THE SEQUENCE OF RESIDUES 1-26 OF HUMAN SERUM TRANSFERRIN

Michael R. SUTTON and Keith BREW

Department of Biochemistry, University of Leeds, 9 Hyde Terrace, Leeds LS2 9LS, England

Received 2 November 1973

1. Introduction

The transferrins are a group of proteins, widely distributed in vertebrates, that function in the binding and transport of ferric iron. Human serum transferrin consists of a single polypeptide chain of molecular weight 77 000 [1, 2] which contains two similar binding sites for iron and other metals [3–6]. Some workers have speculated that transferrin may have originated during phylogeny as the result of the doubling in size of the structural gene for a smaller precursor, which contained a single metal binding site (see ref. [7]).

An investigation of the primary structure of human transferrin was initiated in this laboratory to evaluate this hypothesis, and to open a path for chemical studies of the binding sites of the protein. As the polypeptide chain of transferrin is of a large size for direct sequence studies, we have cleaved the protein with cyanogen bromide to produce seven peptide fragments, which have been isolated in high yield and characterised with respect to molecular weights, amino acid compositions and N-terminal groups [8]. The sequences of these fragments are being assembled separately and will be used to reconstruct the complete amino acid sequence of the original molecule.

Only one of the CNBr fragments (designated CN-6) possesses the same amino terminal residue (valine) as the intact transferrin molecule [1, 9]. This 26-residue fragment must therefore be derived from the amino terminal region of the protein. We report here the complete sequence of CN-6, assembled from the structure of overlapping tryptic and chymotryptic peptides.

2. Materials and methods

2.1. *Materials*

Iron-free human serum transferrin was purchased from the Sigma Chemical Company Ltd. It was found to be substantially homogenous (>95%) by polyacrylamide gel electrophoresis and used without further purification.

The preparation of the S-carboxamidomethyl derivative of CN-6 will be described elsewhere [8].

Bio-Gel P4 (200–400 mesh) and AG50X4 (30–35 μ) were obtained from Bio-Rad Laboratories and carboxymethyl cellulose (CM-52) from Whatman. TPCK-treated trypsin was purchased from Worthington Biochemical Corporation and chymotrypsin (twice recrystallised) and carboxypeptidases A and B (DFP treated) from Sigma.

2.2. *Methods*

Procedures used for enzymic digestions, peptide separations and acid hydrolysis have been described previously [10, 11]. Amino acid analyses were carried out with a Bio-cal BC-200 automatic amino acid analyzer. In one case a group of peptides were fractionated by ion exchange chromatography with a 0.9 X 10 cm column of carboxymethyl cellulose at 45°C. The column was equilibrated with 25 mM ammonium acetate pH 5.0 and the sample dissolved and applied in this buffer. Peptides were eluted with a linear gradient composed of 250 ml each of 25 mM ammonium acetate pH 5.0 and 200 mM ammonium acetate pH 5.0.

Sequence analysis of peptides was performed by the dansyl chloride–Edman degradation procedure as described previously [10]. The position of tryptopha-

North-Holland Publishing Company – Amsterdam

Table 1
Amino acid compositions, N-terminal residues and yields of tryptic and chymotryptic peptides from the amino-terminal CNBr-fragment of human transferrin

Peptide	T1	T2	T3	T4	T5	C1	C2	C3	CN-6
CM-Cys			1.00(1)	0.76(1)			1.65(2)		2.0(2)
Asp	1.00(1)				1.14(1)	1.23(1)		1.17(1)	2.0(2)
Thr[a]		1.10(1)	0.87(1)			0.90(1)	0.94(1)		1.8(2)
Ser[a]			1.03(1)	1.15(1)			1.28(2)		1.7(2)
Glu			1.67(2)	1.37(1)			2.90(3)		3.1(3)
Pro	0.93(1)					0.61(1)			1.1(1)
Ala			2.07(2)				2.25(2)		2.0(2)
Val	1.23(1)	1.01(1)	0.77(1)			1.74(2)	0.85(1)		2.5(3)
Phe				0.78(1)					0.95(1)
Trp			[1][b]			[1][b]			ND
His			1.23(1)		0.94(1)		0.98(1)	0.75(1)	1.6(2)
Lys	0.97(1)		1.00(1)			1.04(1)	0.72(1)		1.8(2)
Arg		0.92(1)		0.77(1)		0.74(1)		0.94(1)	1.9(2)
N-terminal	Val	Thr	Trp	Ser	Asp	Val	Ala	Arg	Val
Yield (%)	31	4	4	5	10	10	6	4	

Figures in parentheses after the molar ratios represent the assumed number of residues.

[a] Not corrected for partial destruction.
b From amino acid sequence.

nyl residues, which are not determined by this procedure, were inferred from the loss of absorption at 280 nm by the peptide after the appropriate Edman degradation step, or when in a C-terminal position in a peptide, identified as the dansyl derivative by omitting acid hydrolysis.

The amide contents of peptides were deduced from their electrophoretic mobilities at pH 6.5 [12] or by total enzymic digestion with aminopeptidase M [10].

3. Results and discussion

The amino acid compositions and N-terminal residues of CN-6 and the peptides obtained from it by digestion with trypsin or chymotrypsin are given in table 1. Sequence analysis of these peptides has permitted the deduction of a complete and unambiguous sequence for CN-6 as shown in fig. 1.

