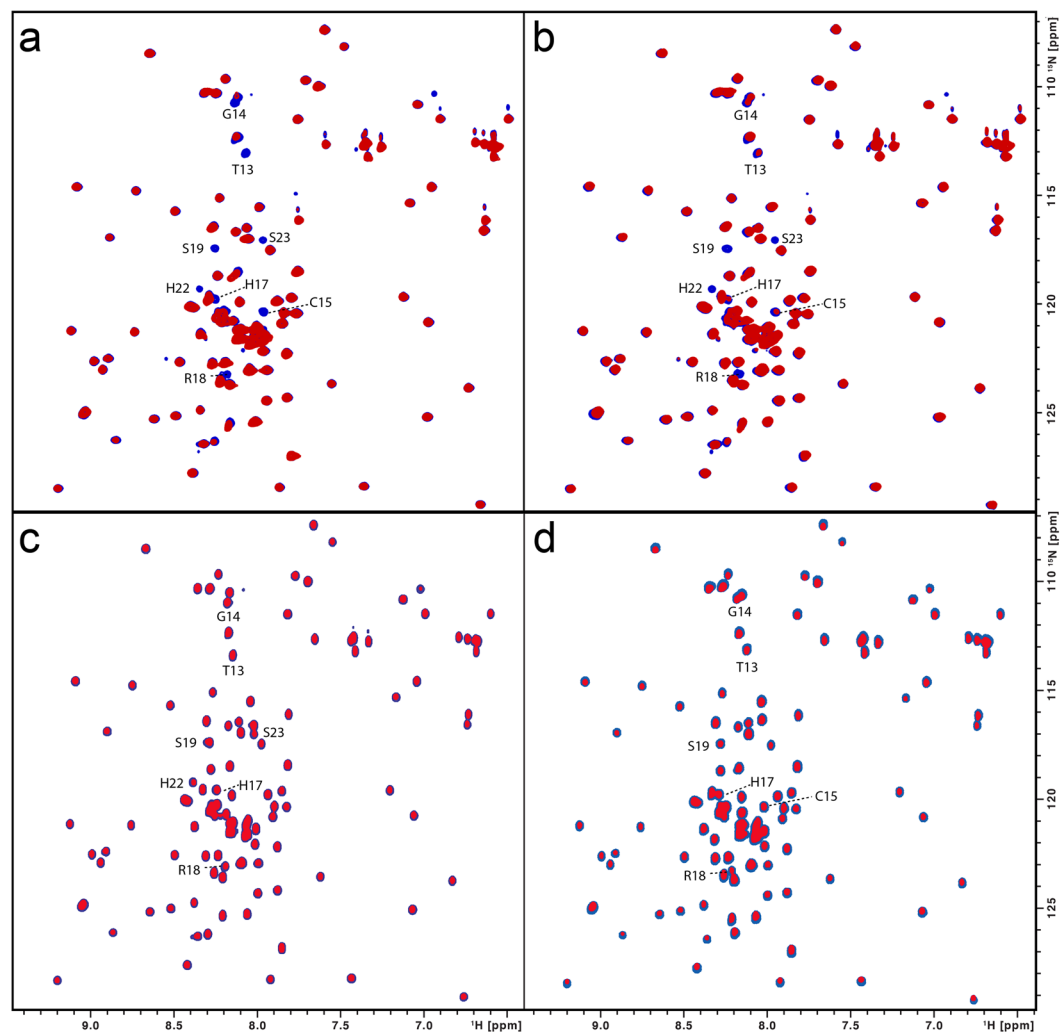


Figure 5. (a) Superimposition of the ^1H , ^{15}N -HSQC spectra of the $_{\text{GB1}}\text{CBS}(1-40)$ fusion protein ($100\ \mu\text{M}$) without (blue) and with hemin ($50\ \mu\text{M}$; red). Interacting residues are indicated. (b) Superimposition of the ^1H , ^{15}N -HSQC spectra of the $_{\text{GB1}}\text{CBS}(1-40)$ fusion protein ($100\ \mu\text{M}$) without (blue) and with Ga-PPIX ($100\ \mu\text{M}$; red). Interacting residues are indicated. (c) Superimposition of the ^1H , ^{15}N -HSQC spectra of the C15S mutant fusion protein ($100\ \mu\text{M}$) without (blue) and with hemin ($50\ \mu\text{M}$; red). (d) Superimposition of the ^1H , ^{15}N -HSQC spectra of the H22L mutant fusion protein ($100\ \mu\text{M}$) without (blue) and with hemin ($50\ \mu\text{M}$; red).

