Journal of Bioenergetics and Biomembranes, Vol. 27, No. 2, 1995

Molecular Abnormalities of Coproporphyrinogen Oxidase in Patients with Hereditary Coproporphyria

Bernard Grandchamp,[1] Jerome Lamoril,[1] and Hervé Puy[1]

Received January 18, 1995; accepted February 1, 1995

Genetic defects of coproporphyrinogen oxidase (CPO) lead to hereditary coproporphyria, an inherited autosomal dominant porphyria. The recent cloning of human cDNAs and of the gene encoding CPO permits deducing the primary structure of the CPO protein and elucidating the molecular basis of HC in some families.

KEY WORDS: Coproporphyrinogen oxidase; hereditary coproporphyria; cDNA; gene; mutations.

INTRODUCTION

Coproporphyrinogen oxidase (CPO) [EC 1.3.3.3.] is a soluble protein that is localized in the intermembrane space of mitochondria in mammalian cells (Grandchamp *et al.*, 1978; Elder and Evans, 1978). It catalyzes the sixth step in heme biosynthesis, the conversion of the two propionate groups at positions 2 and 4 of coproporphyrinogen III to two vinyl groups, thus producing sequentially harderoporphyrinogen IX and protoporphyrinogen IX (see Kappas *et al.*, 1989 for review). The active enzyme is thought to be a homodimer of approximately 70–74 kDa (Camadro *et al.*, 1986; Bogard *et al.*, 1989).

Hereditary coproporphyria (HC) is an inherited autosomal dominant disease with incomplete penetrance and clinically resembles two other forms of hepatic porphyria, acute intermittent porphyria and variegate porphyria. It is characterized by attacks of neurological dysfunction associated with abdominal pain, hypertension, tachycardia, and peripheral neuropathy. Acute attacks are usually precipitated by drugs, alcohol, or caloric deprivation. Photosensitivity is occasionally present. Clinical manifestations are accompanied by abnormal excretion of products derived from the heme pathway, δ-aminolevulinate and porphobilinogen in urine and coproporphyrinogen III in both feces and urine (Kappas *et al.*, 1989). Enzymatic studies have shown that a mean 50% decrease of CPO activity is found in the cells from coproporphyric patients as well as from asymptomatic carriers of the gene defect (Kappas *et al.*, 1989).

The recent cloning of human cDNAs and of the gene encoding CPO permits deducing the primary structure of the CPO protein and elucidating the molecular basis of HC in some families.

PRIMARY STRUCTURE OF HUMAN CPO

Sequencing of human cDNAs (Taketani *et al.*, 1994; Martasek *et al.*, 1994a) and of the CPO gene (Delfau-Larue *et al.*, 1994) revealed the amino acid sequence of the protein. A comparison of the human protein with those from various sources including, mouse (Kohno *et al.*, 1993), soybean (Madsen *et al.*, 1993), *Saccharomyces cerevisiae* (Zagorec *et al.*, 1988) and *Salmonella typhimurium* (Xu and Elliott, 1994) demonstrated that the protein is highly conserved during evolution (Fig. 1). The amino-terminal part of the enzyme is the most divergent while the carboxy-terminal part is highly conserved. The amino-terminal domains of mammalian CPO is thought to be a presequence that directs the protein into the intermembrane space of mitochondria; in contrast,

[1] INSERM U409 et Association Claude Bernard, Faculté Xavier Bichat, 16 rue Henri Huchard, 75018 Paris, France.

0145-479X/95/0400-0215$07.50/0 © 1995 Plenum Publishing Corporation

Saccharomyces cerevisiae is a cytosolic enzyme (Camadro *et al.*, 1986).

There is still some controversy about the length of the presequence of the human protein. Taketani *et al.*, (1994) proposed that the presequence was made of 31 amino acids on the basis of sequence similarity between a human stretch of amino acid deduced from DNA sequencing and the amino-terminal sequence of the mature bovine enzyme. However, the mature murine enzyme contains an additional stretch of 21 amino acids as its NH_2 terminus as compared to the bovine enzyme (Martasek *et al.*, 1994a). Depending on whether the amino-terminal sequence of the human enzyme is deduced from sequence similarity with the bovine protein or the murine protein, the human mature protein would consist of either 323 (Taketani *et al.*, 1994) or 344 (Martasek *et al.*, 1994a) residues. The extremity of the human mature enzyme needs to be sequenced to clarify this point.

