A highly basic N-terminal extension of the mitochondrial matrix enzyme ornithine transcarbamylase from rat liver

Peter McIntyre, Lynda Graf, Julian Mercer+, Gregory Peterson+, Peter Hudson* and Nicholas Hoogenraad*°

Department of Biochemistry, La Trobe University, Bundoora, Victoria 3083, +Birth Defects Research Institute, Royal Children's Hospital, Parkville, Victoria 3052, and **CSIRO, Division of Protein Chemistry, Parkville, Victoria 3052, Australia*

Received 3 September 1984

We have deduced the amino acid sequence of the N-terminal leader peptide of the mitochondrial enzyme ornithine transcarbamylase from a cDNA clone obtained from a rat liver cDNA library. The sequence is remarkable in being highly basic, having 4 arginine, 3 lysine and 1 histidine with no acidic residues in a total of 32 residues. The leader sequence has no extensive hydrophobic stretches, has 72% homology with the leader peptide of human ornithine transcarbamylase [1], and in terms of its basic character resembles the N-terminal extensions on a number of fungal mitochondrial [2–5] and pea chloroplast [6] proteins. Thus the basic nature of these leader peptides may constitute the signal for mitochondrial import.

Mitochondrial import *Leader sequence* *Urea cycle*

1. INTRODUCTION

Many cytoplasmic precursors to mitochondrial proteins are made as larger polypeptides which are processed to their mature forms after import into the mitochondria [7]. Ornithine transcarbamylase (OTC), one of the urea cycle enzymes in ureotelic animals, is synthesized as a precursor of approximately 40 kDa, including an N-terminal extension of 3.4 kDa [8,9]. Isolated mitochondria import pOTC during which it is proteolytically cleaved to the mature 36 kDa subunit [10–12]. The N-terminal extension of pOTC contains both the proteolytic cleavage sites required for the maturation of the protein in the mitochondrion and also at least some information required for the transport to, and uptake of the enzyme by the mitochondria, since only pOTC and not the mature enzyme can be taken up by mitochondria and the mature enzyme does not affect uptake and processing of pOTC [10]. Recently the amino acid sequences of the N-terminal portion of 4 mitochondrial proteins from fungi and one chloroplast protein from pea have been deduced from nucleotide sequences of the cloned genes [2–6] and the complete sequence of human pOTC has been determined [1]. In each case the sequence contains a number of basic residues and except for ribulose 1,5-bisphosphate carboxylase from pea, which contains a single acidic residue, no other leader sequence contains acidic residues. Thus, from these examples it would appear that a basic N-terminal peptide sequence is required for the import of proteins into mitochondria and chloroplasts. Isoelectric focussing information also has suggested that the leader sequence of pOTC contains a number of basic residues [13] and the sequence analysis reported here confirms this suggestion: 8 out of 32 residues in the pOTC leader peptide were found to be basic.

° To whom correspondence should be addressed

Abbreviations: OTC, ornithine transcarbamylase (EC 2.1.3.3); pOTC, precursor form of OTC; cDNA, DNA complementary to RNA; kb, kilobases; bp, base pairs

