AN 88 AMINO ACID LONG C-TERMINAL SEQUENCE OF HUMAN LACTOTRANSFERRIN

M.-H. METZ-BOUTIGUE, J. JOLLÈS, J. MAZURIER*, G. SPIK*, J. MONTREUIL* and P. JOLLÈS

Laboratoire des Protéines, Université de Paris V, 45 rue des Saints-Pères, 75270 Paris Cedex 06 and **Laboratoire de Chimie Biologique et Laboratoire associé au CNRS, Université des Sciences et Techniques de Lille-I, 59655 Villeneuve d'Ascq Cedex, France*

Received 13 April 1982

1. Introduction

We have reported the alignment of all the cyanogen bromide fragments (FI–FVII) of human lactotransferrin also called lactoferrin [1], which is a glycoprotein of ~80 000 M_r consisting of a single polypeptide chain to which 2 carbohydrate groups are attached [2]. We characterized N- and C-terminal domains and described 70% of the sequence [1]. The C-terminal cyanogen bromide fragment of lactotransferrin, called FIII [1], was claimed to be constituted by 111 ± 16 amino acids with an M_r of 12 700 ± 1900, but so far it has not yet been studied in detail. Once more only the elucidation of the sequence was able to indicate the exact relative molecular mass. The establishment of the primary structure of FIII and the prediction of its secondary structure reported in this note allowed a new series of comparisons with human serum transferrin [3,4] and hen ovotransferrin [5] including the situation of possible glycosylation sites.

2. Materials and methods

Sephadex G-25 fine and G-50 fine were obtained from Pharmacia. Trypsin (twice crystallized) was purchased from Worthington, the *Staphylococcus aureus* V_8 protease from Miles, the mouse submaxillary gland protease from Boehringer and citraconic anhydride from Sigma. All reagents (analytical grade) were purchased from Prolabo or Merck, except those employed for the sequencer which were obtained from Pierce or SDS (Marseilles).

Human lactotransferrin was obtained according to [6]. CNBr cleavage was done as in [7]. The CNBr-treated apolactotransferrin was reduced as in [8] and alkylated with iodoacetamide. CNBr fragment FIII was purified as in [1,9].

Partial acid hydrolysis was carried out in 0.03 M HCl during 2 h at 110°C [10]. Citraconylation [11] and tryptic digestion (6 h) were performed according to [1]. Digestion by *Staphylococcus aureus* V_8 protease (18 h) [12] or by mouse submaxillary gland protease (2 h) [13] were achieved at 37°C in 0.05 M ammonium bicarbonate (pH 8) with an enzyme:substrate ratio of 1:50.

The methods used for the purification of the peptides, for their analysis, their automated degradation in a Beckman Sequencer 890 C and their digestion by carboxypeptidases A and B (5 min) have been described [1]. The reported secondary structure was predicted as in [14].

3. Results

The sequence of the C-terminal CNBr fragment FIII of human lactotransferrin (residues 1–88; fig.1) was established according to the following strategy:

(a) Determination of a long N-terminal sequence by automated Edman degradation (41 amino acids);
(b) Isolation of tryptic peptides after citraconylation;
(c) Alignment of the latter by studying peptides obtained by partial acid hydrolysis or by digestion with *Staphylococcus aureus* V_8 protease or mouse submaxillary gland protease;
(d) Digestion by carboxypeptidases A and B.

3.1. *N-terminal sequence*

A long N-terminal sequence (residues 1–41) was determined by automated Edman degradation [1,15] (fig.1). The amino acid located at position 32 was