The amino terminal position of T1 and C1 is clearly indicated by their possession of N-terminal valine. T5, the only tryptic peptide devoid of lysine and arginine, must occupy the C-terminal position in the fragment.

From the electrophoretic mobilities of peptides the only incomplete amide assignment was with peptide T3, where a residue each of glutamic acid and glutamine were found to be present. An aliquot of this peptide was subsequently subjected to five Edman degradation cycles and then subjected to amino acid analysis after incomplete hydrolysis with aminopeptidase M. The following composition was obtained: $Thr_{0.77}$, $Ser+Gln_{0.27}$ $Glu_{1.0}$ $Ala_{0.69}$ $His_{0.23}$ $Lys_{0.35}$. Residue 12 is therefore identified as glutamine and residue 14 as glutamic acid.

It is interesting to note that in a separate investigation of the cyanogen bromide fragments of human transferrin, Bezkorovainy and Grohlich [13] obtained a fragment similar to CN-6, which they designate C-1-b. The fragment has an N-terminal residue of valine, but shows in hydrolysates the presence of leucine and tyrosine, and must therefore be judged to be a slightly impure form of CN-6. No sequence studies of the fragment have been reported.

Comparison of the sequence of CN-6 with the reported N-terminal sequences of other serum pro-

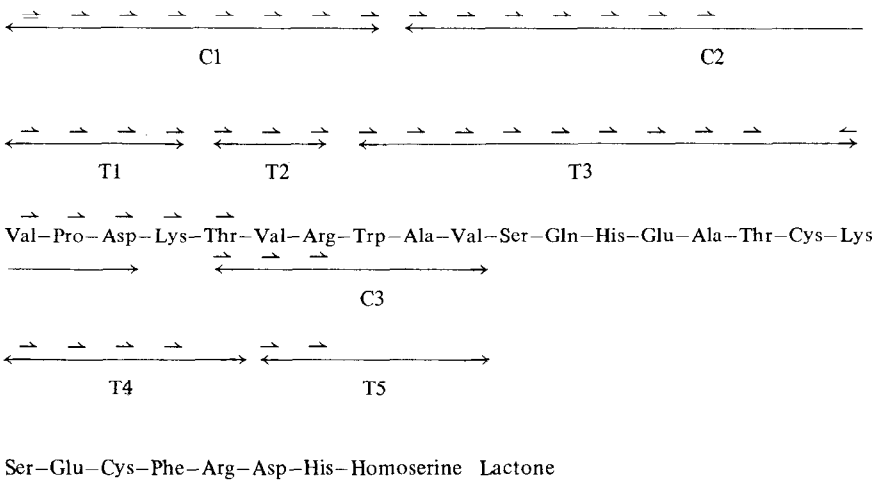

Fig. 1. The amino acid sequence of the amino-terminal fragment derived from sequence analysis of overlapping tryptic (T) and chymotryptic (C) peptides; $\rightharpoonup$ indicates a sequence determined by the dansyl–Edman procedure; $\leftharpoondown$ indicates a sequence derived from a time-course carboxypeptidase digest.

teins shows a slight similarity with the sequences of haptoglobin 1 α chain [14] and α_1-glycoprotein [15] which are both homologous with the immunoglobulin group. Clearly more extensive sequence studies are required to evaluate the possibility of distant homology between transferrin and this group. Such studies are currently in progress.

Acknowledgements

We are grateful to Mrs. Margaret Burnley for carrying out amino acid analyses and to the Science Research Council for a grant in support of this work, including a research studentship for M.R.S.

References

[1] Mann, K.G., Fish, W.W., Cox, A.C. and Tanford, C. (1970) Biochemistry 9, 1348–1354.
[2] Greene, F.C. and Feeney, R.E. (1968) Biochemistry 7, 1366–1371.
[3] Aisen, P., Leibman, A. and Reids, H.A. (1966) J. Biol. Chem. 241, 1666–1671.
[4] Ulmer, D.D. and Vallee, B.L. (1963) Biochemistry 2, 1335–1340.
[5] Woodworth, R.C., Morallee, K.G. and Williams, R.J.P. (1972) Biochemistry 9, 839–842.
[6] Luk, C.K. (1971) Biochemistry 10, 2838–2843.
[7] Feeney, R.E. and Allison, R.G. (1969) Evolutionary Biochemistry of Proteins (Wiley-Interscience N.Y) pp. 151–152.
[8] Sutton, M.R. and Brew, K. (1973) in preparation.
[9] Eriksson, S. and Sjoquist, J. (1960) Biochim. Biophys. Acta, 45 290–296.
[10] Findlay, J.B.C. and Brew, K. (1972) Eur. J. Biochem. 27, 65–86.
[11] Brew, K. (1972) Eur. J. Biochem. 27, 341–353.
[12] Offord, R.E. (1966) Nature 211, 591–593.
[13] Bezkorovainy, A. and Grohlich, D. (1973) Biochim. Biophys. Acta, 310. 365–375.
[14] Black, J.A. and Dixon, G.H. (1968) Nature 218, 736–741.
[15] Schmid, K., Kaufman, H., Isemura, S., Bauer, F., Emura, J., Motoyama, T., Ishiguro, M. and Nanno, S. (1973) Biochemistry 12, 2711–2724.