Isolation of a human ceruloplasmin cDNA clone that includes the N-terminal leader sequence

Julian F.B. Mercer and Andrew Grimes

Birth Defects Research Institute, Royal Children's Hospital, Flemington Road, Parkville, Victoria, 3052, Australia

Received 29 May 1986

A number of cDNA clones encoding human ceruloplasmin were identified using two mixed oligonucleotide probes. One of these clones was shown by DNA sequence analysis to span from the complete N-terminal leader sequence to 114 amino acids short of the C-terminus. The leader sequence consists of 19 primarily hydrophobic amino acids. Northern blot analysis of RNA from human liver showed two species of ceruloplasmin mRNA; a minor species of 3600 nucleotides and a major one of 4400 nucleotides.

Ceruloplasmin *Leader sequence* *mRNA size* *DNA sequence*

1. INTRODUCTION

Human CP is a blue glycoprotein of about 132 kDa containing 6 copper atoms per molecule, 7–8% carbohydrate and carries 90–95% of the blood plasma copper [1,2]. Takahashi et al. [3] have determined the complete primary structure of CP showing that it consists of a single polypeptide chain of 1046 amino acid residues. CP synthesis and/or secretion is altered by a number of physiological factors including inflammation [4], pregnancy [5] and development [6]. Experimentally plasma CP levels have been increased by administration of copper [7], estrogen [8], glucocorticoids [9] and turpentine [10]. To investigate further the molecular mechanisms involved in the regulation of CP levels, analysis of mRNA levels and the CP gene is required. The purification of CP mRNA by immunoprecipitation of polysomes was reported by Gaitskhoki et al. [11], who showed that this mRNA was translated in a wheat germ protein synthesis system to yield a protein of M_r 84000. This result and other results from this group led them to propose that CP was made as smaller molecular mass precursors and maturation of the protein involves ligation of the two precursor fragments [12,13]. All of this work, however, relied upon the specificity of the CP antibody.

Here we report the isolation of a CP cDNA clone from a human liver cDNA library. Partial DNA sequence analysis confirmed the identity of the clone and showed that it spanned almost the entire coding region from the N-terminal leader sequence to 114 amino acids short of the C-terminus, strongly suggesting that CP is synthesized as a single polypeptide.

Abbreviations: CP, ceruloplasmin, iron (II):oxygen oxidoreductase, EC 1.16.3.1; cDNA, DNA complementary to RNA; kb, kilobases; bp, base pairs

2. MATERIALS AND METHODS

2.1. *Screening of a human liver cDNA library*

Two mixed oligonucleotides predicted from the published amino acid sequence of human CP [3] were synthesized by L. Graf and N. Hoogenraad of LaTrobe University using an Applied Biosystems 381A DNA synthesizer. The oligonucleotides were cleaved from the resin, deprotected and without further purification, end-labelled with T_4-polynucleotide kinase and [γ-^{32}P]ATP to a specific activity $>10^8$ cpm/μg [14]. The oligonucleotides were used to screen

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

replicate filters of a λgt10 library of cDNA from human liver RNA [15]. The hybridizations were carried out in 6 × SSC (0.9 M NaCl, 0.09 M Na citrate, pH 8.0), 0.125% skim milk powder, 0.1% Triton X-100, initially at 50°C cooling to 20°C over 18 h. The filters were washed twice for 30 min with 6 × SSC at 20°C and twice for 15 min with 6 × SSC at 37°C and exposed to Kodak XAR5 film for 18 h using an intensifying screen. Ten positive phages were picked, purified and the DNA was isolated [16]. After digestion with *Eco*RI the insert sizes were analyzed by electrophoresis on 1% agarose gels.

2.2. *DNA sequence analysis*

The *Eco*RI fragments from the clones were subcloned into M13 mp18 and the DNA sequence was determined using the dideoxy-chain termination method [17].

2.3. *Northern blot analysis of RNA from human livers*

Total RNA was isolated from human livers as described by Wake and Mercer [18]. The RNA (10 μg) was fractionated on a 1.8% (w/v) agarose gel containing formaldehyde [19], and transferred to nitrocellulose [20]. The CP mRNA was detected by hybridization with the 0.69 kb *Eco*RI fragment from λhCP.1 (100 ng) which had been labelled by nick translation [20] to a specific activity of $>10^8$ dpm/μg with [α-^{32}P]dCTP. The RNA ladder from Bethesda Research Laboratories was used to provide size markers.