Furthermore, there is also some controversy

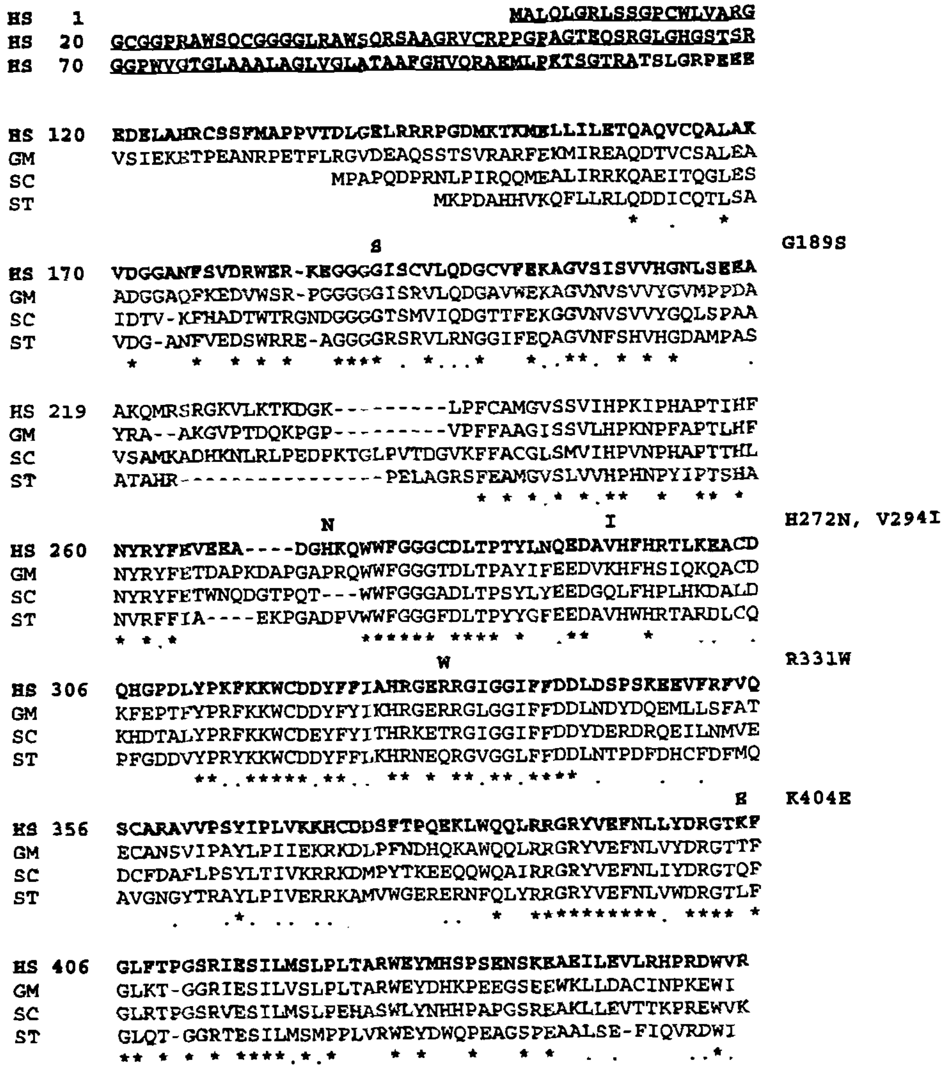

Fig. 1. Alignment of predicted amino acid sequences from human (HS), soybean (GM), *S. cerevisiae* (SC), and *Salmonella typhimurium* (ST) CPXs. The Clustal V software was used to align the sequences. The amino-terminal part of the soybean CPO which does not show any similarity with the human protein is not represented. Stars indicate identical amino acids and dots indicate conservative substitutions. The deduced amino acid sequence of the putative CPO mitochondrial presequence is underlined. Amino acids encoded by allelic variants are represented above the human sequence. Two of these variations are common polymorphisms (H272N, V294I); the three other variations are mutations which modify the enzymatic activity.

