# THE N-TERMINAL SEQUENCE OF HUMAN LACTOTRANSFERRIN: ITS CLOSE HOMOLOGY WITH THE AMINO-TERMINAL REGIONS OF OTHER TRANSFERRINS

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

*Laboratoire des Protéines, Université de Paris V, 45 rue des Saints-Pères, F-75270 Paris Cedex 06, and **Laboratoire de Chimie Biologique de l'Université des Sciences et Techniques de Lille-I, B.P. No. 36, F-59650 Villeneuve d'Ascq*

Received 24 July 1976

## 1. Introduction

Human lactotransferrin (also called lactoferrin) [1–3] consists of a single polypeptide chain of molecular weight 76 500 [4,5] and possesses two carbohydrate groups which are conjugated to the protein by an asparaginyl-*N*-acetylglucosamine linkage [6]: their structures were established by Spik et al. [7,8]. However, the present knowledge of the amino acid sequence of human lactotransferrin remains very limited: some cyanogen bromide fragments have been purified and analyzed [9] and six short cysteic acid containing peptides [10] as well as short tryptic and chymotryptic glycopeptides [7,8,11] have been sequenced. Finally glycine has been identified as the *N*-terminal amino acid [10].

The present paper is dealing with the establishment of the N-terminal sequence of human lactotransferrin and of the corresponding cyanogen bromide fragment. Extensive similarities were thus demonstrated between the N-terminal regions of the ovo-, serum- and lactotransferrins.

## 2. Materials and methods

Human lactotransferrin was obtained according to Chéron et al. [12]. The protein was reduced following the procedure of Crestfield et al. [13]; iodoacetamide was employed for the alkylation step. The preparation and purification of the cyanogen bromide (CNBr) fragments were reported by Mazurier et al. [9 and unpublished results].

The N-terminal amino acids of the protein and of the CNBr-fragment were identified by the dansylation procedure. Automated Edman degradation was carried out in a Socosi Sequencer Model PS-100, by the quadrol double-cleavage method [14]; dithioerithrytol was added in 1-chlorobutane (10 mg/l). The thiazolinones were converted into the corresponding phenylthiohydantoin (PTH)-amino acids; the latter were characterized, as previously reported [15], by thin-layer chromatography, by gas-liquid chromatography (Beckman GC 45 chromatograph) and by chromatography of the free amino acids (Technicon amino acid Autoanalyzer) regenerated from the PTH-amino acids (150°C; 24 h; 6 M HCl containing 1/2000 2-mercaptoethanol).

The chymotryptic digestion of CNBr-fragment F-V′ was performed during 24 h at 37°C in 0.05 M $NH_4HCO_3$ with an enzyme/substrate ratio of 1/20. The chymotryptic peptides were purified by paper electrophoresis (Whatman No. 1) at pH 6.5 (pyridine–water–acetic acid, 100 : 900 : 4, v/v/v) and 50 V/cm.

## 3. Results

### 3.1. *N-terminal sequence of human lactotransferrin*

By the dansylation procedure, Gly was characterized as the N-terminal amino acid of the protein. By automated Edman degradation, the ten first amino acids could be identified (table 1).

### 3.2. *The N-terminal cyanogen bromide fragment of human lactotransferrin: characterization of the first 24 amino acids of the protein*

7 CNBr-fragments (F-I to F-VII) have so far been

Table 1

N-terminal sequences of human lactotransferrin and of the cyanogen bromide fragment F-V′ (N-terminal moiety of the intact lactotransferrin molecule). Methods of identification were as follows: (a) +, PTH-derivative determined by thin-layer chromatography; cysteine was determined as PTH-S-carboxamidomethylcysteine; (b) PTH-amino acid determined by gas–liquid chromatography, yield %; (c) amino acid determined with an Autoanalyzer after regeneration, % of recovery.

