Purification and Characterization of the Seven Cyanogen Bromide Fragments of Human Serum Transferrin

By MICHAEL R. SUTTON and K. BREW* Department of Biochemistry, University of Leeds, 9 Hyde Terrace, Leeds LS2 9LS, U.K.

(Received 1 October 1973)

1. Procedures are described for the isolation of seven distinct cyanogen bromide fragments in high yield from human serum transferrin. 2. Cyanogen bromide-cleaved transferrin is separated into three fragments (CN-A, CN-B and CN-C) by gel filtration with Sephadex G-100. 3. Four peptides are obtained from CN-A (the largest fragment) after reduction and carboxamidomethylation, by gel filtration in acidic solvents. Two peptides are similarly obtained from fragment CN-B, whereas fragment CN-C is a single cystine-free peptide. 4. The molecular weights of the seven peptides, as determined by polyacrylamide-gel electrophoresis in the presence of sodium dodecyl sulphate, by sedimentation-equilibrium ultracentrifugation and by sequence studies, range from 3100 to 27000. Together they account for a molecular weight of 76200 for transferrin. 5. The two largest fragments contain the carbohydrate attachment sites of the protein, and the smallest fragment is derived from the *N*-terminus. 6. The amino acid compositions and *N*-terminal groups of the fragments are reported and the results compared with those of previous investigations.

Human serum transferrin consists of a single polypeptide chain of molecular weight about 77000 (Greene & Feeney, 1968; Mann et al., 1970) to which are attached two identical oligosaccharide moieties by N-glycosidic linkages with asparagine side chains (Jamieson, 1965). The molecule possesses two similar binding sites for Fe³⁺ and other metal ions (Schade et al., 1949; Ulmer & Vallee, 1963; Woodworth et al., 1972). Various workers have speculated that the transferrin structure may have originated during phylogeny as a result of the doubling in size of the structural gene for a smaller polypeptide precursor which possessed a single metalbinding site (Feeney & Allison, 1969; Palmour & Sutton, 1971). To evaluate this hypothesis and to produce a basis for the study of the structures of the binding sites of transferrin, we have initiated an investigation of the amino acid sequence of the human protein. The polypeptide chain of a transferrin is large for direct sequence studies, and therefore as a preliminary step we have isolated and characterized a complete set of fragments produced by cleavage of the protein with CNBr.

Bezkorovainy & Grohlich (1973) have reported the results of a similar study in which six peptides were isolated after CNBr cleavage of transferrin, two of these being prepared in small amounts by elution

Vol. 139

from polyacrylamide gels. A comparison with our study suggests that at least one peptide was lost on separation and that incomplete resolution of the other peptides was obtained with their procedures. Our preparative procedures permit the isolation of seven distinct peptides of molecular weights 3100 to 27000 in high yield, on a sufficient scale for sequence analysis. The combined molecular weights of the fragments account for the entire polypeptide chain of transferrin.

Materials and Methods

Materials

Iron-free human transferrin was purchased from Sigma Chemical Co., St. Louis, Mo., U.S.A., and, being found to be substantially pure (>95%) by polyacrylamide-gel electrophoresis, was used without further purification. CNBr was obtained from Koch-Light Laboratories Ltd., Colnbrook, Bucks., U.K., and 2-mercaptoethanol, iodoacetamide, dansyl chloride and other reagents were from BDH Chemicals Ltd., Poole, Dorset, U.K. Iodoacetamide was freed of iodine by repeated solution in water and drying *in vacuo* over NaOH pellets. Sephadex G-100 (fine grade), G-50 (superfine grade) and G-25 (fine grade) were purchased from Pharmacia Fine Chemicals, Uppsala, Sweden, and CM-cellulose (CM-52) was from Whatman Biochemicals Ltd.,

^{*} Present address: Department of Biochemistry, University of Miami School of Medicine, P.O. Box 875, Biscayne Annex, Miami, Fla. 33152, U.S.A.