GB1 allows usage of easy purification protocols and is characterised with respect to structure and chemical shift assignments^{45,56,57}. This fusion approach was also selected since the amino acid sequence of GB1 does not contain further cysteine or histidine residues which constitute classical heme-binding residues. In addition, NMR experiments performed solely on the GB1 protein with heme revealed no interaction between the components (data not shown). The results obtained indicate that the resonance assignments of the intrinsically disordered peptide (IDP) can be conveniently carried out in GB1-fusion since the GB1 signals can be essentially filtered out from the IDP signals. Thus, spectral overlap problems can be minimised without the need for unlabelling approaches^{58,59} or ^{13}C -detected experiments³⁴. NMR spectroscopy revealed that the forty N-terminal residues constitute an intrinsically disordered region of the protein cystathionine- β -synthase. This is in accordance with the observation that all available X-ray studies did not deliver a structural description of this molecular stretch. In addition, this study demonstrates for the first time that the disordered N-terminal region of CBS contributes heme-binding capacities *via* a second binding site, the CP-based motif at cysteine-15 and a histidine residue at position 22. The functional assay reveals that this heme-binding site at cysteine-15 and histidine-22 as second axial heme ligand increases the efficacy of the enzyme by approx. 30%. The intrinsically disordered N-terminal peptide of CBS is thus involved in transient heme interactions leading to a hexacoordinated complex. Transient binding was observed in both situations, with heme carrying the paramagnetic iron center as well as its diamagnetic substitute, Ga-PPIX. This also stresses the utility of both the $\text{Fe}^{3+}/\text{Ga}^{3+}$ compounds in heme binding studies. The results on $_{\text{GB1}}\text{CBS}(1-40)$ also imply the caveat that a complete functional description of molecules has to carefully consider also regions truncated to facilitate crystallisation for X-ray studies which are often intrinsically disordered, highly dynamic protein regions potentially serving as interaction sites^{21,51,60}.