Published by Elsevier Biomedical Press
00145793/82/0000–0000/$02.75

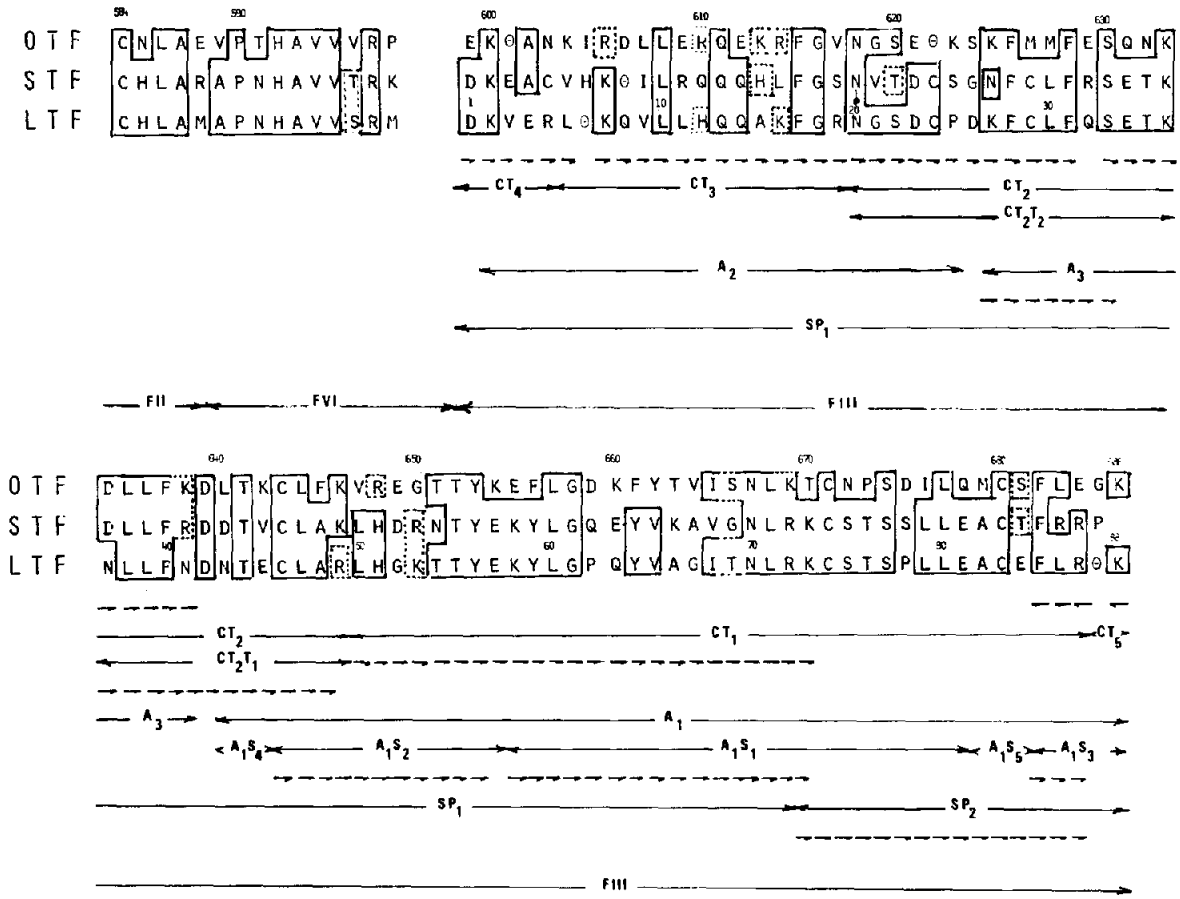

Fig.1. The amino acid sequence of the C-terminal CNBr fragment FIII of human lactotransferrin (LTF) (residues 1–88). Comparison with the homologous C-terminal sequences of human serum transferrin [3] (STF) and hen ovotransferrin [5] (OTF) (residues 599–686). The sequences preceding the C-terminal fragments are also indicated: (→) amino acid determined by automated Edman degradation; (←) amino acid determined by carboxypeptidase digestion; (CT) tryptic peptides from citraconylated fragment FIII; (T) tryptic peptides; (A) peptides obtained after partial acid hydrolysis; (S and SP) peptides obtained by digestion with *Staphylococcus aureus* V_8 protease or mouse submaxillary gland protease, respectively; (N) carbohydrate carrying asparagine residue. The homologous residues (identical residues, ; conservative changes,) are boxed; (Θ) deletion.

identified in another series of experiments (see section 3.3).

