BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **226,** 226–230 (1996)
ARTICLE NO. 1337

# Protoporphyrinogen Oxidase: Complete Genomic Sequence and Polymorphisms in the Human Gene[1]

Hervé Puy, Anne-Marie Robréau, Robert Rosipal,[2]
Yves Nordmann,[3] and Jean-Charles Deybach

*Centre Français des Porphyries, INSERM U409, Hôpital Louis Mourier, 92701 Colombes Cedex, France*

Received July 24, 1996

Variegate porphyria (VP) is an autosomal dominant disorder of heme synthesis caused by a partial deficiency of protoporphyrinogen oxidase (PPOX). Human cDNA encoding PPOX has been recently sequenced and the gene has been cloned, assigned to chromosome 1q23, and its exon/intron organization has been characterized. We report here the complete nucleotide sequence of the Human PPOX gene. Including 660 bp of its promotor region, the PPOX gene spans 5.5 kb. Introns vary in size from 84 bp to 507 bp. Two exonic and 3 intronic biallelic sequence variations have been characterized. 

The enzyme protoporphyrinogen oxidase (PPOX; EC 1.3.3.4) acts at the penultimate step in the heme biosynthetic pathway, and catalyzes the oxidation of protoporphyrinogen IX to protoporphyrin IX within the inner mitochondrial membrane (1, 2). A partial deficiency of PPOX activity is responsible for variegate porphyria (VP), an autosomal dominant disease with incomplete penetrance (3-4). Clinical features of VP include both chronic photodermatitis and intermittent neurovisceral episodes which resemble those found in other acute hepatic porphyrias, acute intermittent porphyria and hereditary coproporphyria (2).

Recently, Human PPOX cDNA has been cloned, sequenced and expressed (5-6). It consists of an open reading frame 1431 nucleotides long, encoding a 477 amino-acid protein. The gene has been mapped by FISH to chromosome 1q23, in contrast with a previous linkage assignment on chromosome 14 (7-9). The PPOX gene contains 13 exons and all exon/intron boundaries have been characterized 8 bp upstream and downstream of each exon (7-8). However, a discrepancy about the size of the gene still exists between Taketani et al. (7) and Roberts et al. (8) who respectively estimate the gene to span about 8 kb and 4.5 kb.

In the present study, we describe the complete genomic sequence of the Human PPOX gene and report five intragenic dimorphisms.

## MATERIALS AND METHODS

*Subjects.* Peripheral blood was collected from 50 unrelated control subjects of French Caucasian origin. Genomic DNA was extracted from peripheral blood as previously described (10).

*DNA sequencing.* Oligonucleotides used were obtained from Genset (Paris, France). The regions corresponding to

[1] The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases.

[2] Cooperation: Université Paris VII, France–Univerzita Karlova, Praha, Czech Republic.

[3] To whom correspondence should be addressed: Pr Y. Nordmann, Hôpital Louis Mourier, 92701 Colombes Cedex, France. Fax : (33) 1 47 60 67 03.

Abbreviations: PPOX, Protoporphyrinogen oxidase; VP, Variegate porphyria.



0006-291X/96 $18.00

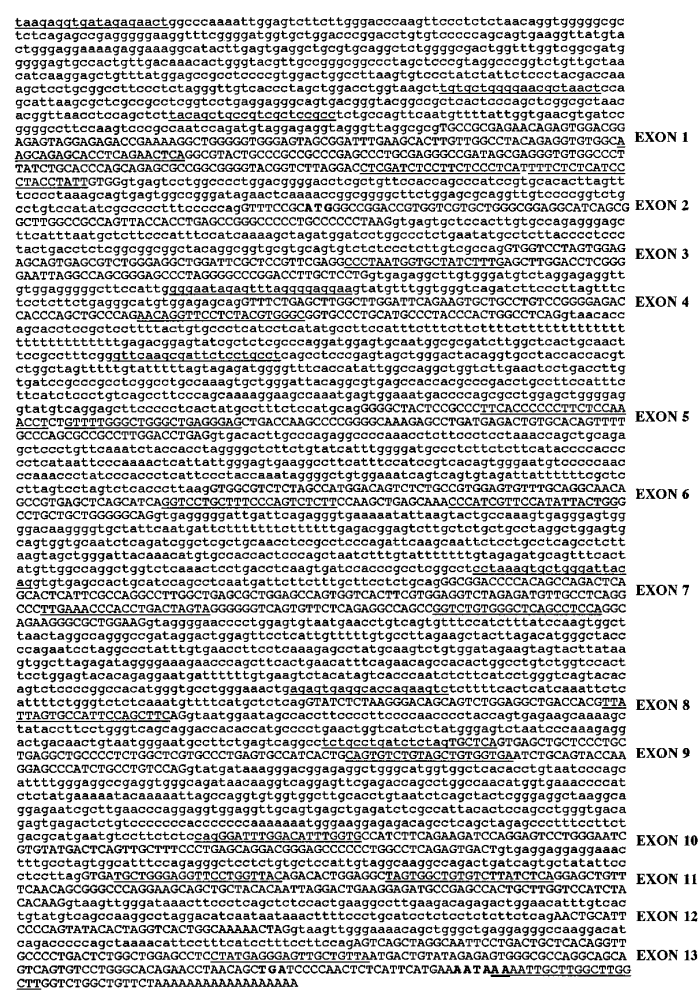


**FIG. 1.** Nucleotide sequence of the PPOX gene. Coding regions and introns are shown by upper- and lowercase letters, respectively. The regions corresponding to amplimers or sequencing primers are underlined and exon locations are notified in the right margin. The initiation and termination codon, in exon 2 and 13, respectively, as well as the polyadenylation signal, are in boldface.