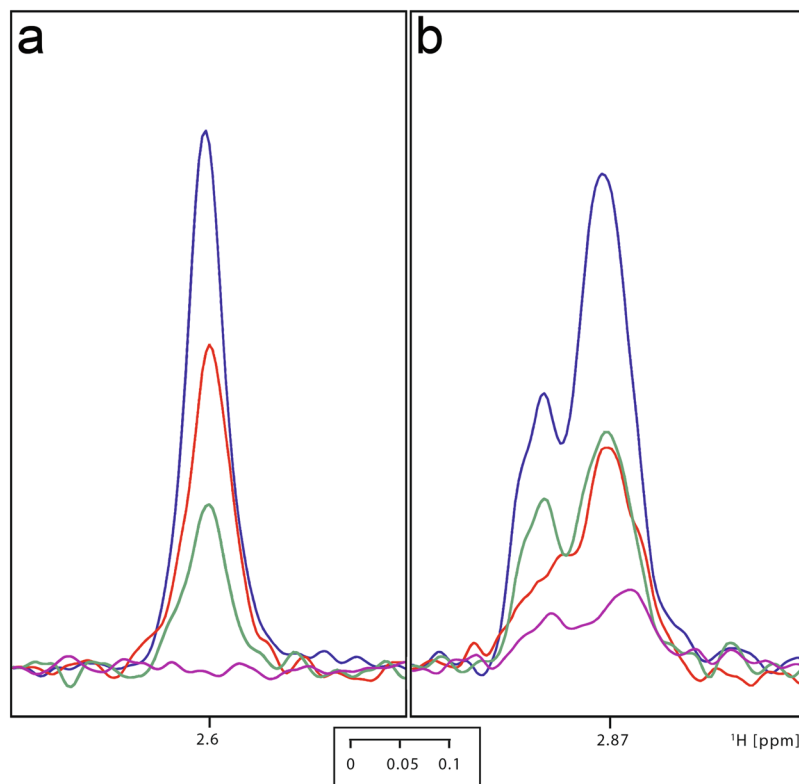


Figure 6. Superimposition of the spectral cross sections from $[^1\text{H},^{13}\text{C}]$ -HSQC spectra of $_{\text{GB1}}\text{CBS}(1-40)$ ($100\ \mu\text{M}$) taken at C^β positions corresponding to the C15 (a) and H17/H22 (b) residues in the free (blue), Ga-PPIX ($100\ \mu\text{M}$; red), Fe-PPIX ($50\ \mu\text{M}$; green) and Fe-PPIX ($100\ \mu\text{M}$; pink) exposed states.

Amino acid sequence comparison of the human and naked mole-rat CBS shows a deletion of the second heme-coordinating residue (histidine-22 in the human IDPR) in the latter species which occurred during evolution from a common ancestor. This might be explained by a lower need for CBS efficacy in sulfur metabolism induced by lower H_2S levels arising from a reduced dietary uptake of methionine in naked mole-rats⁶¹. As higher H_2S levels are connected e.g. to cancer¹², this could be one reason for the cancer resistance and longevity of naked mole-rats. However, human diet is rich in methionine, thus, evolutionary a contribution of a second heme-binding site might have become necessary to achieve the required efficacy for sulfur processing and to prevent pathological conditions. Hence, the second heme-binding site in CBS described in this study might become an attractive target for drug intervention to prevent pathological conditions without losing the complete biological functions of CBS, thus possibly preventing negative effects on the organism as observed for CBS gene deactivating mutations^{6,62}.

Methods

Protein expression and purification. CBS(1–413) protein was produced based upon the published protocol⁴⁴. The sample of the CBS(1–413)C15S mutant was generated using site-directed primers, expressed, purified as described below for $_{\text{GB1}}\text{CBS}(1-40)$ except no GB1 tag was used. Superdex 75 10/300 GL column (GE Healthcare, Freiburg, Germany) was used for size exclusion chromatography of CBS (1–413) and its mutant. To study the CBS(1–40) region, $_{\text{GB1}}\text{CBS}(1-40)$ was expressed *via* a fusion protein approach with the streptococcal protein GB1. It was cloned in pET28a vector with GB1 domain attached to N-terminal of CBS(1–40) with flexible linker (TEV protease site - ENLYFQG) to avoid proteolytic cleavage of CBS(1–40) and additionally increasing the yield with stability. $_{\text{GB1}}\text{CBS}(1-40)$ was expressed in *E. coli* BL21(DE3) cells which were grown up to an $\text{OD}_{600\text{nm}}$ of 0.7 in the LB medium. Cells were pelleted and transferred to M9 media containing $^{15}\text{NH}_4\text{Cl}$ and $^{13}\text{C}_6$ -glucose that will result in labelled $_{\text{GB1}}\text{CBS}(1-40)$. Protein expression was induced at OD_{600} of 0.7 by 0.3 mM IPTG for 18 h at 18 °C. Cells were resuspended in lysis buffer (50 mM Tris/HCl, 300 mM NaCl, 5 mM imidazole, 2 mM β -mercaptoethanol (pH 7.5)) followed by french press to lyse the cells and centrifuged at $10000 \times g$ to collect the supernatant cell lysate. Ni-NTA agarose resin was used to capture the His₆-tagged $_{\text{GB1}}\text{CBS}(1-40)$ protein followed by wash and elution with 10 mM and 250 mM imidazole, respectively. The His-tag was cleaved from the protein in overnight dialysis buffer (20 mM Tris/HCl, 150 mM NaCl, 2 mM DTT (pH 7.5)) by the addition of 5 U/mg of thrombin (Sigma-Aldrich, Taufkirchen, Germany) at 4 °C. The cleaved protein was concentrated to 1 mL using a 3-kDa Sartorius vivaspin filter and injected onto a 16/60 Hiload S75 size exclusion chromatography column (GE Healthcare, Freiburg, Germany) pre-equilibrated with 20 mM Tris/HCl, pH 6.9, 150 mM NaCl, 2 mM DTT to achieve the highest purity. The fractions containing the $_{\text{GB1}}\text{CBS}(1-40)$ were pooled together, concentrated, buffer

