View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Volume 286, number 1,2, 55–57 FEBS 09910 July 1991
© 1991 Federation of European Biochemical Societies 00145793/91/$3.50
ADONIS 001457939100637K

Cloning and sequencing the urocanase gene (hutU) from *Pseudomonas putida*

Martin Fessenmaier[1], Rainer Frank[2], János Retey[1] and Carsten Schubert[1]

[1] *Chair of Biochemistry, Institute of Organic Chemistry, University of Karlsruhe, D-7500 Karlsruhe, Germany* and [2] *Center for Molecular Biology, University of Heidelberg, D-6900 Heidelberg, Germany*

Received 23 April 1991; revised version received 13 May 1991

A clone harbouring the entire urocanase gene (hutU) was obtained from a genomic library of *Pseudomonas putida* using oligonucleotide probes synthesised on the basis of known flanking sequences. One subunit of urocanase consists of 556 amino acids and has a molecular mass of 60 771 Da.

Urocanase; hutU Gene; PCR; *Pseudomonas putida*

1. INTRODUCTION

Urocanase from *P. putida*, a homodimer of M_r 122 000, catalyses the addition of water to urocanate to form β-(5-oxoimidazol-4-yl)propionate. Our structural [1] and mechanistic [2] studies on this enzyme prompted us to determine the amino acid sequence of urocanase by deducing it from the base sequence of the corresponding gene. Here we report the isolation of a fragment of the hut (histidine utilizing) gene cluster containing the complete hutU gene. After amplification by PCR the whole sequence of the hutU gene was determined by the method of Sanger [3].

2. MATERIALS AND METHODS

Urocanase, isolated by classical methods [4], was submitted to Edman degradation. N-Terminal sequence analysis was performed in a gas-phase sequencer constructed and operated as described in [5]. The initial yield was 250 pmol and 21 degradation cycles were performed yielding 21 unambiguously identified residues.

Genomic DNA was obtained from a *P. putida* mutant nic II [2] by the cleared lysate method according to [6]. After partial digestion with *Mbo*I and size fractionation, DNA fragments ranging from 10 to 20 kbp were cloned into λ-EMBL3-arms (provided by Promega). In vitro-packaging was performed with an in vitro-packaging extract, also provided from Promega. Recombinant phages have been identified and isolated via plaque-hybridisation with end-labelled oligonucleotides [7]. The oligonucleotides were synthesised with a DNA-synthesizer (Applied Biosystems). After digestion with several different restriction enzymes, DNA was transferred to a nitrocellulose membrane (Schleicher Schüll, BA85) by capillary-blotting [8], and hybridised with oligonucleotides [6]. For autoradiography Fuji FX 100 films were used. PCR was carried out in a Coy-Temp-Cycler using the following profile: 25 cycles were run with denaturation at 94°C, 60 s, annealing at 39.5°C, 30 s, extension at 72°C for 140 s, followed by one single extension step with denaturation at 94°C for 30 s, annealing at 39.5°C for 30 s and extension at 72°C for 5 min. Following a fill-in reaction according to [7] the blunt-ended PCR product was cut with *Sal*I. Sequenase and nucleotides were from United States Biochemical (Cleveland, OH).

Correspondence address: J. Rétey, Lehrstuhl für Biochemie, Institut für Organische Chemie der Universität Karlsruhe, Postfach 6980, D-7500 Karlsruhe 1, Germany. Fax: (49) (721) 608 4290.

3. RESULTS AND DISCUSSION

Urocanase from *Pseudomonas putida* was purified to electrophoretic homogeneity [1] and the N-terminal amino acid residues were determined by Edman degradation (Fig. 1).

A *P. putida* DNA-library in a λ-EMBL 3 vector was prepared according to the 'Protocols and Application Guide (1989/90)' of Promega (Madison, USA). Screening of the library was carried out by plaque hybridization with synthetic oligonucleotides. The following oligonucleotides were prepared:

5′ GGTGTGTTTGTATATTGACCAGG 3′ (1)
5′ AGCGGGCACATGACTTGGTTGC 3′ (2)

(1) is a sequence in the repressor binding region upstream from the hutU structural gene [9], (2) is complementary to a sequence upstream from the hutH structural gene [10] and turned out to be located within the C-terminal region of the hutU gene. End-labelling of the oligonucleotides (1) and (2) with [γ-^{32}P]ATP and T_4-polynucleotide kinase was carried out according to L.G. Davis et al. [7]. The radioactive oligonucleotides were used both for screening the library and for Southern blots.

