A yeast system for expression of human cystathionine \( \beta \)-synthase: Structural and functional conservation of the human and yeast genes

(homocystinuria/inherited metabolic disorder/\textit{Saccharomyces cerevisiae}/CYS4)

WARREN D. KRUGER* AND DAVID R. COX*

Department of Psychiatry, University of California at San Francisco, San Francisco, CA 94143

Communicated by Ira Herskowitz, March 14, 1994 (received for review May 17, 1993)

ABSTRACT Human cystathionine \( \beta \)-synthase (CBS; EC 4.2.1.22) deficiency results in a recessive genetic disorder whose clinical and biochemical manifestations vary greatly among affected individuals. In an effort to identify and analyze mutations in the human CBS gene, we have developed a yeast expression system for human CBS. We have cloned and sequenced a human cDNA that codes for CBS and have expressed the human CBS protein in yeast cells lacking endogenous CBS. The human enzyme produced in yeast is functional both \textit{in vitro} and \textit{in vivo}. We have also cloned and sequenced the yeast gene, \textit{CYS4}, that codes for CBS. The predicted human and yeast CBS proteins are 38\% identical and 72\% similar to each other, as well as sharing significant similarity with bacterial cysteine synthase. These results demonstrate the evolutionary conservation of CBS and establish the utility of a yeast expression system for studying human CBS.

Cystathionine \( \beta \)-synthase (CBS) deficiency in humans is a rare autosomal recessive disease, which is the most common cause of homocystinuria. Individuals with this disorder have a variety of clinical phenotypes including ocular, skeletal, neurological, and cardiovascular defects, which are believed to result primarily from elevated plasma homocysteine levels. Age of onset, symptoms, and severity of the disease can vary greatly among patients (1). Some, but not all, individuals with homocystinuria respond to increased dietary levels of vitamin B\(_6\) with an increase in residual enzyme activity associated with a lowering of plasma homocysteine levels. Within sibships, almost complete concordance is observed for responsiveness and nonresponsiveness to vitamin B\(_6\) (2). Concordance within sibships is also observed in the residual CBS activity found in cultured fibroblast cells derived from affected individuals (3). Residual enzyme activity in fibroblast extracts ranges from undetectable to 10\% of normal controls. Generally, vitamin B\(_6\) responders have higher levels of residual CBS activity than nonresponders (3–6). These observations suggest that different mutations within the CBS gene alter the enzyme activity in discrete ways; thus, different CBS mutations may be related to the clinically observed heterogeneity found in patients with CBS deficiency. A life-threatening complication of CBS deficiency is thromboembolism (1). Untreated vitamin B\(_6\) responsive and nonresponsive individuals have a 4\% risk per year of a thrombotic event (2). Thrombosis can occur at any age and can involve both large and small arteries. Studies also suggest that individuals heterozygous for mutations in CBS are at increased risk for premature peripheral and cerebral arterial disease (7–10). These studies show a correlation between reduced levels of CBS activity and premature vascular disease. Since \( \approx 1 \) in 200 people in the general population is a heterozygote (1), it has been suggested that these individuals may account for a large fraction of those at risk for stroke. However, testing this hypothesis is difficult due to the lack of an accurate assay that identifies heterozygotes (11).

To identify and study mutations that cause homocystinuria, an effective assay for the human CBS gene product is required. Such an assay system must be able to distinguish DNA polymorphisms with no phenotypic effect from mutations that affect enzymatic function. In addition, the assay must be able to isolate individual alleles for study.

We report here the development of such an experimental system for the study of human CBS using \textit{Saccharomyces cerevisiae}. We describe the cloning and sequencing of a human CBS-encoding cDNA\(^1\) and show that expression of this cDNA is able to functionally complement a yeast strain that lacks endogenous CBS activity. Additionally, we have cloned and sequenced the yeast gene encoding CBS (\textit{CYS4}). The predicted yeast and human CBS proteins are 72\% similar to each other and have significant similarity to the predicted rat CBS protein and bacterial cysteine synthase. These results show the strong evolutionary conservation of CBS and suggest the utility of this yeast system for detection and analysis of mutations in the human CBS coding region.

MATERIALS AND METHODS

Yeast Strains and Genetic Methods. Yeast strains WC6-6a (MAT\(a\) ura3-52 cys4-1 cys2-1) and WC4-4a (MAT\(a\) ura3-52) were constructed by crossing JW1-2c (MAT\(a\) cys2-1 cys4-1) (12) and CY70 (MAT\(a\) trp1 ura3-52 his3 leu2) (13). Standard yeast genetic methods were performed essentially as described (14).