exchanged into 20 mM sodium phosphate buffer, pH 6.9 and the concentration measured at OD₂₈₀ using UV/Vis spectroscopy. Samples were lyophilised or directly used for the biophysical and NMR experiments.

UV/Vis spectroscopy. A Multiskan GO microplate spectro-photometer (Thermo Scientific, Dreieich, Germany) was employed for UV/Vis measurements. The hemin solution was prepared by dissolving 1 mM hemin in 30 mM NaOH followed by incubation in the dark for 30 minutes on ice. Subsequently, the hemin solution was diluted with the respective buffer to the concentrations used. _{GB1}CBS(1–40) (5 μM) protein was incubated in the dark with different concentration of hemin (0.2–30 μM) for 60 minutes before measuring at 300–600 nm. HEPES-buffer (100 mM, pH 7.0) was used for dissolving protein and hemin.

Enzyme activity assay and protein concentration. The enzyme activity assay was performed on CBS(1–413) and the CBS(1–413)C15S mutant using established protocols¹⁰. Fluorescence was read at 450 nm with excitation at 365 nm using an Infinite M1000 microplate reader (Tecan, Männedorf, Switzerland). The protein concentration was determined by UV/Vis spectroscopy.

NMR spectroscopy. Solution NMR experiments were performed on a Bruker Avance III spectrometers with proton frequencies of 600 MHz at a temperature of 283 K. Data were acquired, processed and analyzed with Topspin (Bruker, Rheinstetten, Germany). The resonance assignment was performed using a combination of experiments such as HNN, HNCO/HNCACO, HNCA/HNCOCA, HCN, HBCBCGCDHD, HNCACB, HBHANH, [¹H,¹⁵N]-HSQC, [¹H,¹³C]-HSQC etc. experiments with _{GB1}CBS(1–40) dissolved at concentrations of 700 μM. [¹H,¹³C]-HSQC spectra of the _{GB1}CBS(1–40) were recorded in the free, Ga(III)-protoporphyrin IX (Ga-PPIX) and hemin (Fe-PPIX) exposed state at concentrations of 100 μM and 50 μM, respectively. Ga(III)-protoporphyrin IX chloride and hemin were used as obtained from Frontier Scientific (Logan, USA). Unless otherwise stated, the _{GB1}CBS(1–40) and CBS samples were dissolved in 20 mM phosphate buffer and 5% D₂O.

Third-party methods. IUPred. Version 1.0 with default parameters was used.

Data availability. The authors declare that all data supporting the findings of this study are available within the article and its Supplementary Information files, or are available from the corresponding author upon request. The chemical shift assignments of CBS(1–40) have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) under accession code 27351.