Positive clones were amplified and the recombinant λ-DNA samples isolated. Digestion with *Bam*HI/*Eco*RI of one sample resulted in four clear-cut fragments of 9, 6, 3.7 and 2.8 kbp. Of these only fragment **2** (6 kbp) hybridized with the radioactive oligomers (1) and (2). The total size of the insert in λ-EMBL 3 is thus 21.5 kbp. Restriction analysis of fragment **2**

Published by Elsevier Science Publishers B.V.

```
1     GTTTGTATATTGACCACGACATACCCGCCCGCGGACGACCGCAGAGGACCTTTCCCGTGACCGACAACAACAAATACCGT
                                                                     T  D  N  N  K  Y  R
81    GACGTTGAAATCCGTGCCCCACGTGGCAACAAGCTGACCGCCAAAAGCTGGCTGACCGAAGCGCCACTGCGCATGCTGAT
       D  V  E  I  R  A  P  R  G  N  K  L  T  A  K  S  W  L  T  E  A  P  L  R  M  L  M
161   GAACAACCTCGACCCACAGGTTGCGGAAAACCCGAAAGAACTGGTGGTGTACGGCGGTATCGGCCGCGCCGCCCGCAACT
        N  N  L  D  P  Q  V  A  E  N  P  K  E  L  V  V  Y  G  G  I  G  R  A  A  R  N
241   GGGAATGCTACGATAAGATCGTCGAAACCCTGACCCGCCTGGAAGACGACGAAACCCTGCTGGTGCAGTCGGGCAAGCCG
      W  E  C  Y  D  K  I  V  E  T  L  T  R  L  E  D  D  E  T  L  L  V  Q  S  G  K  P
321   GTCGGCGTGTTCAAGACCCACAGCAATGCCCCGCGCGTGCTGATTGCCAACTCCAACCTGGTGCCGCACTGGGCCAACTG
       V  G  V  F  K  T  H  S  N  A  P  R  V  L  I  A  N  S  N  L  V  P  H  W  A  N  W
401   GGAACACTTCAACGAACTGGACGCCAAAGGCCTGGCCATGTACGGCCAGATGACCGCCGGCAGCTGGATCTACATCGGCA
        E  H  F  N  E  L  D  A  K  G  L  A  M  Y  G  Q  M  T  A  G  S  W  I  Y  I  G
481   GCCAGGGCATCGTCCAGGGCACTTATGAAACCTTCGTCGAAGCCGGTCGCCAGCACTACGGCGGCACCGTGAAAGCGAAG
      S  Q  G  I  V  Q  G  T  Y  E  T  F  V  E  A  G  R  Q  H  Y  G  G  T  V  K  A  K
561   TGGGTACTCACCGCAGGCCTCGGCGGCATGGGCGGCGCCCAGCCACTGGCTGCTACCCTGGCTGGCGCCTGCTCGCTGAA
       W  V  L  T  A  G  L  G  G  M  G  G  A  Q  P  L  A  A  T  L  A  G  A  C  S  L  N
641   CATCGAATGCCAGCAGAGCCGCATCGACTTCCGTCTGGAAACCCGCTACGTCGACGAGCAGGCCACCGACCTCGACGACG
        I  E  C  Q  Q  S  R  I  D  F  R  L  E  T  R  Y  V  D  E  Q  A  T  D  L  D  D
721   CCCTGGTGCGCATCGCCAAATACACCGCCGAAGGCAAGGCCATCTCCATCGCCCTGCACGGCAACGCTGCCGAAATCCTG