3. RESULTS AND DISCUSSION

3.1. *Isolation of a human CP cDNA clone*

Using the complete amino acid sequence of human CP [3], we chose two amino acid regions as suitable for oligonucleotide probes for screening purposes. Fig.1 shows that the two sequences are close together (amino acids 667–672 and 690–695) thus maximizing the chance that positive clones would give a signal with both probes. In both cases a mixture of sixteen 17-mers were synthesized as described in section 2 and used without further

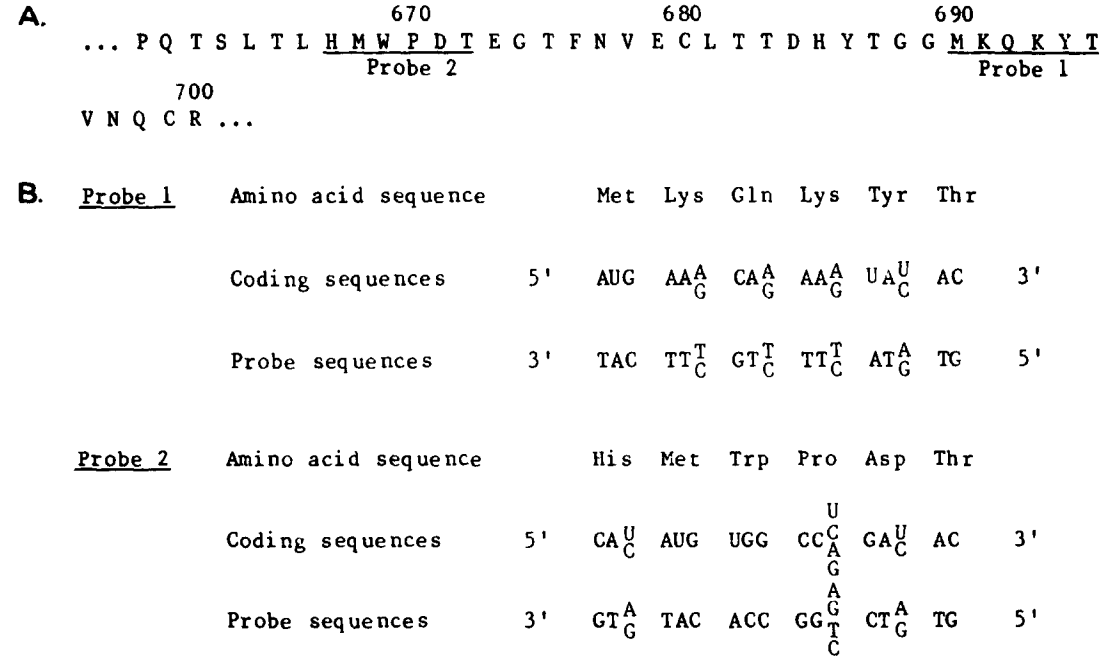


Fig.1. Probes used for CP screen. (A) A portion of the amino acid sequence of CP from amino acid 660 to 700 is shown in single letter code. The amino acids used to deduce the sequences of the oligonucleotide probes are underlined. (B) Sequences of the mixed oligonucleotide probes used to screen for CP clones. The amino acid sequence and deduced coding sequences are shown in conventional orientation, the mixed probes are written 3′–5′.

purification. A screen of 80000 plaques of a human cDNA library in λgt10 using the end-labelled probes yielded 42 clones that gave a positive signal with each probe mixture. Ten of these were isolated and one (λhCP.1) was found to contain three *Eco*RI fragments of 1.74, 0.69 and 0.43 kb, giving a total insert size of 2.86 kb. The other clones all had a common 0.69 kb band and four included the 0.43 kb band.

3.2. *Anatomy and partial DNA sequence analysis of CP cDNA*

To confirm the identity of the CP cDNA, DNA sequences from four regions of the clone including each *Eco*RI fragment were obtained and the predicted amino acid sequences compared with the human CP amino acid sequence [3]. Each *Eco*RI fragment yielded sequences that matched exactly with the published amino acid sequence (fig.2B), confirming the identity of the clone and showing that the cDNA clone had not resulted from a fusion of unrelated cDNA fragments during the cloning procedures. Fig.2A shows the regions of the protein encoded by the three *Eco*RI fragments. The 1.74 kb fragment extends beyond the N-terminus of the mature protein to include the complete leader sequence of CP (see below). The 0.43 kb fragment terminates 114 amino acids from the C-terminus. The *Eco*RI sites are shown above the line and the amino acid sequences used to predict the primer structure are shown as P1 and P2 in fig.2A. The fact that all the clones terminated at *Eco*RI sites found in λhCP.1 suggests that during the library construction, the *Eco*RI site methylation was incomplete, resulting in digestion of the cDNA during the linker removal. Unfortunately, none of the clones contained any fragments extending beyond the 3′-*Eco*RI site so we were unable to obtain sequences from this region.