References

- Miles, E. W. & Kraus, J. P. Cystathionine beta-synthase: structure, function, regulation, and location of homocystinuria-causing mutations. *J. Biol. Chem.* **279**, 29871–4 (2004).
- Banerjee, R. & Zou, C. G. Redox regulation and reaction mechanism of human cystathionine-beta-synthase: a PLP-dependent hemesensor protein. *Arch. Biochem. Biophys.* **433**, 144–56 (2005).
- Refsum, H., Ueland, P. M., Nygard, O. & Vollset, S. E. Homocysteine and cardiovascular disease. *Annu. Rev. Med.* **49**, 31–62 (1998).
- Mills, J. L. *et al.* Homocysteine metabolism in pregnancies complicated by neural-tube defects. *Lancet* **345**, 149–51 (1995).
- Clarke, R. *et al.* Folate, vitamin B12, and serum total homocysteine levels in confirmed Alzheimer disease. *Arch. Neurol.* **55**, 1449–55 (1998).
- Kraus, J. P. *et al.* Cystathionine beta-synthase mutations in homocystinuria. *Hum. Mutat.* **13**, 362–75 (1999).
- Kraus, J. P. CBS mutation database. <http://cbs.lfi.cuni.cz>.
- Evande, R., Ojha, S. & Banerjee, R. Visualization of PLP-bound intermediates in hemeless variants of human cystathionine beta-synthase: evidence that lysine 119 is a general base. *Arch. Biochem. Biophys.* **427**, 188–96 (2004).
- Taoka, S. & Banerjee, R. Characterization of NO binding to human cystathionine beta-synthase: possible implications of the effects of CO and NO binding to the human enzyme. *J. Inorg. Biochem.* **87**, 245–51 (2001).
- Thorson, M. K., Majtan, T., Kraus, J. P. & Barrios, A. M. Identification of cystathionine beta-synthase inhibitors using a hydrogen sulfide selective probe. *Angew. Chem. Int. Ed. Engl.* **52**, 4641–4 (2013).
- Wallace, J. L. & Wang, R. Hydrogen sulfide-based therapeutics: exploiting a unique but ubiquitous gasotransmitter. *Nat. Rev. Drug Discov.* **14**, 329–45 (2015).
- Hellmich, M. R., Coletta, C., Chao, C. & Szabo, C. The therapeutic potential of cystathionine beta-synthetase/hydrogen sulfide inhibition in cancer. *Antioxid. Redox Signal.* **22**, 424–48 (2015).
- Hellmich, M. R. & Szabo, C. Hydrogen Sulfide and Cancer. *Handb. Exp. Pharmacol.* **230**, 233–41 (2015).
- Meier, M., Janosik, M., Kery, V., Kraus, J. P. & Burkhard, P. Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *EMBO J.* **20**, 3910–6 (2001).
- Ereno-Orbea, J., Majtan, T., Oyenarte, I., Kraus, J. P. & Martinez-Cruz, L. A. Structural basis of regulation and oligomerization of human cystathionine beta-synthase, the central enzyme of transsulfuration. *Proc. Natl. Acad. Sci. USA* **110**, E3790–9 (2013).
- Ojha, S., Wu, J., LoBrutto, R. & Banerjee, R. Effects of heme ligand mutations including a pathogenic variant, H65R, on the properties of human cystathionine beta-synthase. *Biochemistry* **41**, 4649–54 (2002).
- Kuhl, T. & Imhof, D. Regulatory Fe(II/III) heme: the reconstruction of a molecule's biography. *ChemBioChem* **15**, 2024–35 (2014).
- Zhang, L. & Guarente, L. Heme binds to a short sequence that serves a regulatory function in diverse proteins. *EMBO J.* **14**, 313–20 (1995).
- Zhang, L. In *Heme Biology: The Secret Life of Heme in Regulating Diverse Biological Processes* (ed. Zhang, L.), (World Scientific, Singapore, 2011).
- Hira, S., Tomita, T., Matsui, T., Igarashi, K. & Ikeda-Saito, M. Bach1, a heme-dependent transcription factor, reveals presence of multiple heme binding sites with distinct coordination structure. *IUBMB Life* **59**, 542–51 (2007).
- Watanabe-Matsui, M. *et al.* Heme binds to an intrinsically disordered region of Bach2 and alters its conformation. *Arch. Biochem. Biophys.* **565**, 25–31 (2015).
- Suenaga, T. *et al.* Charge-state-distribution analysis of Bach2 intrinsically disordered heme binding region. *J. Biochem.* **160**, 291–298 (2016).
- McCorvie, T. J. *et al.* Inter-domain communication of human cystathionine beta-synthase: structural basis of S-adenosyl-L-methionine activation. *J. Biol. Chem.* **289**, 36018–30 (2014).