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

2. MATERIALS AND METHODS

2.1. *Screening of rat liver cDNA library*

We have previously isolated a 1450 bp OTC cDNA clone and subcloned this into M13 mp9 [14]. A probe specific for the 5′-end of this clone was made by priming on the M13 template using M13 sequencing primer and [α-^{32}P]dATP [15]. Under conditions of limiting dATP (i.e., with no unlabelled dATP), the primed extensions were found to be less than 300 bases in length. We used this 5′ specific probe to screen a rat liver cDNA library in λgt10 [16] and selected a 600 bp clone encoding part of the N-terminal extension of pOTC as well as having sequence overlap with the 1450 bp clone [14]. We subcloned a 44 bp *Taq*1-*Eco*R1 fragment from the 5′-end of the 600 bp clone into M13 mp9 and used single-stranded template to make a hybridisation probe [17] and used this to re-screen the λgt10 cDNA library. Using this approach we isolated a cDNA clone with a 1200 bp insert encoding the entire precursor peptide sequence of pOTC as well as 103 bases of 5′-untranslated sequence. In addition, this cDNA clone overlapped with the DNA sequence from the 5′-end of the 1450 bp clone.

```
          -190       -180       -170       -160       -150       -140
           .          .          .          .          .          .
5'  TGCAACTGAAATGCATTCTCTAGCTTGCCAGTGGCCCCCACTGCCTGCCTGCCTGCAGGAA

        -130       -120       -110       -100           -90          -80
         .          .          .          .             .            .
    CTACTCTAGACATAGATTCCTCCTCCACTCTAGCAAGAG AAG ATG CTG TCT AAT TTG AGG
                                              met leu ser asn leu arg

           -70           -60           -50           -40           -30           -20
            .             .             .             .             .             .
    ATC CTG CTC AAC AAG GCA GCT CTT AGA AAG GCT CAC ACT TCC ATG GTT CGA AAT TTT CGG
    ile leu leu asn lys ala ala leu arg lys ala his thr ser met val arg asn phe arg

           -10          -1 1           10           20           30           40
            .            . .           .            .            .            .
    TAT GGG AAG CCA GTC CAG AGT CAA GTA CAG CTG AAA GGC CGT GAC CTC CTC AAC CTG AAG
    tyr gly lys pro val gln ser gln val gln leu lys gly arg asp leu leu thr leu lys
                           ↑
           50           60           70           80           90          100
            .            .            .            .            .            .
    AAC TTC ACA GGA GAG GAG ATT CAG TAC ATG CTA TGG CTC TCT GCA GAT CTG AAA TTC AGG
    asn phe thr gly glu glu ile gln tyr met leu trp leu ser ala asp leu lys phe arg

          110          120          130          140          150          160
           .            .            .            .            .            .
    ATC AAA CAG AAA GGA GAA TAC TTG CCT TTA TTG CAA GGG AAA TCC TTA GGG ATG ATT TTT
    ile lys gln lys gly glu tyr leu pro leu leu gln gly lys ser leu gly met ile phe

          170          180          190          200          210
           .            .            .            .            .
    GAG AAA AGA AGT ACT CGA ACA AGA CTG TCC ACA GAA ACA CGC TTC GCT   ...3'
    glu lys arg ser thr arg thr arg leu ser thr glu thr arg phe ala
```

Fig.1. Nucleotide and amino acid sequences of the N-terminal extension of rat OTC. The clones from which this sequence was obtained were selected from a cDNA library in bacteriophage λgt10 from liver of Sprague-Dawley rats as described in section 2. The amino terminus of the mature protein (serine) is indicated by the arrow.

2.2. *Sequence analysis of the pOTC cDNA clones*

DNA containing pOTC coding information was sequenced by dideoxy chain termination method after excision from the λgt10 vector and subcloning into M13 mp9 [15].

2.3. *Hybridization of cloned cDNA to RNA fractionated on agarose gels*

RNA from rat liver was fractionated on a 1.2% (w/v) agarose gel containing formaldehyde [18]. The RNA was transferred to nitrocellulose [19] and pOTC mRNA was detected by hybridization with ^{32}P-labelled nick-translated cDNA (175 ng) at a specific activity of 4×10^8 dpm/μg [20].

3. RESULTS AND DISCUSSION

3.1. *Sequence of N-terminal leader of pOTC from rat liver*

Fig.1 shows the nucleotide sequence of the 5′-end of the DNA for pOTC from rat liver and the corresponding amino acid sequence for the 32 amino acids in the leader peptide. Also shown are an additional 210 bases encoding the N-terminus of the mature enzyme. The sequence was obtained from two clones; a 600 bp clone contained sequence corresponding to only 24 of the amino acids of the leader and a 1200 bp clone contained the complete sequence for all 32 amino acids of the leader plus an additional 103 bases 5′- to the initiation codon. Both clones had overlapping sequence with the 1450 bp clone which we have previously shown to correspond to the mature protein [14]. The most striking feature of the leader peptide is the overall positive nature of the extension due to a preponderance of basic residues (8 out of 32) and complete absence of acidic amino acids. This contrasts with the leader sequences of proteins exported from cells, which are highly hydrophobic in character [21]. The sequence of the N-terminus of the mature protein shows quite a different characteristic being less basic; there are 13 basic and 7 acidic residues in the first 70 amino acids.