3.2. *Tryptic digestion of citraconylated fragment FIII*

After filtration of the tryptic digest on Sephadex G-25, 5 fragments were characterized: CT1 (residues 50–87); CT2 (residues 20–49), CT3 (residues 6–19); CT4 (residues 1–5); and CT5 (Lys) (fig.1). The Arg–Lys bond (72–73) was not cleaved probably because of the citraconylated lysine residue. Two fragments, CT1 and CT2, were studied. A long N-terminal sequence of CT1 (residues 50–73) was determined with the sequencer (fig.1). The presence of an Asn–Gly sequence at the N-terminal end of fragment CT2 avoided its automated degradation. A similar observation was reported in [16]. An asparaginylglycine sequence undergoes cyclization in an acid medium [17] and the resulting cyclic imide is not susceptible to Edman degradation. After decitraconylation, fragment CT2 was submitted to tryptic digestion. Only 2 peptides were obtained, CT2T1 (residues 37–49) and CT2T2 (residues 20–36). Cleavage of the Lys–Phe (residues 27–28) bond was not observed, and we suggest that the presence of residues 25 and 26 (Pro and Asp) prevent the action of trypsin. The sequence of peptide CT2T1 was automatically determined (fig.1).

3.3. *Partial acid hydrolysis of fragment FIII*

After partial acid hydrolysis of fragment FIII, 3 peptides were characterized on Sephadex G-25: A1 (residues 43–88); A2 (residues 2–25); and A3 (residues 27–41). The presence of an Asn–Thr sequence at the N-terminal end of fragment A1 did not allow its automated degradation. To determine its primary structure, peptide A1 was digested by *Staphylococcus aureus* V_8 protease. After filtration on Sephadex G-25, 5 peptides were characterized: A1S1 (residues 58–81), A1S2 (residues 46–57), A1S3 (residues 85–88), A1S4 (residues 43–45) and A1S5 (residues 82–84). Peptide A1S2 allowed to determine the overlapping sequence between peptides CT2T1 and CT1 (fig.1). Peptide A3 was submitted to automated Edman degradation in order to determine the residue located at position 32 (glutamine) (fig.1).

3.4. *Digestion of fragment FIII by mouse submaxillary gland protease*

Following digestion of fragment FIII the cleavage of a unique Arg–X bond was observed (Arg 72–Lys 73). After chromatography on Sephadex G-50, automated Edman degradation of the shorter peptide SP2 (residues 73–88) allowed to complete the sequence (residues 73–87) of fragment FIII and to join peptides A1S5 and A1S3.

3.5. *Digestion of fragment FIII by carboxypeptidases A and B*

After 5 min the enzymes liberated the following amino acids (yields % between parentheses): Phe (39), Leu (50), Arg (57), Lys (64). The sequence of fragment FIII was thus established.

3.6. *Predicted secondary structure*

The C-terminal sequence of lactotransferrin is composed of CNBr fragments FII → FVI → FIII [1]. The results reported in this note allowed to predict the secondary structure of a large C-terminal fragment (103 amino acids) with its α-helixes, β-sheets and β-turns (fig.2).

4. Discussion

The comparison of the C-terminal sequence of lactotransferrin with those of serum and ovotransferrins allowed to establish remarkable sequence homologies (64% and 42%, respectively). The secondary structures of the C-terminal fragments of the 3 transferrins were also closely homologous with their cystine residues in register (residues 584, 643, 671 and 680; see fig.1) (fig.2).

Asparagine residue 20 of fragment FIII belongs to a code sequence N G S and is situated in a β-turn (fig.1,2); nevertheless it is not glycosylated. The same absence of glycosylation occurs in the homologous ovotransferrin sequence whereas in the corresponding serum transferrin sequence the asparagine residue carries a sugar group.

Finally an unusually high occurrence of paired amino acids in fragment FIII should be mentioned.