Maidstone, Kent, U.K. Reduced and carboxymethylated proteins as standards for polyacrylamidegel electrophoresis were gifts from Dr. G. S. Bell, Department of Biophysics, University of Leeds.

Methods

Cleavage with CNBr. Apotransferrin was treated with CNBr in 70% (v/v) formic acid as described previously (Findlay & Brew, 1972). After drying *in vacuo* over NaOH pellets, the digest was dissolved in 5% (v/v) formic acid at a concentration of 20 mg/ml for column chromatography.

Reduction and S-carboxamidomethylation. Polypeptide fragments were dissolved at a concentration of 10–15 mg/ml in 8 m-urea (deionized) containing 0.2 m-Tris-HCl, pH8.6, and reduction was initiated by the addition of 1% (v/v) mercaptoethanol. After 4 h at room temperature, 30 mg of iodoacetamide/ml was added and the mixture left for 15 min. Reaction was terminated by adding excess of mercaptoethanol and leaving for 30 min, and the modified peptides were separated from reagents by gel filtration with columns of Sephadex G-25 equilibrated with 5% formic acid, or were applied directly to columns for separation.

Polyacrylamide-gel electrophoresis in the presence of sodium dodecyl sulphate. The molecular weights of the larger CNBr peptides were determined by comparison of their mobilities on electrophoresis in polyacrylamide gels in the presence of sodium dodecyl sulphate as described by Weber & Osborn (1969). The following reduced carboxymethylated proteins were used as standards in the determination: bovine glutamate dehydrogenase (53000), heavy and light chains of pooled bovine immunoglobulin G (50000 and 25000 respectively), ovalbumin (43000), pepsin (35000), chymotrypsinogen (25700), myoglobin (17200) and cytochrome c (11700). Two reduced aminoethylated CNBr fragments of known structure were used as additional standards in the lower-molecular-weight range, guinea-pig α -lactalbumin CNBr 1 (10000: Brew, 1972) and human α-lactalbumin CNBr D (7200; Findlay & Brew, 1972).

Polyacrylamide gels (10%) were prepared with 0.6 and 1.2% bisacrylamide and after electrophoresis were stained for protein with 1% Coomassie Brilliant Blue in 12.5% (w/v) trichloroacetic acid and subsequently destained with 5% (v/v) acetic acid or were stained for carbohydrate with fuchsin reagent after periodate treatment (Zacharius *et al.*, 1969).

Calibration curves [log (molecular weight) versus mobility of protein relative to Bromophenol Blue] were linear in the higher-molecular-weight range (>15000) but showed curvature below this, even with 1.2% bisacrylamide. Values given for the molecular weights are the average of five to ten determinations.

Sedimentation-equilibrium ultracentrifugation. The molecular weights of peptides CN-1, CN-3 and CN-4 were determined by low-speed sedimentation equilibrium with a Beckman model E analytical ultracentrifuge. Freeze-dried peptides were dissolved at a concentration of 0.1-0.2% (w/v) in 6m-urea (deionized) containing 0.1M-sodium cacodylate buffer, pH7.4, and were centrifuged at 16000 (CN-1), 24000 (CN-3) or 26000 (CN-4) rev./min at 20°C until equilibrium was attained. The rotor was then accelerated to 60000 rev./min for a short period to deplete the meniscus, thereby permitting determination of the absolute value of the fringe number (i_m) at the lower speed. Molecular weights were then calculated as described by La Bar (1965) by using values for the partial specific volume (v) obtained from the amino acid composition (Cohn & Edsall, 1943). It is realized that as carbohydrate is present in fragment CN-1, the value for \bar{v} calculated in this way will be a slight overestimate, thereby leading to an overestimate of the molecular weight of this fragment. The close agreement between the values for the molecular weight obtained in this way and by gel electrophoresis nevertheless leads us to consider that the error is not very large.