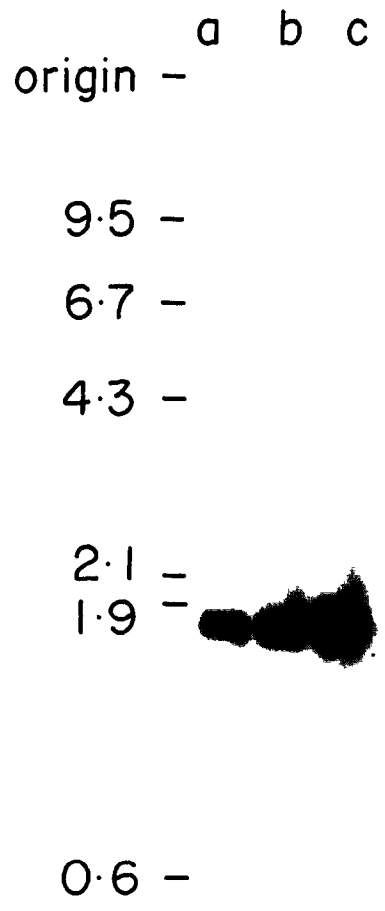

Fig.2. RNA blot analysis using a cDNA pOTC clone. RNA was transferred from 1.2% (w/v) agarose/formaldehyde gels to nitrocellulose filters as described in the text. Total RNA (lane a, 5 μg; b, 7.5 μg; c, 10 μg) from adult rat liver was probed with a [^{32}P]cDNA probe consisting of the complete coding sequence for mature OTC and the N-terminal extension shown in fig.1. Autoradiography was for 96 h. Sizes are given in kb and were obtained from denatured λ*Hin*dIII markers.

3.2. *Properties of the pOTC mRNA*

The pOTC mRNA from Northern blot analysis compared with λ *Hin*dIII markers, is 1700 bases in length (fig.2). The 36 kDa mature OTC subunit would account for approx. 1100 bases and the 3.4 kDa N-terminal extension for the 96 bases found to code for the 32 amino acids of the leader (fig.1). The 3′-untranslated region contains 356 bases ([22] and our unpublished data). This leaves approx. 150 bases of 5′ untranslated mRNA and we present the sequence of 103 nucleotides which

was obtained from the longest cDNA clone in the 5′-untranslated region (fig.1).

3.3. *Comparison of mitochondrial leader sequences*

The findings of a basic N-terminal amino acid sequence on rat OTC, human OTC [1], 4 mitochondrial proteins from fungi and one chloroplast protein from pea [2–6] and that mature mitochondrial proteins cannot compete with precursor forms for mitochondrial uptake [7] suggest that the basic leader peptide may constitute an 'address' for directing the protein to the appropriate mitochondrial compartment. However, there is little overall sequence homology between the leaders of rat pOTC and the proteins from fungi and pea (fig.3) that might provide any insight into the role of the leader in directing the import of these proteins into mitochondria. A comparison of the leader sequences from rat pOTC to that of human shows 72% sequence homology; rat pOTC has 8 basic residues compared with 5 in human pOTC and all 5 basic residues in human pOTC coincide with the same residues in rat pOTC. Perhaps more striking is the much higher degree of homology between amino acid residues at the N-terminus of the mature enzymes; there is 95% homology in the first 70 amino acid residues (fig.4). Although the significance of specific residues in the pOTC leader in the import process has yet to be determined, it may be that the N-terminus of the mature protein also plays a role in the import process.