| | Human lactotransferrin | | | | CNBr-fragment F-V′ | | | |
|---|---|---|---|---|---|---|---|---|
| Sequence position | a | b | c | Residue | a | b | c | Residue |
| 1 | Gly | Gly (18) | Gly (20) | Gly | Gly | Gly (15) | Gly (18) | Gly |
| 2 | Arg* | | Arg (4) | Arg | | | Arg (7) | Arg |
| 3 | Arg* | | Arg (4.2) | Arg | | | Arg (7) | Arg |
| 4 | Arg* | | Arg (4.8) | Arg | | | Arg (9) | Arg |
| 5 | Arg* | | Arg (4) | Arg | | | Arg (9) | Arg |
| 6 | Ser | Ser (11) | | Ser | Ser | Ser (17) | | Ser |
| 7 | Val | Val (12) | Val (12) | Val | Val | Val (24) | Val (19) | Val |
| 8 | Gln | | Glu (8) | Gln | Gln | | Glu (14) | Gln |
| 9 | Trp | Trp (5) | | Trp | Trp | Trp (15) | | Trp |
| 10 | Cys | | Cys (1) | Cys | Cys | Cys | Cys (2) | Cys |
| 11 | | | | | Ala | Ala (18) | Ala (14) | Ala |
| 12 | | | | | Val | Val (18) | Val (14) | Val |
| 13 | | | | | Ser | Ser (13) | | Ser |
| 14 | | | | | Gln | | Glu (10) | Gln |
| 15 | | | | | Pro | Pro (5) | Pro (3) | Pro |
| 16 | | | | | Glu | | Glu (4) | Glu |
| 17 | | | | | Ala | Ala (7) | Ala (5) | Ala |
| 18 | | | | | Thr | Thr | | Thr |
| 19 | | | | | Lys | | Lys (4) | Lys |
| 20 | | | | | Cys | Cys | Cys(0.5) | Cys |
| 21 | | | | | Phe | Phe (2) | Phe (2) | Phe |
| 22 | | | | | Gln | | Glu (1) | Gln |
| 23 | | | | | Trp | Trp (0.5) | | Trp |
| 24 | | | | | Gln | | | Gln |

* Characterized by thin-layer chromatography during 3.5 h (instead of 0.5 h) in solvent 7 (15) and staining with the Sakaguchi reagent.

Table 2

Mobility m (at pH 6.5) values, yields and structures of the chymotryptic peptides of fragment F-V′ of human lactotransferrin. $m$ = 0 for Gly; $m$ = +1 for Arg; m = −1 for $CySO_3H$.

| Peptide | $m$ | Yield (%) | Structure |
|---|---|---|---|
| C-1 | +1.15 | 35 | Gly–Arg–Arg–Arg–Arg |
| C-2 | 0 | 35 | Ser–Val–Gln–Trp |
| C-3 | −0.25 | 50 | Cys–Ala–Val–Ser–Gln–Pro–Glu–Ala–Thr–Lys–Cys–Phe |
| C-4 | −0.56 | 15 | PyrGlu*–Trp |
| C-5 | +0.58 | 32 | Gln–(Asx, Ser, Ala, Val, $Arg_2$)–HSer* |
| C-6** | +0.33 | 34 | Gln–Trp–Gln–(Asx, Ser, Ala, Val, $Arg_2$)–HSer* |

* PyrGlu, pyrrolidonecarboxylic acid; HSer, homoserine.
** C-6 = C-4 + C-5.

Table 3
Homology of the amino-terminal regions of human lactotransferrin (this study), human serum transferrin [16] and hen ovotransferrin [18] -0-: deletion.

| | 1 | | | | | | | | 10 | | | | | | | | | | 20 | | | | | | | 25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lactotransferrin | Gly– | Arg– | Arg– | Arg– | Arg– | Ser– | Val– | Gln– | Trp– | Cys– | Ala– | Val– | Ser– | Gln– | Pro– | Glu– | Ala– | Thr– | Lys– | Cys– | Phe– | Gln– | Trp– | Gln– | – | |
| Serum transferrin | Val– | Pro– | Asp– | Lys– | -0- | – | Thr– | Val– | Arg– | Trp– | Cys– | Ala– | Val– | Ser– | Glu– | His– | Glu– | Ala– | Thr– | Lys– | Cys– | Gln– | Ser– | Phe– | Arg– | Asp– | |
| Ovotransferrin | Ala– | Pro– | Pro– | Lys– | -0- | – | Ser– | Val / Ile | Arg– | Trp– | Cys– | Thr– | Ile– | Ser– | Ser– | Pro– | Glu– | Gln– | Lys– | Lys– | Cys– | Asn– | Asn– | Leu– | Arg– | Asp– | |

isolated [9]. During further purification of fragment F-V, a peptide F-V′ was obtained by filtration on Biogel P6 with 10% acetic acid as eluent followed by chromatography on Dowex 1 X 2 with a pyridine (10%)–acetic acid buffer of pH 8.4.

Its amino acid composition was established (residues/mole between parentheses): Asp (1), Thr (1), Ser (3), Glu (5), Pro (1), Gly (1), Ala (3), Cys (2), Val (3), Phe (1), Trp (2), Lys (1), Arg (6), H.Ser (1). F-V′ was submitted to automated Edman degradation. Table 1 indicates that this fragment constitutes the N-terminal region of human lactotransferrin; a sequence constituted by the first 24 amino acids was established, and this result is reported for the first time.