Amino acid analyses. Samples of protein or CNBr peptides were hydrolysed with 6M-HCl at 110°C in sealed evacuated Pyrex tubes. Peptides were hydrolysed only for 24h and no correction was made for destruction of threonine and serine or for the incomplete release of other amino acids. The amino acid composition of whole apotransferrin was determined by the analysis of samples hydrolysed for 24, 48 and 96h, serine and threonine values being extrapolated to zero time and assuming complete release of isoleucine and valine at 96h. All analyses were performed with a BioCal BC200 automatic amino acid analyser by the single-column procedure. Tryptophan was not determined directly by amino acid analysis. The tryptophan contents given are from sequence studies of the CNBr fragments (M. R. Sutton & K. Brew, unpublished work).

N-Terminal determination. N-terminal residues of purified fragments were determined by the dansyl chloride procedure as described previously (Findlay & Brew, 1972). With the larger CNBr peptides (CN-1, CN-2 and CN-3) dansylation was carried out in the presence of sodium dodecyl sulphate.

Results

Separation of CNBr peptides

CNBr digests of apotransferrin (400 mg) dissolved in 18 ml of 5% formic acid were fractionated by gel filtration on columns ($4 \text{ cm} \times 130 \text{ cm}$) of Sephadex G-100 equilibrated with 5% formic acid. The columns were eluted at a flow rate of 25 ml/h and 7.5 ml



Fig. 1. Separation of CNBr fragments of human transferrin on Sephadex G-100





Fig. 2. Separation of reduced and carboxamiaomethylated CN-A peptides on Sephadex G-100

Experimental details are given in the text.

fractions were collected and monitored by the E_{280} (see Fig. 1). Three fragments were obtained (CN-A, CN-B and CN-C, in order of elution) together with a shoulder on the leading edge of the first peak. This proved to contain only small amounts of aggregated material and was not further treated. Fragment CN-C alone showed a single N-terminal group (glycine). Fragment CN-A (approx. 250 mg of material derived from 400mg of transferrin) was reduced, carboxamidomethylated and re-fractionated by gel filtration with Sephadex G-100 under the above conditions. Essentially similar results were obtained when the material was desalted, freeze-dried and redissolved in 5% formic acid (15ml) or when the reaction mixture was applied directly to the column. Four fragments were separated (Fig. 2), the first two being well resolved (CN-1 and CN-2) and the second two (CN-3 and CN-4) incompletely separated. Fragments CN-3 and CN-4 were pooled,



Fig. 3. Purification of fragments CN-3 and CN-4 by recycling on a column of Sephadex G-50

Experimental details are given in the text.



Fig. 4. Separation of reduced and carboxamidomethylated peptides on Sephadex G-50

Experimental details are given in the text.

freeze-dried, redissolved in 6ml of 3% formic acid and separated by recycling with a column of Sephadex G-50 equilibrated with 3% formic acid. The mixture was applied to a column ($2.8 \text{ cm} \times 85 \text{ cm}$) of Sephadex G-50 (superfine grade), fitted with upward-flow adaptors and a peristaltic pump. The column was eluted at a flow rate of 20ml/h and the effluent passed through a u.v. monitoring device (LKB Uvicord II) before recycling to the column via the peristaltic pump. After the second partially resolved peak had been completely eluted and reloaded, the column was eluted with fresh 3% formic acid and 6.5ml fractions were collected. Fragments CN-3 and CN-4 were pooled as shown (Fig. 3). Fragment CN-B was reduced, carboxamidomethylated and applied directly to a column $(4 \text{ cm} \times 75 \text{ cm})$ of Sephadex G-50 equilibrated with 5% formic acid. The column was eluted at a flow rate of 25 ml/h and two well separated peptides (CN-5 and CN-6) were obtained (Fig. 4).

Fraction CN-C was devoid of cystine on amino acid analysis and contained a single N-terminal group. A sample of fragment CN-C was dissolved in 25 mmammonium acetate, pH 5.0, and applied to a column (0.9 cm × 10 cm) of CM-cellulose equilibrated at 45°C with the same buffer. The column was eluted with a linear gradient composed of 250ml each of 25mmammonium acetate and 200mm-ammonium acetate, pH5.0. Two components were thereby separated which had essentially identical amino acid compositions and the same N-terminal residue. We attribute the apparent heterogeneity of the fragment to partial lactonization of the C-terminal homoserine (derived from methionine on CNBr cleavage) or to partial deamidation, and assume that it is a single fragment (CN-7). Sequence analysis (M. R. Sutton & K. Brew, unpublished work) has confirmed this view.