Besides containing information for import of pOTC into mitochondria, the N-terminal peptide also contains sequence information for the proteolytic maturation of the precursor peptide to its mature form. The proteolytic processing of pOTC has been shown to occur in vitro by a 2-step process via an intermediate of 37 kDa [11,12,23,24]. The first cleavage of the 40 kDa precursor to the 37 kDa intermediate is carried out by a chelator-sensitive protease which has been partially purified from the matrix of rat liver mitochondria [25].

Protein	Organism	Location	Sequence	References
1. proteolipid s.u. of ATP synthase	Neurospora crassa	inner membrane	MASTRVLASRLASQMAASAK VARPAVRVAQVSKRTIQTGS PLQTLKRTQMTSIVNATTRQ AFQKRA↑YSS...	[2]
2. cytochrome c peroxidase	yeast	inter membrane space	MTTAVRLLPSLGRTAHKRSL YLFS(A)$_{10}$TFAYSQSHKRS SSSPGGGSNHGWNNWGKAAA LAS↑TTP...	[3]
3. pre-mRNA processing enzyme	yeast	?	MTVLTAPSGATQLYFHLLRK SPHNRLVVSHQTRRHLMGFV RNALGLD...	[4]
4. 70 Kda membrane protein	yeast	outer membrane	MKSPITRNKTAILATVAATG TAIGAYYYYNQLQQQQQRGK KNTINK...	[5]
5. small s.u. of Rib b-P carbox.	pea	thylakoid space	MASMISSSAVTTVSRASRGQ SAAVAPFGGLKSMTGFPVKK VNTDITSITSNGGRVKC ↑MQ V...	[6]
6. ornithine transcarbamylse	human	matrix	MLFNLRILLNNAAFRNGHNF MVRNFRCGQPLQ↑NKV...	[1]
7. ornithine transcarbamylase	rat	matrix	MLSNLRILLNKAALRKAHTS MVRNFRYGKPVQ↑SQV...	This paper

Fig.3. Comparison of N-terminal amino acid sequences of mitochondrial and chloroplast proteins. Basic residues are underlined. The arrow indicates the position of proteolytic maturation of the precursor, where this is known.

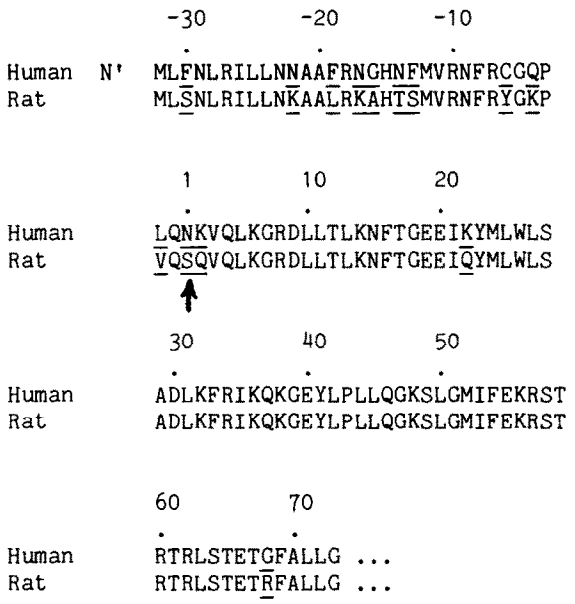

Fig.4. Comparison of amino acid sequence of human [1] and rat ornithine transcarbamylase. The arrow indicates the N-terminus of the mature proteins. Non-homologous residues are underlined.

Since the intermediate produced by this cleavage is 1 kDa larger than the mature protein, it would appear that the most likely cleavage site lies between positions −10 and −7 (fig.1). The final maturation of the peptide involves a second cleavage at the N-terminus of the mature protein: between glutamine and serine at positions −1 and 1 (fig.1). This cleavage would appear to be highly specific since there is no evidence for heterogeneity at the amino terminus of mature OTC, yet in human pOTC the cleavage occurs between glutamine and asparagine [1]. Secondary structure predictions [26] predict an α-helical region for the 20 amino acids at the N-terminus of the leader followed by a β-bend around glycine −5 (fig.4). This glycine residue as well as proline at −3 and glutamine at −1 are conserved in both human [1] and rat pOTC (fig.4) and may be important in inducing a structure in the protein which directs the specific cleavages observed.