24. Vicente, J. B. *et al.* A Clinically Relevant Variant of the Human Hydrogen Sulfide-Synthesizing Enzyme Cystathionine beta-Synthase: Increased CO Reactivity as a Novel Molecular Mechanism of Pathogenicity? *Oxid. Med. Cell. Longev.* **2017**, 8940321 (2017).
25. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
26. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M. & Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **41**, 6573–82 (2002).
27. Rezaei-Ghaleh, N., Blackledge, M. & Zweckstetter, M. Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *ChemBioChem* **13**, 930–50 (2012).
28. Uversky, V. N. Proteins without unique 3D structures: biotechnological applications of intrinsically unstable/disordered proteins. *Biotechnol. J.* **10**, 356–66 (2015).
29. Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. & Jones, D. T. The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138–9 (2004).
30. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635–45 (2004).
31. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* **37**, 215–46 (2008).
32. He, B. *et al.* Predicting intrinsic disorder in proteins: an overview. *Cell Res.* **19**, 929–49 (2009).
33. Deng, X., Eickholt, J. & Cheng, J. A comprehensive overview of computational protein disorder prediction methods. *Mol. Biosyst.* **8**, 114–21 (2012).
34. Gray, F. L., Murai, M. J., Grembecka, J. & Cierpicki, T. Detection of disordered regions in globular proteins using ¹³C-detected NMR. *Protein Sci.* **21**, 1954–60 (2012).
35. Mantylahti, S., Hellman, M. & Permi, P. Extension of the HA-detection based approach: (HCA)CON(CA)H and (HCA)NCO(CA)H experiments for the main-chain assignment of intrinsically disordered proteins. *J. Biomol. NMR* **49**, 99–109 (2011).
36. Felli, I. C. & Pierattelli, R. Recent progress in NMR spectroscopy: toward the study of intrinsically disordered proteins of increasing size and complexity. *IUBMB Life* **64**, 473–81 (2012).
37. Wiedemann, C. *et al.* A Set of Efficient nD NMR Protocols for Resonance Assignments of Intrinsically Disordered Proteins. *ChemPhysChem* **17**, 1961–8 (2016).
38. Wiedemann, C. *et al.* HN-NCA heteronuclear TOCSY-NH experiment for ¹H(N) and ¹⁵N sequential correlations in ((¹³C), (¹⁵N)) labelled intrinsically disordered proteins. *J. Biomol. NMR* **63**, 201–12 (2015).
39. Goradia, N. *et al.* An approach to NMR assignment of intrinsically disordered proteins. *ChemPhysChem* **16**, 739–46 (2015).
40. Narayanan, R. L. *et al.* Automatic assignment of the intrinsically disordered protein Tau with 441-residues. *J. Am. Chem. Soc.* **132**, 11906–7 (2010).
41. Novacek, J., Zidek, L. & Sklenar, V. Toward optimal-resolution NMR of intrinsically disordered proteins. *J. Magn. Reson.* **241**, 41–52 (2014).
42. Felli, I. C., Pierattelli, R. & Tompa, P. Intrinsically Disordered Proteins. In *NMR of Biomolecules: Towards Mechanistic Systems Biology* (eds Bertini, I., McGreevy, K. S. & Parigi, G.) 137–152 (Wiley-VCH Verlag & Co. KGaA, Weinheim, 2012).
43. Gibbs, E. B., Cook, E. C. & Showalter, S. A. Application of NMR to studies of intrinsically disordered proteins. *Arch. Biochem. Biophys.* **628**, 57–70 (2017).
44. Oliveriusova, J., Kery, V., Maclean, K. N. & Kraus, J. P. Deletion mutagenesis of human cystathionine beta-synthase. Impact on activity, oligomeric status, and S-adenosylmethionine regulation. *J. Biol. Chem.* **277**, 48386–94 (2002).
45. Gronenborn, A. M. *et al.* A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* **253**, 657–61 (1991).
46. Brewitz, H. H. *et al.* Role of the Chemical Environment beyond the Coordination Site: Structural Insight into Fe(III) Protoporphyrin Binding to Cysteine-Based Heme-Regulatory Protein Motifs. *ChemBioChem* **16**, 2216–24 (2015).