      A  L  V  R  I  A  K  Y  T  A  E  G  K  A  I  S  I  A  L  H  G  N  A  A  E  I  L
801   CCAGAGCTGGTCAAGCGTGGCGTGCGCCCGGACATGGTCACCGACCAGACCAGCGCCCACGACCCGCTCAACGGCTACCT
       P  E  L  V  K  R  G  V  R  P  D  M  V  T  D  Q  T  S  A  H  D  P  L  N  G  Y  L
881   GCCTGCCGGCTGGACCTGGGAGCAGTACCGCGACCGCGCGCAGACCGAACCGGCTGCAGTGGTCAAGGCCGCCAAGCAGT
        P  A  G  W  T  W  E  Q  Y  R  D  R  A  Q  T  E  P  A  A  V  V  K  A  A  K  Q
961   CGATGGCCGTGCACGTGCAGGCCATGCTGGACTTCCAGAAGCAGGGCGTCCCGACCTTCGACTATGGCAACAACATCCGC
      S  M  A  V  H  V  Q  A  M  L  D  F  Q  K  Q  G  V  P  T  F  D  Y  G  N  N  I  R
1041  CAGATGGCCAAGGAAGAGGGCGTGGCCAATGCCTTCGACTTCCCGGGCTTCGTCCCTGCCTATATCCGTCCGCTGTTCTG
       Q  M  A  K  E  E  G  V  A  N  A  F  D  F  P  G  F  V  P  A  Y  I  R  P  L  F  C
1121  CCGCGGCGTAGGCCCTTTCCGCTGGGCGGCGCTGTCTGGCGAGGCCGAGGACATCTACAAGACCGACGCCAAGGTCAAGG
        R  G  V  G  P  F  R  W  A  A  L  S  G  E  A  E  D  I  Y  K  T  D  A  K  V  K
1201  AACTGATCCCCGACGACGCCCACCTGCACCGCTGGCTGGACATGGCCCGCGAGCGCATCAGCTTCCAGGGCCTGCCGGCA
      E  L  I  P  D  D  A  H  L  H  R  W  L  D  M  A  R  E  R  I  S  F  Q  G  L  P  A
1281  CGTATCTGCTGGGTTGGCCTGGGCCTGCGCGCCAAGCTGGGCCTGGCCTTCAACGAAATGGTCCGCAGCGGCGAGCTGTC
       R  I  C  W  V  G  L  G  L  R  A  K  L  G  L  A  F  N  E  M  V  R  S  G  E  L  S
1361  GGCACCGGTCGTGATCGGCCGTGACCACCTGGACTCGGGTTCGGTCTCCAGCCCCAACCGCGAAACCGAAGCCATGCGTG
        A  P  V  V  I  Q  R  D  H  L  D  S  G  S  V  S  S  P  N  R  E  T  E  A  M  R
1441  ATGGCTCGGACGCTGTGTCCGACTGGCCGCTGCTCAACGCCCTGCTGAACACCGCAGGCGGCGCTACCTGGGTGTCGCTG
      D  G  S  D  A  V  S  D  W  P  L  L  N  A  L  L  N  T  A  G  G  A  T  W  V  S  L
1521  CACCATGGTGGTGGCGTGGGCATGGGCTTCTCTCAGCACTCGGGCATGGTCATCGTCTGCGACGGCACCGATGAAGCCGC
       H  H  G  G  G  V  G  M  G  F  S  Q  H  S  G  M  V  I  V  C  D  G  T  D  E  A  A
1601  CGAGCGCATCGCCCGTGTACTGACCAACGACCCAGGGACTGGCGTCATGCGCCACGCCGATGCCGGTTATGACATCGCCA
        E  R  I  A  R  V  L  T  N  D  P  G  T  G  V  M  R  H  A  D  A  G  Y  D  I  A
1681  TCGACTGCGCCAAGGAGCAGGGCCTGGACCTGCCGATGATCACTGGCTGATTGCCAC
      I  D  C  A  K  E  Q  G  L  D  L  P  M  I  T  G  -
```