Fragment F-V′ was further submitted to a chymotryptic digestion. Table 2 indicates the mobilities, yields and structures of six chymotryptic peptides (C-1 to C-6) in the order as they appear in the sequence. Peptide C-6 was composed of peptides C-4 + C-5; in C-4, glutamine was characterized as pyrrolidonecarboxylic acid. An unusual chymotryptic split was observed after the fourth arginine residue of F-V′. The study of peptides C-1 to C-6 corroborated the data established by automated Edman degradation.

## 4. Discussion

### 4.1. *Comparison of the amino terminal regions of 3 transferrins*

From the N-terminal sequence of human serum transferrin which we previously published [16] and present data concerning human lactotransferrin, the alignments proposed in table 3 could be constructed; they demonstrate strikingly the extensive similarity between the N-terminal regions of the two proteins. We point out more particularly the satisfactory alignment of the cysteinyl residues present in both sequences: this was not possible when the N-terminal sequence of serum transferrin proposed by Sutton and Brew [17] was considered. Our results were in accordance with the highly conservative behaviour of cysteinyl residues during evolution, a fact illustrated by several other published phylogenetic protein data. Table 3 also includes William's recently established long N-terminal sequence of hen ovotransferrin quoted, without details, by Bluard-Deconinck et al. [18] despite the fact that it differs from previously published data by Elleman and Williams [19]; indeed the *N*-terminal sequence of hen transferrin presents again a close homology with the corresponding sequences of the two human transferrins.

## Acknowledgments

This research was supported in part by the C.N.R.S. (ER. 102 and LA 217), the I.N.S.E.R.M. (group U-116) and the D.G.R.S.T. (contract 75.7.1334).

## References

[1] Montreuil, J. and Mullet, S. (1960) C. R. Acad. Sci. Paris 250, 1736–1737.
[2] Montreuil, J., Tonnelat, J. and Mullet, S. (1960) Biochim. Biophys. Acta 45, 413–421.
[3] Masson, P. L., Heremans, J. F., Prignot, J. J. and Wauters, G. (1966) Thorax 21, 538–544.
[4] Querinjean, P., Masson, P. L. and Heremans, J. F. (1971) Eur. J. Biochem. 20, 420–425.
[5] Spik, G. (1971) Ann. Nutrit. 25, A81–A90.
[6] Spik, G., Monsigny, M. and Montreuil, J. (1966) C. R. Acad. Sci. Paris 263, 893–896.
[7] Spik, G., Vandersyppe, R., Montreuil, J., Han, K. K. and Tetaert, D. (1974) FEBS Lett. 38, 213–216.
[8] Spik, G., Vandersyppe, R., Fournet, B., Bayard, B., Charet, P., Bouquelet, S., Strecker, G. and Montreuil, J. (1975) Actes du Colloque C.N.R.S. sur la Méthodologie des Glycoconjugués, Villeneuve d'Ascq, 20–27 Juin 1973, C.N.R.S. éd., Paris, 483–499.
[9] Mazurier, J., Spik, G. and Montreuil, J. (1974) FEBS Lett. 48, 262–265.
[10] Bluard-Deconinck, J. M., Masson, P. L., Osinski, P. A. and Heremans, J. F. (1974) Biochim. Biophys. Acta 365, 311–317.
[11] Graham, I. and Williams, J. (1975) Biochem. J. 145, 263–279.
[12] Chéron, A., Mazurier, J., Fournet, B. and Montreuil, J. Biochim. Biophys. Acta (in press).
[13] Crestfield, A. M., Moore, S. and Stein, W. H. (1963) J. Biol. Chem. 238, 622–627.
[14] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.

[15] Jollès, J., Schoentgen, F., Hermann, J., Alais, Ch. and Jollès, P. (1974) Eur. J. Biochem. 46, 127–132.
[16] Boutigue, M. H., Jollès, J., Charet, P., Montreuil, J. and Jollès, P. (1976) Biochimie 58, 891–892.
[17] Sutton, M. R. and Brew, K. (1974) FEBS Lett. 40, 146–148.
[18] Williams, J. in Bluard-Deconinck, J. M., Osinski, P., Querinjean, P., Masson, P. and Heremans, J. F. (1975) Proc. 1st Int. Conf. 'Solid Phase methods in Protein Sequence Analysis' (Laursen ed.) Boston, p. 203–209.
[19] Elleman, J. C. and Williams, J. (1970) Biochem. J. 116, 515–535.