Cloning of Human CBS cDNA. A human cDNA library derived from HepG2 cells constructed in vector pAB23-BXN (15) was used as a template for PCR amplification using the CBS-specific primers 5'-TCCATTATATGAAATGTCC (forward) and 5'-TGTTGCTCAGCATCCTCC (reverse) (16). PCR was carried out in 67 mM Tris (pH 8.8), 6.7 mM Mg\(_{2+}\), 16.6 mM NH\(_4\)SO\(_4\), 10 mM 2-mercaptoethanol, and 6.7 \( \mu \)M EDTA. The cycling parameters were 30 sec at 94°C, 30 sec at 50°C, and 30 sec at 72°C for 40 cycles. Bacterial cells containing the human cDNA inserts were diluted in liquid medium and used to inoculate cultures at \( \approx 50,000 \) cells per tube. Miniprep DNA was prepared and analyzed for the presence of a CBS-encoding cDNA by PCR, and the product was visualized by agarose gel electrophoresis. One of the positive pools was diluted and used to inoculate 10 new pools at 8000 cells each. One of these pools tested positive by PCR.

Abbreviations: CBS, cystathionine \( \beta \)-synthase; ORF, open reading frame.

\(^1\)Present address: Department of Genetics, School of Medicine, Stanford University, Stanford, CA 94305.

\(^{1}\)The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L134577 and L134578).
and this pool was plated out for individual colonies, which were subsequently screened by hybridization with the PCR-generated human CBS fragment as probe. Probe was prepared by random priming (17), and hybridizations were carried out using Biotrans membranes (ICN) under conditions recommended by the manufacturer. The final clone was designated pHCBS.

Cloning of CBS and Construction of a CBS Null Allele. Previous studies have shown that yeast containing cys4-1 do not exhibit CBS activity (12). Therefore, strain WC6-6a was transformed with a yeast genomic library constructed in YCP50 (18) and plated out on SC-URA (14) plates supplemented with 20 μl of glutathione (30 mg/ml) as a source of cysteine. Approximately 2000 transformants were then replica plated to unsupplemented SC-URA plates. Positive colonies were restested and cured of the plasmid by passage on 5-fluoroorotic acid plates (19). DNA was prepared from the yeast and used to transform Escherichia coli. The recovered plasmid was named pYCBS. Restriction fragments were subcloned into YCP50 as indicated in Fig. 1.

A null allele of CBS was constructed as follows: Initially, the 3.2-kb BamHI-Sph I fragment containing CBS was subcloned into pUC18. The resulting construct was then cut with Bgl II and a 2.2-kb LEU2 fragment with BamHI ends was inserted. The LEU2 fragment was derived from pCD7-B, a derivative of pRB758. The resulting clone, pUC:YCBS:: LEU2 (Fig. 1C), was digested with Sph I and used to transform CY70. Nineteen out of 20 Leu+ colonies examined had a Cys− phenotype. Proper integration was confirmed by hybridization.

Sequencing. A 2-kb EcoRI fragment, a 1.2-kb EcoRI-Sph I fragment, and a 3.2-kb BamHI-Sph I fragment from the yeast clone pYCBS (containing the CBS gene) were subcloned into pVZ1+ for sequencing. The entire EcoRI fragment containing the human CBS cDNA from clone pHCBS1 was also cloned into pVZ1+. pVZ1+ is pBS+ with an expanded polylinker (20). All sequencing was done using CsCl-purified double-stranded DNA as template and Sequenase. The following primers, in addition to the vector T7 and T3 primers, were used in the CBS sequencing: 5′-GACCAATATAAAGATATATGAT 5′-CTTGGCATTGGCTACGTTGAC, 5′-TCATTTAATAGGATTCTTCTG, 5′-AATGCTTAAACTCTGACG, and 5′-GTGGGCTTA AAATTGAAACC.

The following primers were used in the human CBS sequencing: 5′-CTGGCATTGGCTACGTTGAC, 5′-GACCAATATAAAGATATATGAT 5′-CTTGGCATTGGCTACGTTGAC, 5′-TCATTTAATAGGATTCTTCTG, 5′-AATGCTTAAACTCTGACG, and 5′-GTGGGCTTA AAATTGAAACC. The following primers were used in the human CBS sequencing: 5′-CTTGGCATTGGCTACGTTGAC, 5′-GACCAATATAAAGATATATGAT 5′-CTTGGCATTGGCTACGTTGAC, 5′-TCATTTAATAGGATTCTTCTG, 5′-AATGCTTAAACTCTGACG, and 5′-GTGGGCTTA AAATTGAAACC.