ACKNOWLEDGEMENTS

We acknowledge financial support from the Australian National Health and Medical Research Council. We thank Geoff Howlett for supplying a cDNA library from Sprague-Dawley rat liver [16].

REFERENCES

[1] Horwich, A.L., Fenton, W.A., Williams, K.R., Kalousek, F., Kraus, J.P., Doolitle, R.F., Konigsberg, W. and Rosenberg, L.E. (1984) Science 224, 1068–1074.
[2] Viebrock, A., Perz, A. and Sebald, W. (1982) EMBO J. 1, 565–571.
[3] Kaput, J., Goltz, S. and Blobel, G. (1982) J. Biol. Chem. 257, 15054–15058.
[4] Faye, G. and Simon, M. (1983) Cell 32, 77–78.
[5] Hase, T., Reizman, H., Suda, K. and Schatz, G. (1983) EMBO J. 2, 2169–2172.
[6] Cashmore, A.R. (1983) in: Genetic Engineering in Plants (Kossuge, T. et al. eds) pp.29–38, Plenum, New York.
[7] Hay, R., Bohni, P. and Gasser, S. (1984) Biochim. Biophys. Acta 779, 65–87.
[8] Conboy, J.G., Kalousek, F. and Rosenberg, L.E. (1979) Proc. Natl. Acad. Sci. USA 76, 5742–5727.
[9] Mori, M., Miura, S., Tatibana, M. and Cohen, P.P. (1980) J. Biochem. (Tokyo) 88, 1829–1836.
[10] Mori, M., Miura, S., Morita, T., Takiguchi, M. and Tatibana, M. (1982) Mol. Cell Biochem. 49, 97–111.
[11] Kraus, J.P., Conboy, J.G. and Rosenberg, L.E. (1981) J. Biol. Chem. 256, 10739–10742.
[12] Conboy, J.G. and Rosenberg, L.E. (1981) Proc. Natl. Acad. Sci. USA 78, 3073–3077.
[13] Miura, S., Mori, M., Morita, T. and Tatibana, M. (1982) Biochem. Int. 4, 201–208.
[14] McIntyre, P., Mercer, J.F.B., Peterson, M.G., Hudson, P. and Hoogenraad, N. (1984) Eur. J. Biochem., in press.
[15] Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5436–5467.
[16] De Jong, F.A., Howlett, G., Aldred, A.R., Fidge, N. and Schreiber, G. (1984) Biochem. Biophys. Res. Commun. 119, 657–662.
[17] Hu, N. and Messing, J. (1982) Gene 17, 271–277.
[18] Rave, N., Crkvenjakov, R. and Boedtker, H. (1979) Nucleic Acids Res. 6, 3559–3567.
[19] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[20] Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) Proc. Natl. Acad. Sci. USA 72, 1184–1188.
[21] Kreil, G. (1981) Annu. Rev. Biochem. 50, 317–348.
[22] Horwich, A.L., Kraus, J.P., Williams, K., Kalousek, F., Konigsberg, W. and Rosenberg, L.E. (1983) Proc. Natl. Acad. Sci. USA 80, 4258–4262.

[23] Mori, M., Miura, S., Tatibana, M. and Cohen, P.P. (1980) Proc. Natl. Acad. Sci. USA 77, 7044–7048.
[24] Mori, M., Morita, T., Miura, S. and Tatibana, M. (1981) J. Biol. Chem. 256, 8263–8266.
[25] Miura, S., Mori, M., Amaya, Y. and Tatibana, M. (1982) Eur. J. Biochem. 122, 641–647.
[26] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.