3.3. *Structure of the leader sequence of CP*

The predicted leader sequence of CP was obtained by sequencing both strands of the cDNA. This was achieved by subcloning the 140 nucleotide *Eco*RI-*Nco*I fragment from the 5′-end of the 1.74 kb *Eco*RI fragment into M13 mp8 thus also allowing the sequence to be read back from the *Nco*I site. The leader peptide is a highly hydrophobic region of 19 amino acids (fig.2B, sequence 1). It is likely that the methionine at −19 is the initiator since there is an in-frame termination signal three codons upstream. In addition, the methionine is followed by a lysine, also characteristic of signal peptides [21]. The N-terminal lysine of the mature protein is preceded by an alanine, as is often found in this position [22].

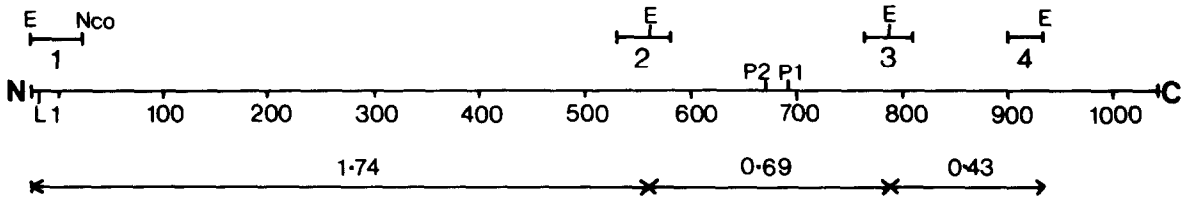
3.4. *Characterization of human CP mRNA*

Liver RNA from 3 individuals was electrophoresed in formaldehyde agarose gels and transferred to nitrocellulose. As shown in fig.3, probing with nick translated 0.69 kb *Eco*RI fragment showed two mRNA species in each sample, the larger species was predominant and ran slightly ahead of the 28 S rRNA. The same two bands were observed with poly(A)$^+$ RNA (fig.3B, lane 1) and no signal was detected with total RNA from brain stem (fig.3B, lane 2) showing that the upper band was not due to hybridization with 28 S rRNA. Using RNA standards, we estimate the size of the two CP mRNAs to be approx. 4400 and 3600 nucleotides, both significantly larger than that required to encode the precursor CP (3195 nucleotides). We are not sure why the signal given by the sample in lane 2 was less intense. From the ethidium bromide stained gel, all three samples were loaded equivalently. The mRNA in lane 2 was somewhat degraded as evidenced by the smearing at lower molecular masses, however, this does not seem sufficient to account for the differences. Both samples in lanes 1 and 3 were from accident victims, whereas the child in lane 2 died from methylene tetrahydrofolate reductase deficiency. The samples in lanes 1 and 2 were from 7-year-old males. The origin of sample 3 is not recorded except as accident victim. As yet we have no information about the effects of illness or age differences on CP mRNA levels.

3.5. *Conclusions*

The isolation of a CP clone that contains virtually the entire coding region, and the detection of two mRNA species in humans, which are both large enough to encode CP, refutes the hypothesis of Puchkova et al. [12] and Prozorovski et al. [13] that CP is made as two smaller molecular mass proteins that are ligated to produce the mature molecule. It is possible that the antibody used by

A.