amplimers or sequencing primers were underlined (Figure 1). The PCR products were purified using the Wizard PCR preps DNA purification system (Promega-Biotech, Madison, WI, USA) and direct sequencing was performed in both orientations using a fmol DNA sequencing kit (Promega-Biotech). Samples from five unrelated subjects were studied to determine the normal intronic sequences.

*Polymorphism study*. Sequence variations were first identified by direct sequencing. Allele frequency was obtained from the study of 50 unrelated subjects of Caucasian origin using either DNA restriction or sequence analysis.

*DNA restriction analysis. Bss*SI and *Hinf*I restriction analysis were performed as follows: 10 $\mu$l of amplified product were incubated with 1 U of *Bss*SI or *Hinf*I (New England Biolabs Inc., Beverly, MA, USA) in the appropriate buffer at 37° C for 3 h. *Sfc*I restriction analysis were performed using 10 $\mu$l of amplified product incubated with 1.5 U of *Sfc*I (New England Biolabs Inc.) in the appropriate buffer at 37° C for 1 h.

## RESULTS AND DISCUSSION

The sequence of genomic PPOX is depicted in figure 1 : the 13 exons are located within a 5,5 kb region including 660 bp upstream from the initiation of translation site. This sequence has been deposited in GenBank (Assignment of accession number is in process). We characterize 12

TABLE 1
Structural Organization of the Human ppox Gene

| Exon | Exon size (bp) | Genomic positions* | Intron | Intron size (bp) | Genomic position* |
|---|---|---|---|---|---|
| 1 | 264 | −439/−176 | 1 | 167 | −175/−9 |
| 2 | 95 | −8/87 | 2 | 166 | 88/253 |
| 3 | 135 | 254/388 | 3 | 136 | 389/524 |
| 4 | 116 | 525/640 | 4 | 507 | 641/1147 |
| 5 | 133 | 1148/1280 | 5 | 304 | 1281/1584 |
| 6 | 145 | 1585/1729 | 6 | 416 | 1730/2145 |
| 7 | 191 | 2146/2336 | 7 | 476 | 2337/2812 |
| 8 | 61 | 2813/2873 | 8 | 185 | 2874/3058 |
| 9 | 119 | 3059/3177 | 9 | 384 | 3178/3561 |
| 10 | 111 | 3562/3672 | 10 | 100 | 3673/3772 |
| 11 | 150 | 3773/3922 | 11 | 137 | 3923/4059 |
| 12 | 43 | 4060/4102 | 12 | 84 | 4103/4186 |
| 13 | 183 | 4184/4369 | | | |

* Numbering of the exon/intron junctions starts from the initial base of the initiation codon described by Taketani et al. (7).

introns with relatively short sequences ranging from 84 bp to 507 bp (Table 1). The differences reported in the size of the PPOX gene between Taketani et al. (about 8 kb) and Roberts et al. (about 4.5 kb) can be explained by an incorrect approximation in the size of introns 4, 7, and 9 (7-8) (Table 1). The distribution of each nucleotide along the PPOX gene sequence is homogeneous with 21% A, 28% C, 27% G, and 25% T.

In the course of the genomic sequence determination, five intragenic polymorphic sites were found. The variant nucleotides and the genomic positions are indicated in figure 2. Two polymorphisms are exonic : the −414 A/G is located in exon 1, which has a 5′ untranslated sequence (5). The 3101 G/A polymorphism lies within exon 9. This dimorphism corresponds to the one already described at position 1188 in PPOX cDNA (11). The base change leads to an arginine to histidine substitution at codon 304 (R304H), and creates a new *Bss*SI restriction site.

Three polymorphisms are intronic : the 821 G/A dimorphism in intron 4 was detected by *Hinf*I restriction analysis; the two last polymorphisms are located in intron 6 (1909 A/C, and

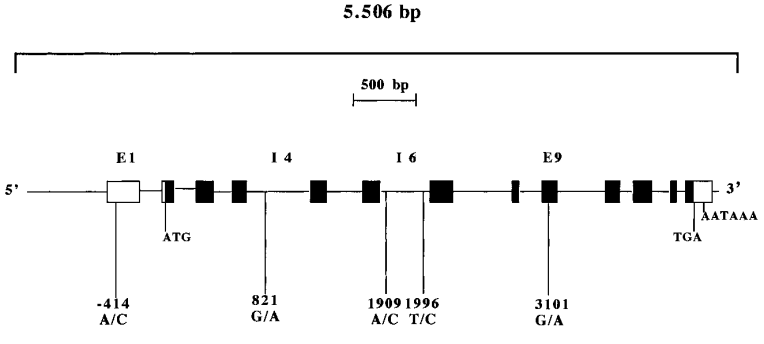


**FIG. 2.** Exon/intron organization of the Human PPOX gene and location of 5 polymorphisms. Numbering of the polymorphisms starts from the initial base of the initiation codon described by Taketani et al. (7). Exons are indicated by boxes: closed and open boxes represent protein-coding and untranslated regions, respectively.