47. Kuhl, T. *et al.* Analysis of Fe(III) heme binding to cysteine-containing heme-regulatory motifs in proteins. *ACS Chem. Biol.* **8**, 1785–93 (2013).
48. Schubert, E. *et al.* Spectroscopic studies on peptides and proteins with cysteine-containing heme regulatory motifs (HRM). *J. Inorg. Biochem.* **148**, 49–56 (2015).
49. Brewitz, H. H. *et al.* Heme interacts with histidine- and tyrosine-based protein motifs and inhibits enzymatic activity of chloramphenicol acetyltransferase from *Escherichia coli*. *Biochim. Biophys. Acta* **1860**, 1343–53 (2016).
50. Shimizu, T. Binding of cysteine thiolate to the Fe(III) heme complex is critical for the function of heme sensor proteins. *J. Inorg. Biochem.* **108**, 171–7 (2012).
51. Sahoo, N. *et al.* Heme impairs the ball-and-chain inactivation of potassium channels. *Proc. Natl. Acad. Sci. USA* **110**, E4036–44 (2013).
52. Dimster-Denk, D., Tripp, K. W., Marini, N. J., Marqusee, S. & Rine, J. Mono and dual cofactor dependence of human cystathionine beta-synthase enzyme variants *in vivo* and *in vitro*. *G3 (Bethesda)* **3**, 1619–28 (2013).
53. Dosztanyi, Z., Csizmok, V., Tompa, P. & Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433–4 (2005).
54. Anthis, N. J. & Clore, G. M. Visualizing transient dark states by NMR spectroscopy. *Q. Rev. Biophys.* **48**, 35–116 (2015).
55. Clore, G. M. Seeing the invisible by paramagnetic and diamagnetic NMR. *Biochem. Soc. Trans.* **41**, 1343–54 (2013).
56. Sommer, L. A., Meier, M. A. & Dames, S. A. A fast and simple method for probing the interaction of peptides and proteins with lipids and membrane-mimetics using GB1 fusion proteins and NMR spectroscopy. *Protein Sci.* **21**, 1566–70 (2012).
57. Wilton, D. J., Tunnicliffe, R. B., Kamatari, Y. O., Akasaka, K. & Williamson, M. P. Pressure-induced changes in the solution structure of the GB1 domain of protein G. *Proteins* **71**, 1432–40 (2008).
58. Zhou, P. & Wagner, G. Overcoming the solubility limit with solubility-enhancement tags: successful applications in biomolecular NMR studies. *J. Biomol. NMR* **46**, 23–31 (2010).
59. Kobayashi, H. *et al.* Segmental isotope labeling of proteins for NMR structural study using a protein S tag for higher expression and solubility. *J. Biomol. NMR* **52**, 303–13 (2012).
60. Uversky, V. N. The multifaceted roles of intrinsic disorder in protein complexes. *FEBS Lett.* **589**, 2498–506 (2015).
61. Dziegielewska, M. *et al.* Low sulfide levels and a high degree of cystathionine beta-synthase (CBS) activation by S-adenosylmethionine (SAM) in the long-lived naked mole-rat. *Redox Biol.* **8**, 192–8 (2016).
62. de Franckis, R., Kraus, E., Kozich, V., Sebastio, G. & Kraus, J. P. Four novel mutations in the cystathionine beta-synthase gene: effect of a second linked mutation on the severity of the homocystinuric phenotype. *Hum. Mutat.* **13**, 453–7 (1999).

Acknowledgements

We thank Dr. Klaus Huse, Dr. Karol Szafranski and Maja Dziegielewska (FLI) for helpful discussions. Financial support by the Deutsche Forschungsgemeinschaft (DFG) within FOR 1738 (to D.I. and O.O.) is gratefully acknowledged. The FLI is a member of the Leibniz Association (WGL) and is financially supported by the Federal Government of Germany and the State of Thuringia.

Author Contributions

O.O. and D.I. designed the research; A.K., A.W., N.G., P.B. and R.R. performed the experiments; A.K., R.R. and O.O. analysed the data; A.K., R.R. and O.O. wrote the paper, all authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-20841-z>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018