B.

```
                                -19                                 -10
1                               Met Lys Ile Leu Ile Leu Gly Ile Phe Leu Phe Leu Cys
    ATT CCA GCC TGA GAA GAA ATG AAG ATT TTG ATA CTT GGT ATT TTT CTG TTT TTA TGT
                                 1                                  10
    Ser Thr Pro Ala Trp Ala Lys Glu Lys His Tyr Tyr Ile Gly Ile Ile Glu Thr Thr
    AGT ACC CCA GCC TGG GCG AAA GAA AAG CAT TAT TAC ATT GGA ATT ATT GAA ACG ACT
                                20
    Trp Asp Tyr Ala Ser Asp His Gly...
    TGG GAT TAT GCC TCT GAC CAT GGG...

    530                                 540
2...Ile Phe Thr Gly Leu Ile Gly Pro Met Lys Ile Cys Lys Lys Gly Ser Leu His Ala
...ATA TTC ACT GGG CTT ATT GGG CCA ATG AAA ATA TGC AAG AAA GGA AGT TTA CAT GCA
       550                                  560
    Asn Gly Arg Gln Lys Asp Val Asp Lys Glu Phe Tyr Leu Phe Pro Thr Val Phe Asp
    AAT GGG AGA CAG AAA GAT GTA GAC AAG GAA TTC TAT TTG TTT CCT ACA GTA TTT GAT
             570                                 580
    Glu Asn Glu Ser Leu Leu Leu Glu Asp Asn Ile Arg Met...
    GAG AAT GAG AGT TTA CTC CTG GAA GAT AAT ATT AGA ATG...

                            770                                 780
3...Tyr Arg Gln Tyr Thr Asp Ser Thr Phe Arg Val Pro Val Glu Arg Lys Ala Glu Glu
...TAT CGG CAG TAT ACT GAT AGC ACA TTC CGT GTT CCA GTG GAG AGA AAA GCT GAA GAA
                                790                                 800
    Glu His Leu Gly Ile Leu Gly Pro Gln Leu His Ala Asp Val Gly Asp Lys Val Lys
    GAA CAT CTG GGA ATT CTA GGT CCA CAA CTT CAT GCA GAT GTT GGA GAC AAA GTC AAA
                                   810
    Ile Ile Phe Lys Asn Met Ala Thr...
    ATT ATC TTT AAA AAC ATG GCC ACA...

    900                                 910
4...Leu Phe Leu Val Phe Asp Glu Asn Glu Ser Trp Tyr Leu Asp Asp Asn Ile Lys Thr
...CTG TTT CTA GTT TTT GAT GAG AAT GAA TCT TGG TAC TTA GAT GAC AAC ATC AAA ACA
       920                                  930
    Tyr Ser Asp His Pro Glu Lys Val Asn Lys Asp Asp Glu Glu Phe...
    TAC TCT GAT CAC CCC GAG AAA GTA AAC AAA GAT GAT GAG GAA TTC...
```

Fig.2. (A) Anatomy of the ceruloplasmin clone. The solid line represents the CP protein sequence with the N-terminus (N) on the left and C-terminus on the right (C), the amino acid numbers with the N-terminus of the mature protein taken as 1 are shown beneath; the letter L below the line indicates the beginning of the leader sequence. Also below the line the regions spanned by the three *Eco*RI fragments (1.74, 0.69 and 0.43 kb) of λhCP.1. Above the line the letters E and Nco refer to the *Eco*RI and *Nco*I sites, respectively. P2 and P1 are the positions of the amino acid sequences used to generate the probes shown in fig.1. The lines labelled 1–4 are the regions of DNA sequence shown in B. (B) Partial DNA sequence of the CP clone. The numbers 1–4 refer to the regions indicated in (A). Amino acids are numbered from the mature N-terminus based on the sequence of Takahashi et al. [3]. In most cases the sequence is compiled from several separate readings of one strand, but the leader sequence (sequence 1) was read from both strands. The *Eco*RI sites are underlined.