Construction of pHCBS and pHCBS. The plasmid pAB23BXN is a 2-μm shuttle vector containing a multiple cloning site placed in between glyceraldehyde-3-phosphate dehydrogenase promoter and terminator sequences (15). The original cDNA clone, pHCBS, had the human CBS cDNA insert in the antisense orientation relative to the promoter in the vector pAB23BXN. Therefore, pHCBS1DNA was digested with Bgl II, which released the insert, and subsequently treated with T4 ligase. One of the resulting clones contained the insert in the correct orientation and was called pHCBS. To delete the 5′ untranslated region of the CBS cDNA, we made use of an Sph I site containing the predicted start ATG for the human CBS gene. The plasmid pVZ1+:HCBS was partially digested with Sph I, and a resulting 2.2-kb fragment containing the human CBS cDNA and 3′ polylinker sequences (containing a Sal I site) was isolated and inserted into pUC18. The insert was removed from this plasmid by digestion with Sal I and was subsequently inserted into pBX-X at the Xho I site. Plasmid pBX-X is identical to pAB23BXN except for replacement of the multiple cloning site with a single Xho I site; it was created by digestion of pHCBS1 with Bgl II, isolation of the vector backbone, and ligation in the presence of Xho I linker DNA.

CBS Enzyme Assays. Yeast extracts were prepared from 100-ml cultures grown to midlogarithmic phase in SC-URA medium. Cells were harvested at 4°C, washed once with ice-cold lys buffer (20 mM NaPO4, pH 7.5), suspended in 500 μl of lys buffer, and transferred to tubes containing 500 μl of acid-washed glass beads (0.2 μm). Cells were lysed using a Biospec Products (Bartlesville, OK) minibeadbeater for 2 min (4 30-sec pulses). Each lysed cell suspension was then centrifuged in a Fisher microcentrifuge for 10 min to remove insoluble material. Protein concentrations were determined with the Bio-Rad protein assay reagent, using bovine serum albumin as an internal standard. CBS enzyme assays were conducted as described (21), except assay volumes were 200 μl, and the reaction time was 4 hr. Cystathionine was measured as described (21) by mixing 70 μl of the assay reaction with 1 ml of Nihydrin reagent and boiling for 5 min. OD520 was measured using a Beckman model DU-50 spectrophotometer. Levels of cystathionine produced were quantitatively determined by comparing results against a standard of known cystathionine concentration. The assay was linear over the time tested.

RESULTS

Cloning of a Human CBS cDNA. A human CBS cDNA clone was isolated using two primers derived from the 3′ untranslated region of the human CBS gene and used to map the gene on chromosome 21 (16). These primers were used to screen by PCR a human HepG2 cDNA library (15) for the presence of a 323-bp product. The library was subsequently divided into subpools and screened again. One subpool was identified that contained the CBS cDNA. This pool was plated out, and the clones were screened by hybridization using the radiolabeled 323-bp CBS PCR product as a probe. Three positive clones were obtained, all of which had an identical 2.5-kb cDNA insert. A map of this clone and the sequencing strategy are described (22).

This sequence has been deposited in GenBank (accession no. L14577). The cDNA contains one very long open reading frame (ORF) of 1656 bp and two short ORFs upstream. The DNA sequence of the long open reading frame has 80% identity with the rat CBS cDNA type 1, which is one of the four splice variants of the rat CBS cDNA previously described (23). At the amino acid level, the similarity is even more striking. Greater than 90% of the amino acids are similar between the predicted rat and human proteins (see Fig. 2).

Cloning of the Yeast Gene Encoding CBS. An S. cerevisiae genomic clone encoding CBS was isolated by complementation of a yeast strain containing the cys4-1 mutation (WC6-6a). Previous work had shown that yeast containing this mutation are unable to form colonies in the absence of exogenous cysteine and that extracts of these strains lack detectable CBS activity in vitro (ref. 12; Table 1). This yeast strain was transformed with a yeast genomic library, and seven colonies were identified that grew in the absence of exogenous cysteine (see Materials and Methods). This phenotype was found to be plasmid dependent. Additionally, extracts made from cells carrying the plasmid had CBS enzyme activity in vitro (Table 1). The plasmids were recovered, and restriction analysis indicated that they all contained identical inserts. Subcloning of the insert followed by transformation into WC6-6a revealed that an Sph I-BamHI subfragment was sufficient for complementation and that an EcoRI site must lie in or very near the gene (see Fig. 1B). Sequencing in both directions from the EcoRI site revealed:

Genetics: Kruger and Cox

cysteine (data not shown). However, when the construct was reengineered such that the 5' untranslated region of the human CBS cDNA was deleted (see Materials and Methods), growth was observed. As shown in Fig. 3, the plasmid expressing the truncated human CBS cDNA, pAHCBS, allows strain WC6-6a to form colonies in the absence of exogenous cysteine, while the same strain containing the expression vector alone, pBX-X, does not. The figure also shows that cells containing pAHCBS do not grow as rapidly as cells containing a plasmid expressing CY54, pYCBS, in the absence of exogenous cysteine. Identical results were observed using a yeast strain carrying a null allele of CY54 (Fig. 1C; data not shown).