TABLE 2
PPOX Intragenic Polymorphisms: Location and Allele Frequencies in Normal French Caucasians

| Genomic position* | Location | Restriction site | Allele frequency ($N = 100$) |
|---|---|---|---|
| −414 A | Exon 1 | | 0.61 |
| C | | | 0.39 |
| 821 G | Intron 4 | *Hinf* I + | 0.94 |
| A | | − | 0.06 |
| 1909 A | Intron 6 | | 0.75 |
| C | | | 0.25 |
| 1996 T | Intron 6 | *Sfc* I − | 0.96 |
| C | | + | 0.04 |
| 3101 G | Exon 9 | *BssS* I − | 0.59 |
| A | | + | 0.41 |

* Numbering of the polymorphisms starts from the initial base of the initiation codon described by Taketani et al. (7).

1996 T/C). The respective allele frequencies of the five dimorphisms were calculated from the data obtained from 50 different chromosomes using either restriction site analysis (821 G/A, 1996 T/C, 3101 G/A) or sequence analysis (−414 A/G, 1909 A/C) (Table 2). The frequencies of each common alleles range from 0.59 to 0.96%. The five polymorphisms are not in linkage disequilibrium in the population studied.

To date, only 4 different mutations and 1 polymorphism have been reported from VP patient cDNA (11-12). In 3 different VP families, we have reported two mutations found in the coding sequence of PPOX gene: a point insertion of a G at position 1022 of the cDNA, and a missense G232R mutation (11). Meissner et al. described a R168C mutation and a R59W mutation with a high prevalence in South Africans VP patients (12). Therefore, the determination of the complete PPOX genomic sequence could facilitate further studies focused on mutation detection in VP patients. In addition, the findings of five intragenic polymorphic sites that segregate independently, should facilitate the detection of presymptomatic heterozygotes in VP families whose specific PPOX mutations have not yet been identified.

## ACKNOWLEDGMENTS

This work was supported by grants from INSERM (U 409) and from Association Française contre les Myopathies. We thank Mrs Sylvie Robréau and Catherine Guyomard for preparing the manuscript.

## REFERENCES

1. Deybach, J. C., Da Silva, V., Grandchamp, B., and Nordmann, Y. (1985) *Eur. J. Biochem.* **149,** 431–435.
2. Kappas, A., Sassa, S., Galbraith, R. A., and Nordmann, Y. (1995) *in* The Metabolic Basis of Inherited Disease (Scriver, C. R., Beaudet, A. L., and Sly, W. S., Eds.), pp. 2103–2159, McGraw–Hill, New York.
3. Brenner, D. A., and Bloomer, J. R. (1980) *N. Engl. J. Med.* **302,** 765–769.
4. Deybach, J. C., de Verneuil, H., and Nordmann, Y. (1981) *Hum. Genet.* **58,** 425–428.
5. Nishimura, K., Taketani, S., and Inokuchi, H. (1995) *J. Biol. Chem.* **270,** 8076–8080.
6. Dailey, T. A., and Dailey, H. A. (1996) *Protein Science* **5,** 98–105.
7. Taketani, S., Inazawa, J., Abe, T., Furukawa, T., Kobno, H., Tokunaga, R., Nishimura, K., and Inokuchi, H. (1995) *Genomics* **29,** 698–703.
8. Roberts, A. G., Whatley, S. D., Daniels, J., Holmans, P., Fenton, J., Owen, M. J., Thompson, P., Long, C., and Elder, G. H. (1995) *Hum. Mol. Genet.* **4,** 2387–2390.
9. Bissport, S., Hitzeroth, H. W., de Wentzel, D. P., Van den Berg, C. W., Senff, H., Wienker, T. F., and Bender, K. (1988) *Hum. Genet.* **79,** 289–290.

10. Higuchi, R. (1989) *in* PCR Technology, Principles and Applications for DNA Amplification, pp. 31–38, Stockton Press, New York.
11. Deybach, J. C., Puy, H., Robréau, A. M., Lamoril, J., Da Silva, V., Grandchamp, B., and Nordmann, Y. (1996) *Hum. Mol. Genet.* **5,** 407–410.
12. Meissner, P. N., Dailey, T. A., Hift, R. J., Ziman, M., Corrigall, A. V., Roberts, A. G., Meissner, D. M., Kirsch, R. E., and Dailey, H. A. (1996) *Nature Genetics* **13,** 95–97.