about the position of the initiating codon. Taketani *et al.* (1994) suggested that the translational start was situated 31 codons upstream of the putative NH_2 terminus of the mature protein. However, sequence analysis of the first exon of the human gene showed that the cloned cDNAs were not full length and that the first in-frame AUG lies within the first exon, 300 bp upstream from the previously described putative initiating codon (Delfau-Larue *et al.*, 1994). It is most likely that the most 5′ AUG encodes the initiating methionine since it is an almost perfect match to the Kozak consensus sequence with a purine at position −3 and a G at position +4 (Kozak, 1984). By comparison with the putative N-terminal sequence of the mature protein based on similarity with the murine enzyme, it is possible to identify a long presequence of 110 amino acids which shares some characteristics with presequences of other proteins situated in the intermembrane space of mitochondria (Glick *et al.*, 1992). The amino terminal portion of these presequences resembles a matrix-targeting signal with a high density of positively charged residues; the carboxy-terminal portion contains a hydrophobic stretch and functions as the intermembrane space targeting domain.

Portions of the human cDNA encoding the putative mature enzyme have been expressed in *E. coli* and are active (Taketani *et al.*, 1994; Martasek *et al.*, 1994a).

ORGANIZATION OF THE GENE CODING FOR HUMAN CPO

Southern analysis of human restriction fragments using a cDNA probe suggested that a single gene is present in the human genome (Martasek *et al.*, 1994a). This gene has been cloned and DNA sequencing of the genomic clones indicated that the human CPO gene spans about 14 kb and consists of seven exons and six introns. The introns vary in size from 269 bp to 5 kb and they all have consensus sequences (C/T) AG–exon–GT at their boundaries.

Multiple transcriptional initiation sites were located using primer extension and RNase protection assays from various mRNA sources, suggesting that a single promoter is active but differentially regulated in erythropoietic and nonerythropoietic cells (Delfau-Larue *et al.*, 1994).

Computer-assisted analysis of the promoter region revealed a very GC rich region (77% in a 350 nt region) containing many potential cis-acting regulatory elements. There was no consensus sequence for TATA or CAAT boxes, as it is frequently the case for promoters of constitutively expressed genes. However, six Sp1 binding sites are situated upstream of the major initiation sites, and two additional recognition sequences are found in exon 1. Furthermore, four GATA sites and the CACCC boxes are evident. Sp1 binding sites are most often present in the promoter region of constitutively or widely expressed genes, while GATA binding sites are commonly found in combination with either CACCC boxes or Sp1 binding sites in the promoter region of genes specifically expressed or up-regulated in erythroid cells (Faisst and Meyer, 1992; Mignotte *et al.*, 1989). Indeed, the level of CPO transcripts had been found to increase during the erythroid differentiation of mouse erythroleukemic cells in culture (Khono *et al.*, 1993).

It is interesting to note that the CPO gene contains two polyadenylation signals separated by 125 bases from each other. The cDNA isolated by Taketani *et al.* (1994) from a human placental library displays a poly A stretch 90 bases downstream of the first polyadenylation signal while two independent clones that we isolated from a foreskin fibroblast cDNA library included the two polyadenylation signals (Delfau-Larue *et al.*, 1994). These differences in the 3′ end of the cDNAs between placenta and fibroblast may reflect alternative usage of different polyadenylation signals. Whether or not this usage is regulated by cell differentiation, as has been described for ferrochelatase (Chan *et al.*, 1993), remains to be investigated.

MOLECULAR ABNORMALITIES OF THE CPO GENE IN COPROPORPHYRIA

Four mutations have been found so far in patients with coproporphyria; two of these mutations were present at the homozygous state in patients with rare variant forms of coproporphyria, while two other mutations were detected at the heterozygous state in patients with the usual form of the disease. In addition, three neutral DNA polymorphisms were detected in the CPO gene. Two of these polymorphisms result in conservative substitution of amino acids in the protein (Martasek *et al.*, 1994b; Fujita *et al.*, 1994). All known allelic variants at the protein level are shown in Fig. 1. The amino acid numbering is from the putative initiating methionine, and the

actual number is increased by 100 as compared to published results (Martasek *et al.*, 1994b; Fujita *et al.*, 1994) because in these studies the initiating methionine was probably misplaced, as discussed above.