Fig. 1. Base sequence of the hutU gene from *Pseudomonas putida*. The N-terminal amino acid residues determined by Edman degradation are underlined. The sequence has been deposited at the EMBL Data Bank in Heidelberg, Germany (Acc. No. X58483 PP HUT U).

revealed three *Sal*I and four *Pst*I restriction sites. (The four *Sal*I fragments had 1.5, 1.2, 0.8 and 0.75 kbp). Using the synthetic oligonucleotides (1) and (2) as primers the gene encoding urocanase was amplified by PCR [11]. The PCR product (≈1600 bp) was examined by restriction analysis (*Pst*I, *Sal*I). All these examinations confirmed the identity of the PCR product with the urocanase gene.

Digestion of the blunt-ended PCR product with *Sal*I gave two fragments of 1.0 and 0.6 kbp, respectively. These were cloned into M13mp18 and M13mp19 vectors and three different clones harbouring the larger and seven harbouring the smaller fragment were sequenced in both directions according to Sanger [3]. Alignment of the known N-terminus with the determined base sequence and search for an adequate open reading frame revealed that the hutU gene has a reading frame different from that of the hutH gene and that in our PCR product 97 bases at the C-terminus were missing. This and a sequence overlapping with the PCR product was determined using an M13 mp18 clone harbouring a *Pst*I/*Pst*I fragment of the originally isolated hutU gene. Furthermore, 205 bases at the C-terminus overlap with the non-coding sequence upstream from the hutH gene published by Consevage and Phillips [10]. Thus the intergenic region between hutU and hutH consists of 110 bp. The base sequence of the hutU gene includes 1674 bases (start and stop codons included) and is depicted in Fig. 1. The start codon is GTG, but the N-terminal methionine is removed in the mature urocanase which thus consists of 556 amino acids and has a molecular mass of 60 771 Da. This compares favourably with the value of 61 000 recently determined for the monomer by SDS electrophoresis [1]. In our sequence (Fig. 1) 7 cysteine residues occur. Matherly and Phillips [12] found 6 cysteine residues by titration of the $NaBH_4$-treated with DTNB and by calculating on the basis of a subunit mass of 55 kDa. Correcting for the exact molecular mass of 60 771 Da, 6.65 Cys residues per subunit are obtained, which is very close to the number found in the now determined complete sequence.

Although genes of the hut operon have been cloned from several organisms [10,14-18] our search for urocanase sequences in deposited DNA banks were negative. A search for sequences homologous to the urocanase sequence presented here revealed about 80% homology to a sequence recently determined upstream from the structural gene encoding rubisco in white clover [13]. This 370 base sequence may be an exon of the urocanase gene in white clover. Another sequence showing 62% homology in a ca. 700 bp region was also found [16]. This sequence is situated downstream to the hutH gene of *Bacillus subtilis* and corresponds with high probability to the N-terminal half of the hutU gene.

Acknowledgement: We thank the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie for financial support.

REFERENCES

[1] Klepp, J. and Rétey, J.R. (1989) Eur. J. Biochem. 185, 615-619.

[2] Klepp, J., Fallert-Müller, A., Grimm, K., Hull, W.E. and Rétey, J. (1990) Eur. J. Biochem. 192, 669-676.

[3] Sanger, F., Nicklen, S., Coulson, A. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

[4] Keul, V., Kaeppeli, F., Ghosh, C., Krebs, T., Robinson, J.A. and Rétey, J. (1979) J. Biol. Chem. 254, 843-851.

[5] Gausepohl, H., Trossin M. and Frank, R. (1986) in: Advanced Methods in Protein Microsequence Analysis (Wittmann-Liebold, B. et al. eds.) Springer-Verlag, Berlin-Heidelberg-New York, pp. 149-160.

[6] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual, Cold Spring Harbor University Press, Cold Spring Harbor, NY.

[7] Davis, G.D., Dibner, M.D. and Battey, J.F. (1986) Basic Methods in Molecular Biology, Elsevier, New York.

[8] Southern, E. (1975) J. Mol. Biol. 98, 503-517.

[9] Hu, L., Allison, S.L. and Phillips, A.T. (1989) J. Bacteriol. 171, 4189-4195.

[10] Consevage, M.W. and Phillips, A.T. (1990) J. Bacteriol. 172, 2224-2229.

[11] Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (1990) PCR Protocols: A Guide to Methods and Applications, Academic Press.

[12] Matherly, L.H. and Phillips, A.T. (1980) Biochemistry 19, 5814-5818.

[13] Ellison, N.W., Yu, P.L. and White D.W.R. (1990) Nucleic Acids Res. 18, 4914.

[14] Hu, L. and Phillips, A.T. (1988) J. Bacteriol 170, 4272-4279.

[15] Allison, S.L. and Phillips, A.T. (1990) J. Bacteriol 172, 5470-5476.

[16] Oda, M., Sugishita, A. and Furukawa, K. (1988) J. Bacteriol. 170, 3199-3205.

[17] Boylan, S.A. and Bender, R.A. (1984) Mol. Gen. Genet. 193, 92-98.

[18] Taylor, R.G., Lambert, M.A., Sexsmith, E., Sadler, S.J., Ray, P.N., Mahuran, D.J. and McInnes, R.R. (1990) J. Biol. Chem. 265, 18192-18199.