Characteristics of the CNBr peptides

All of the CNBr peptides gave single N-terminal groups on N-terminal analysis by the dansyl chloride procedure. The larger fragments (CN1-5) gave single bands on gel electrophoresis, the bands for fragments CN-1 and CN-2 alone showing positive results with the fuchsin-periodate stain. Fragments CN-6 and CN-7 were not examined by gel electrophoresis. The molecular weights of fragments CN-1, CN-2, CN-3, CN-4 and CN-5 were determined by gel electrophoresis. Molecular-weight values for fragments CN-1 and CN-4 determined by sedimentation equilibrium were closely similar to those obtained by gel electrophoresis. However, the molecular weights for fragments CN-3 and CN-5 determined from their electrophoretic mobilities proved to be inconsistent with values obtained by other procedures. Fragment CN-3 gave a value of 12000 by gel electrophoresis and 8900 by sedimentation equilibrium; fragment CN-5 had an electrophoretic mobility corresponding to a molecular weight of 6000, whereas amino acid sequence studies have shown the molecular weight to be 9300 (M. R. Sutton & K. Brew, unpublished work).

Sequence studies on fragments CN-6 and CN-7 (M. R. Sutton & K. Brew, unpublished work) have confirmed the molecular weights indicated by amino acid analysis for these two peptides. The characteristics of the fragments, including their amino acid compositions, are summarized in Table 1.

The distinct nature of the seven fragments is apparent not only from the differences in their sizes and N-terminal groups, but also from the amino acid compositions. This is particularly obvious with fragments CN-5, CN-6 and CN-7, where certain amino acids are completely absent from the acid hydrolysates, and becomes apparent with fragments CN1-4 if the molar ratios of various residues are compared.

The molar yields of the fragments purified as described, based on our estimates of their molecular weights, range from 68 to 109%, indicating that only one of each is obtained on CNBr cleavage of the transferrin molecule. Their combined molecular weight (76200) is in good agreement with that determined for transferrin by physicochemical methods, and the summed amino acid composition is in reasonable agreement with the composition of the intact protein.

Discussion

Two previous studies have been made of the CNBr fragments of human serum transferrin by Jeppsson (1967) and by Bezkorovainy & Grohlich (1973). Jeppsson (1967) achieved a separation of an initial CNBr digest similar to that shown in Fig. 1, but no further fractionation was attempted and the molecular weights given for fragments corresponding to CN-A CN-B and CN-C are clearly erroneous.

In the more recent study by Bezkorovainy & Grohlich (1973), CNBr-treated transferrin was 3-carboxypropionylated and separated by gel filtration with Sephadex G-200 in NaHCO₃ buffer into two pools of material. We consider that one of these clearly corresponds to fragment CN-A, and the second to a mixture of fragments CN-B and CN-C, which was subsequently resolved by gel filtration with Sephadex G-100. De-acylated fragment CN-A was reduced and alkylated but resolved into only two fractions by gel filtration with Sephadex G-100 in a buffer of neutral pH. One fraction so obtained contained carbohydrate, resembled fragment CN2 in amino acid composition and molecular weight but is reported to have N-terminal serine, with contaminating aspartic acid. From the other fraction, two components were obtained by elution from polyacrylamide gels after electrophoresis. The larger of these (mol.wt, 26000) resembles fragment CN-1 in size, N-terminal residue (glycine), carbohydrate content and amino acid composition. The second fragment has a molecular weight of 15000, N-terminal aspartic acid, contained only small amounts of carbohydrate, but resembles fragment CN-2 rather than fragment CN-3 or CN-4 in amino acid composition.

Fragments similar to CN-5, CN-6 and CN-7 were isolated, whose amino acid compositions resemble those reported here, but show signs of crosscontamination (the presence of glycine and leucine in fragment CN-6, histidine in CN-5 and tyrosine and threonine in CN-7).