The first mutation to be identified at the CPO locus was found in a patient previously diagnosed as a homozygous case of coproporphyria (Grandchamp *et al.*, 1977). This patient had a severe form of the disease and constantly excreted massive amounts of coproporphyrin. She was born to first-cousin parents and had a very low residual activity of CPO. Using reverse-transcription, amplification of the cDNA, and direct sequencing of the amplified products, we found a point mutation C to T, resulting in an arginine-to-tryptophane substitution (Martasek *et al.*, 1994b; R331W, Fig. 1). Sequencing cDNAs from both parents demonstrated that they were heterozygous for the base change. Expression studies of normal and mutated cDNAs in a bacterial system demonstrated that this substitution resulted in the synthesis of an unstable protein with a residual catalytic activity. The arginine itself is conserved in mouse (Kohno *et al.*, 1993) and soybean (Madsen *et al.*, 1993) but not in other species. However, residues present at the homologous position are hydrophilic whereas the tryptophan introduced by the mutation is hydrophobic, and this may destabilize the spatial structure of the protein. The finding of an enzyme less stable than normal but still partially active is not surprising in a homozygous patient since it is likely that a complete defect of CPO activity would not be compatible with life. Another amino acid substitution was found in siblings with a variant form of the disease. In 3 sibs (2 boys, 1 girl) with intense jaundice and hemolytic anemia at birth, we previously described a high level of coproporphyrin in the urine and feces (Nordmann *et al.*, 1983). The pattern of fecal porphyrin excretion was atypical because the major porphyrin was harderoporphyrin. Homozygosity was suggested by the fact that the level of lymphocyte corproporphyrinogen III oxidase was 10% of controls in the sibs and 50% of normal in both parents (who showed only mild abnormalities of porphyrin excretion). The mutant enzyme showed abnormal kinetics. cDNA sequencing revealed an A-to-G transition which led to the replacement of a lysine residue by glutamic acid at position 404 (K404E, Fig. 1) in the mutated protein (Lamoril *et al.*, 1995). Expression of the protein encoded by the mutant cDNA was obtained in *E. coli*, and enzymatic studies showed a decreased affinity of the abnormal enzyme for its substrates, accounting for an accumulation of coproporphyrin and harderoporphyrin. These abnormalities reproduced those previously described from enzymatic studies performed on lymphocyte lysates of the patients. Although the amino acid substitution does not involved a highly conserved residue, the introduction of a positively charged amino acid in a highly conserved region may affect the enzyme structure in such a way that the affinity for the tetrapyrroles is diminished.

In patients with the more usual heterozygous form of coproporphyria, the causal mutations have been identified in two cases.

One patient was of Czech origin. He repeatedly presented neurological symptoms with paresis and was hospitalized for 6 months following treatment with barbiturates before the correct diagnosis was established. Sequencing of amplified products from genomic DNA showed a G-to-A transition at position −1 of a splice donor site in exon 6 (IVS6-01). This base substitution affected the splicing and resulted in the skipping of exon 6 during the processing of the primary transcript (Delfau-Larue *et al.*, 1994).

Fujita *et al.* (1994) studied another patient of Caucasian origin with typical signs of HC. cDNA sequencing revealed a single base substitution G-to-A which resulted in an amino acid substitution (G189S, Fig. 1). The mutant enzyme had a residual activity of less than 5% of the normal one when expressed in *E. coli*.

All the mutations were searched for and not detected in other patients with HC, suggesting that there is a high degree of allelic heterogeneity of the disease as has been documented in other porphyrias.