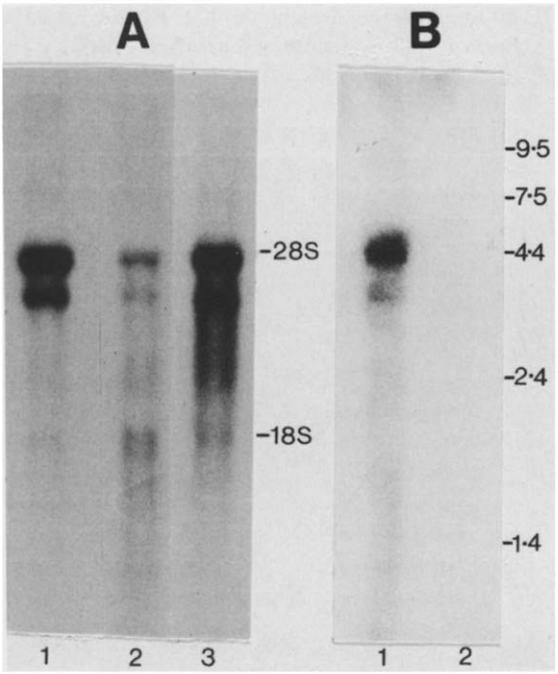


Fig.3. Northern blot analysis of RNA from human liver. Total RNA (10 μg) or poly(A)$^+$ RNA (400 ng) from three individuals was transferred to nitrocellulose from a 1.8% (w/v) agarose gel containing formaldehyde as described in the text and probed with the 0.69 kb *Eco*RI fragment from λhCP.1. (A) Autoradiography was for 72 h. Lanes: 1, hepatic RNA from a 7-year-old accident victim; 2, hepatic RNA from a 7-year-old child who died from methylene tetrahydrofolate reductase deficiency; 3, hepatic RNA from an accident victim. The position of the 18 S and 28 S ribosomal RNA species are indicated. (B) Autoradiography was for 18 h. Lanes: 1, hepatic poly(A)$^+$ RNA from an accident victim; 2, total RNA from human brain stem. The sizes of the RNA standards in kb are indicated on the right.

these workers was not monospecific for CP. Our cDNA clone will be useful for isolation of a rat cDNA clone which will allow us to study the response of CP mRNA to various agents such as copper and hormones.

ACKNOWLEDGEMENTS

We are grateful to Geoff Howlett for supplying the human cDNA library and to Linda Graf and Nick Hoogenraad for synthesis of the oligonucleotides. We thank Greg Peterson for advice about DNA sequencing, Brendon Murray for suggesting conditions for oligonucleotide probe hybridization, Ed Harris for encouragement to proceed with this work and David Danks for advice and criticism of the manuscript. This work was supported in part by a grant from the Australian National Health and Medical Research Council.

REFERENCES

[1] Cousins, R.J. (1985) Physiol. Rev. 65, 238–309.
[2] Frieden, E. (1980) in: Biological Roles of Copper, pp.93–124, Elsevier/North-Holland (Ciba Foundation), Amsterdam, New York.
[3] Takahashi, N., Ortel, T.L. and Putnam, F.W. (1984) Proc. Natl. Acad. Sci. USA 81, 390–394.
[4] Sweeney, H.C. (1967) J. Dental Res. 46, 1171–1176.
[5] Flynn, A. (1980) in: Inflammatory Diseases and Copper (Sorenson, J.R.J. ed.) pp.17–30, Humana Press, Clifton, NJ.
[6] Evans, G.W., Myron, D.R., Cornatzer, N.F. and Cornatzer, W.E. (1970) Am. J. Physiol. 218, 298–300.
[7] Linder, M.C., Houle, P.A., Isaacs, E., Moor, J.R. and Scott, L.E. (1979) Enzyme 24, 23–35.

[8] Evans, G.W., Cornatzer, N.F. and Cornatzer, W.E. (1970) Am. J. Physiol. 218, 613–615.
[9] Weiner, A.L. and Cousins, R.J. (1983) Proc. Soc. Exp. Biol. Med. 173, 486–494.
[10] Meyer, B.J., Meyer, A.C. and Horwitt, M.K. (1958) Am. J. Physiol. 194, 581–584.
[11] Gaitskhoki, V.S., L'vov, V.M., Monakhov, N.K., Puchkova, L.V., Schwartzman, A.L., Frolova, L.J., Skobeleva, N.A., Zagorski, W. and Neifakh, S.A. (1981) Eur. J. Biochem. 115, 39–44.
[12] Puchkova, L.V., Gaitskhoki, V.S., Monakhov, N.K., Temchenko, L.T. and Neifakh, S.A. (1981) Mol. Cell. Biochem. 35, 159–169.
[13] Prozorovski, V.N., Rashkovetski, L.G., Shavlouski, V.B., Vasiliev, V.B. and Neifakh, S.A. (1982) Int. J. Peptide Prot. Res. 19, 40–53.
[14] Khorana, H.G. and Vizsolyi, J.P. (1961) J. Am. Chem. Soc. 83, 675–685.
[15] Nagashima, M., Morris, G., Howlett, G., Fidge, N. and Schreiber, G. (1985) J. Lipid Res., in press.
[16] Davis, R.W., Botstein, D. and Roth, J.R. (1980) Advanced Bacterial Genetics, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.
[17] Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5436–5467.
[18] Wake, S.A. and Mercer, J.F.B. (1985) Biochem. J. 228, 425–432.
[19] Lehrach, H., Diamond, D., Wozney, J.M. and Boetker, M. (1977) Biochemistry 16, 4743–4751.
[20] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[21] Rigby, P.W.J., Diekmann, M., Rhodes, S. and Berg, P. (1977) J. Mol. Biol. 113, 237–251.
[22] Kreil, G. (1981) Annu. Rev. Biochem. 50, 317–348.