Compositions are given as mol of residues/mol of peptide and for the whole protein are rounded off to the nearest integer. Yields are given in mol %, based on the weights of the freeze-dried fragments and assuming the molecular weights given in the Table. N.D. = Not determined. For whole transferrin the best values are averages except for threonine (24h value), serine (extrapolated to 0h), valine and isoleucine (72h values).

				Eroment						Whole tr	ansferrin	
								` ı				Best
-	CN-1	CN-2	CN-3	CN4	CN-S	CN-6	CN-7	Total	24h	48h	72h	value
Asp	26.1 (26)	13.2 (13)	8.4 (8)	5.0 (5)	10.7 (11)	2.0 (2)	5.0 (5)	20	72.0	70.3	70.8	71.0
Thr	10.7 (11)*	4.5 (5)*	3.7 (4)*	4.1 (4)*	1.8 (2)*	1.8 (2)*	0	28	25.3	23.7	24.0	25.3
Ser	9.6 (10)*	6.1 (6)*	3.8 (4)*	3.0 (3)*	4.9 (5)*	1.7 (2)*	3.8 (4)*	34	32.8	29.9	28.2	35.1
Glu	18.8 (19)	9.1 (9)	6.0 (6)	7.0 (7)	6.3 (6)	3.1 (3)	7.0 (7)	57	52.9	52.6	52.8	52.7
Pro	10.9 (11)	6.6 (7)	3.6 (4)	3.4 (3)	5.9 (6)	1.1 (1)	2.7 (3)	35	35.2	35.1	38.2	36.2
Gly	16.7 (17)	9.9 (10)	6.1 (6)	3.4 (3)	4.9 (5)	0	5.0 (5)	8	46.8	45.5	46.6	46.3
Ala	14.8 (15)	9.6 (10)	6.6 (7)	4.4 (4)	13.5 (13)	2.0 (2)	2.1 (2)	53	50.9	50.8	52.2	51.3
<u> </u>	7.0 (7)†	5.9 (6)	3.9 (4)†	4.1 (4)†	2.1 (2)	2.0 (2)†	0	25	29.5	31.1	20.4	30.3
Val	6.5 (7)	3.7 (4)	4.8 (5)	4.1 (4)	10.1 (10)	2.5 (3)	1.2 (1)	34	36.8	40.6	40.3	40.3
Met	0	0	0	0	0	0	0		37.0	7.6	7.6	7.4
Ile	2.3 (2)	2.3 (2)	2.1 (2)	2.3 (2)	3.2 (3)	1.0 (1)	1.0 (1)	12	13.7	13.5	13.6	13.6
Leu	23.3 (23)	11.8 (12)	4.5 (5)	3.5 (4)	5.6 (6)	0	6.5 (7)	57	52.6	51.6	51.9	52.0
Tyr	12.5 (13)	4.7 (5)	5.0 (5)	2.9 (3)	5.8 (6)	0	0	32	24.6	24.0	24.5	24.4
Phe	12.2 (12)	5.4 (5)	1.7 (2)	0.97 (1)	3.0 (3)	0.95 (1)	4.5 (5)	5	27.0	26.0	26.7	26.6
Trp	N.D.	‡(I) [I]	N.D.	N.D.	(0) 0	[1] (1)	[1] (1)					N.D.
His	8.3 (8)	2.5 (3)	1.8 (2)	2.3 (2)	0	1.6 (2)	2.8 (3)	ຊ	17.0	17.1	17.0	17.0
Lys	15.9 (16)	9.8 (10)	8.0 (8)	5.0 (5)	7.4 (7)	1.8 (2)	7.0 (7)	55	48.9	49.2	49.2	49.1
Arg	10.7 (11)	4.3 (4)	1.7 (2)	2.3 (2)	1.1 (1)	1.9 (2)	1.0 (1)	33	22.1	22.3	23.1	22.5
Carbohydrate	+	+	, ,			,	,					
Mol.wt.	27000§	15000§	10068	72008	9290‡	3040	5840‡	76200				76600
Molar yield (%)	87	68	109	93	.89	95	78					
N-Terminus	Gly	Asp	Ser	Tyr	Lys	Val	Gly					Val
: : : : : : : : : : : : : :	•											

* From 24h hydrolysis only.