Acute attacks of HC are often precipitated by additional factors to the CPO deficiency, and most carriers of the gene defect remain asymptomatic. Early detection of gene carriers is important in the prevention of acute attacks as they can be advised to avoid precipitating factors (Kappas *et al.*, 1989). So far, a screening method in HC families is based on the determination of CPO activity in lymphocytes. However, there is some overlap between CPO activities in controls and HC patients, and the CPO assay in lymphocytes or fibroblasts is only performed in a limited number of specialized laboratories. Therefore DNA analysis may be suitable for a more reliable diagnosis. This is now feasible either through identification of causal mutations in HC families or by linkage analysis using common intragenic polymorphisms. These approaches will facilitate the

detection of asymptomatic carriers and consequently the prevention of acute attacks.

REFERENCES

Bogard, M., Camadro, J. M., Nordmann, Y., and Labbe, P. (1989). *Eur. J. Biochem.* **181,** 417–421.

Camadro, J. M., Chambon, H., Jolles, J., and Labbe, P. (1986). *Eur. J., Biochem.* **156,** 579–587.

Chan, R. Y. Y., Shulman, H. M., and Ponka, P. (1993). *Biochem. J.* **292,** 343–349.

Delfau-Laruc, M. H., Martasek, P., and Grandchamp, B. (1994). *Hum. Mol. Genet.* **3,** 1325–1330.

Elder, G. H., and Evans, J. O. (1978). *Biochem. J.* **172,** 345–347.

Faisst, S., and Meyer, S. (1992). *Nucleic Acids Res.* **20,** 3–26.

Fujita, H., Kondo, M., Taketani, S., Nomura, N., Furuyama., K., Akagi, R., Naga, T., Terajima, M., Galbraith, R. A., and Sassa, S. (1994). *Hum. Mol. Genet.* **3,** 1807–1810.

Glick, B. S., Beasley, E. M., and Schatz, G. (1992). *Trends Biochem. Sci.* 453–459.

Grandchamp, B., Phung, N., and Nordmann, Y. (1977). *Lancet* **ii,** 1345–1349.

Grandchamp, B., Phung, N., and Nordmann, Y. (1979). *Biochem. J.* **176,** 97–102.

Kappas, A., Sassa, S., Galbraith, R. A., and Nordmann, Y. (1989). In *The Metabolic Basis of Inherited Diseases* (Scriver, C. R., *et al.*, eds), McGraw-Hill, New York, pp. 1305–1365.

Kohno, H., Furukawa, T., Yoshinaga, T., Tokunaga, R., and Taketani, S. (1993). *J. Biol. Chem.* **268,** 21359–21363.

Kozak, M. (1984). *Nucleic Acids Res.* **12,** 857–872.

Lamoril, J., Martasek, P., Deybach, J. C., Da Silva, V., Grandchamp, B., and Nordmann, Y. (1995). *Hum. Mol. Genet.*, in press.

Madsen O, Sandal, N. N., and Marcker, K. A. (1993). *Plant Mol. Biol.* **23,** 35–43.

Martasck, P., Camadro, J. M., Delfau-Larue, M. H., Dumas, J. B., Montagne, J. J., de Verneuil, G., Labbe, P., and Grandchamp, B. (1994a). *Proc. Natl. Acad. Sci. USA.* **91,** 3024–3028.

Martasek, P., Nordmann, Y., and Grandchamp, B. (1994b). *Hum. Mol. Genet.* **3,** 477–480.

Mignotte, V., Wall, L., de Boer, E., Grosveld, F., and Romeo, P. H. (1989). *Nucleic Acids Res.* **17,** 37–54.

Nordmann, Y., Grandchamp, B., Verneuil, H. de, and Phung, L. (1983). *J. Clin. Invest.* **72,** 1139–1149.

Taketani, S., Kohno, H., Furukawa, T., Yoshinaga, T., and Tokunaga, R. (1994). *Biochim. Biophys. Acta.* **1183,** 547–549.

Xu, K., and Elliott, T. (1993). *J. Bacteriol.* **175,** 4990–4999.

Zagorec, M., Buhler, J. M., Treich, I., Keng, T., Guarente, L., and Labbe-Bois, R. (1988). *J. Biol. Chem.* **263,** 9718–9724.