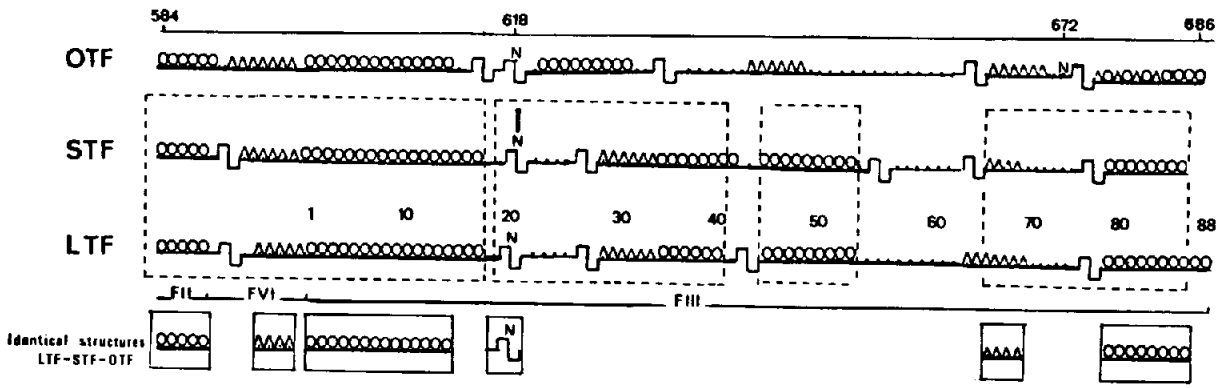

Fig.2. Predicted secondary structures of the C-terminal moieties of human lactotransferrin (LTF), serum transferrin (STF) and ovotransferrin (OTF). Residues are represented in their respective conformational state: helical ( ), β-sheet ( ), section which has the potentiality to adopt either an α-helical or a β-sheet structure ( ), coil (–), β-turn ( ). ( ) Carbohydrate carrying asparagine residue (N). The identical structures are boxed (comparison LTF–STF ; comparison LTF–STF–OTF). For the numbering of the residues of OTF and LTF, see fig.1.

Acknowledgements

This research was supported by the CNRS (ER no. 102 and LA no. 217) and the INSERM (Unité U-116). The excellent technical assistance of Mr Ly Quan Le and Mr J.-P. Decottignies is gratefully acknowledged.

References

[1] Metz-Boutigue, M.-H., Mazurier, J., Jollès, J., Spik, G., Montreuil, J. and Jollès, P. (1981) Biochim. Biophys. Acta 670, 243–254.
[2] Montreuil, J. and Spik, G. (1975) in: Proteins of Iron Storage and Transport in Biochemistry and Medicine (Crichton, R. R. ed) pp. 27–37, Elsevier Biomedical, Amsterdam, New York.
[3] McGillivray, R. T. A., Mendez, E. and Brew, K. (1977) in: Proteins of Iron Metabolism (Brown, E. B. et al. eds) pp. 133–141, Grune and Stratton, New York.
[4] Lineback-Zins, J. and Brew, K. (1982) Proc. Natl. Acad. Sci. USA in press.
[5] Jeltsch, J.-M. and Chambon, P. (1982) Eur. J. Biochem. 122, 291–295.
[6] Chéron, A., Mazurier, J. and Fournet, B. (1977) CR Acad. Sci. Paris 284, 585–588.
[7] Gross, A. and Witkop, J. (1962) J. Biol. Chem. 237, 1856–1860.
[8] Crestfield, A. M., Moore, S. and Stein, W. H. (1963) J. Biol. Chem. 238, 622–627.
[9] Mazurier, J. (1980) Thesis, University of Lille.
[10] Schultz, J., Allison, H. and Grice, M. (1962) Biochemistry 1, 694–698.
[11] Maley, G. F., Bellisario, R. L., Guarino, D. U. and Maley, F. (1979) J. Biol. Chem. 254, 1288–1295.
[12] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[13] Lévy, M., Fishman, L. and Schenkein, I. (1970) Methods Enzymol. 19, 672–681.
[14] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 222–245.
[15] Metz-Boutigue, M.-H., Jollès, J., Jollès, P., Mazurier, J., Spik, G. and Montreuil, J. (1980) Biochim. Biophys. Acta 622, 308–314.
[16] Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H. and Walsh, K. A. (1972) Biochemistry 11, 4493–4502.
[17] Bornstein, P. and Balian, G. (1977) Methods Enzymol. 47, 132–145.