CBS activity was also assayed from whole-cell extracts made from these cells in vitro. CBS activity is detected in vitro in extracts made from yeast expressing the human CBS cDNA, but not from yeast with the expression vector alone (Table 1). This activity, however, is only about 20% of the activity seen in wild-type yeast extracts or in extracts made from the mutant strain carrying the yeast CBS gene on a plasmid. Thus, the reduced levels of activity observed in vitro correlate with the slow growth of the strain in vivo.

**DISCUSSION**

This report describes the cloning and sequencing of a human CBS clone and a yeast genomic clone encoding CBS. We demonstrate that the human gene is able to substitute functionally for the yeast gene in both in vivo and in vitro assays.

CBS catalyzes the formation of cystathionine from homocysteine and serine. This reaction involves the creation of a sulfide linkage between the homocysteine and serine and the production of water as a leaving group. Both the human and yeast CBS proteins are able to carry out this reaction (12, 24). It is therefore not surprising that these two enzymes share a large degree of sequence similarity. The yeast protein is slightly shorter than the human protein (506 vs 552 amino acids). The alignment (Fig. 2) indicates that most of this difference occurs in the N terminus. No splice acceptor sequences were detected upstream of the predicted yeast start codon, suggesting that this difference is genuine. The function of the N terminus of the mammalian CBS protein has not been determined.

The predicted yeast protein shares 38% identity and 72% similarity over its entire length with the human protein. Interestingly, the human and yeast proteins also have significant similarity with bacterial cysteine synthase A, which is encoded by the cysK gene (25). The human and yeast CBS proteins are, respectively, 37% and 34% identical to the cysK gene product over its entire length. Cysteine synthase catalyzes the addition of inorganic sulfide to O-acetylsereine to form cysteine with acetic acid as a leaving group. Thus both CBS and cysteine synthase condense serine with sulfur-containing compounds to form cysteine. Additionally, both mammalian CBS and bacterial cysteine synthase use pyridoxal phosphate as a cofactor (21, 25). There is one exceptionally large block of identity between amino acids 144 and 155 of human CBS in which 10 of 11 amino acids are identical in all four CBS-related proteins. We speculate that this region could be involved in the binding of the common elements: serine, pyridoxal phosphate, and sulfur. Interestingly, mammalian CBS can perform the bacterial reaction in vitro (26). The similarity of these two enzymatic reactions suggests that the condensation of serine with sulfur compounds to produce cysteine must have occurred early in evolution.

Recently, after this work had been submitted, both the human and yeast CBS DNA sequences were published. The yeast CBS DNA sequence published by Cherest et al. (27) and the one published here differ at three positions: 205, 408, and 1239. The coding region for the human CBS cDNA
The ability to complement yeast cells lacking yeast CBS activity with the human cDNA provides an assay system to identify and characterize the effects of various mutations in human CBS. Such an assay system is significant because it allows analysis of mutations at a functional level.

Using the yeast assay for human CBS in combination with reverse transcription-PCR, it should be possible to detect heterozygotes in the general population. These studies could firmly establish whether heterozygotes are actually at increased risk of stroke. Some potential limitations of this approach include misdiagnosis due to mutations introduced by PCR and inability to detect mutations that lie outside the coding region. Additionally, it may be that certain mutations harmful in humans do not give phenotypes in yeast.

Kozich and Kraus (24) have successfully used a bacterial expression system for the human CBS cDNA to identify and analyze two mutations in the CBS gene. This system essentially uses the bacteria as a factory to produce the enzyme and then analyzes the product in vitro. In contrast, the yeast system described here can monitor activity both in vitro and in vivo. Thus the yeast system can examine human CBS function in the context of the entire cysteine metabolic pathway.

The yeast expression system should be useful in analysis of allelic interactions between different human CBS mutations. Because CBS is normally present as a homotetramer (30), certain heterotetramers may have altered function. By allowing analysis of alleles singly and in combination with other alleles, the yeast system should be particularly well suited for such subunit interaction studies.

Similar yeast systems have recently been developed to study other human diseases. A yeast expression system for the study of human galactose-1-phosphate uridylyltransferase has been successfully used to identify and characterize a mutation in a patient with galactosemia (31). Another yeast-based expression system has been described for galactokinase, which may be useful in analyzing mutations that cause cataracts in newborns (23). Given the high degree of conservation in basic metabolic pathways between human
and yeast, it is likely that yeast expression assays can be developed for a wide variety of genetic metabolic disorders.

We thank Anthony Brake and Chiron for the use of the cDNA library. We also thank Chris Clarke for her gift of pRB758. This work was supported by a grant from the National Alliance for Research on Schizophrenia and Depression (NARSAD) to W.D.K. and National Institutes of Health Grant HD24610 to D.R.C.