† Determined as S-carboxymethylcysteine. From sequence studies.

S

Molecular weights from gel electrophoresis.

Molecular weights from sedimentation equilibrium.

It is difficult to reconcile completely the results of this study with ours, especially as the separation patterns of peptides obtained from fragment CN-A on reduction and alkylation are not reported. We have found that fragments CN-3 and CN-4 have pronounced tendencies to aggregate, and it appears possible that in the absence of dissociating solvents they will be eluted from Sephadex G-100 with fragment CN-1, but because of their small size may have migrated off the polyacrylamide gels during the electrophoretic separation of this fraction.

It is, however, clear from their distinct amino acid compositions, *N*-terminal residues and homogeneity on gel electrophoresis that peptides CN-1 to CN-7 are homogeneous and distinct. The procedures described provide a single and relatively rapid method for obtaining these fragments from human transferrin in good yield and therefore provide a good basis for the elucidation of the amino acid sequence of the protein.

Further studies will be required to determine the ordering of the fragments in the transferrin molecule, but as CN-6 is the only fragment with the same *N*-terminal residue as transferrin (valine: Eriksson & Sjoquist, 1960; Mann *et al.*, 1970) it is clearly derived from the *N*-terminal region of the protein. The sequence of this fragment (now completed) is in close agreement with the amino acid composition, as are the largely completed sequences of fragments CN-5 and CN-7 (M. R. Sutton & K. Brew, unpublished work).

There is considerable variation in the reported values for the methionine content of transferrin, ranging from 4 to 10 residues per molecule (Mann et al., 1970; Bezkorovainy et al., 1968). Our amino acid analyses indicate the presence of seven residues of methionine and, if they are correct, eight fragments would be expected on cleaving the molecule with CNBr. The isolation of one fewer fragment than expected, if not attributable to inaccuracies in the methionine determination, can be accounted for if the protein contains a Met-Met sequence or a methionyl bond resistant to CNBr cleavage. In the former case, free homoserine or homoserine lactone would be expected to be present in the digest, and may have been missed in our separation procedures. Neither possibility can therefore be eliminated at present, although an examination of the regions of sequence around the methionyl residues, which will be necessary to determine the alignment of the CNBr fragments, should distinguish between them. As the seven fragments were separated by gel filtration, and their summed molecular weight closely approximates

to that of the entire polypeptide chain of transferrin, it appears unlikely that an eighth fragment can have been lost during purification. The isolation of the seven fragments in high yield and approximately equimolar amounts is entirely consistent with the now generally accepted view that transferrin consists of a single polypeptide chain. In a molecule containing two identical polypeptide chains, fewer CNBr fragments would be expected in higher molar yields (>100%) and a situation where two dissimilar chains are present containing a total of seven methionine residues should give rise to nine fragments on CNBr cleavage, which is again less consistent with our findings.

We thank Mrs. Margaret Burnley for carrying out the amino acid analyses and Dr. Janet Powell for operating the analytical ultracentrifuge. We also thank the Science Research Council for a grant in support of this work, and a research studentship for M. R. S.

References

- Bezkorovainy, A. & Grohlich, D. (1973) Biochim. Biophys. Acta 310, 365-375
- Bezkorovainy, A., Grohlich, D. & Gerbeck, C. M. (1968) Biochem. J. 110, 765-770
- Brew, K. (1972) Eur. J. Biochem. 27, 341-353
- Cohn, E. J. & Edsall, J. T. (1943) Proteins, Amino-Acids and Peptides, p. 370, Reinhold, New York
- Eriksson, S. & Sjoquist, J. (1960) Biochim. Biophys. Acta 45, 290–296
- Feeney, R. E. & Allison, R. G. (1969) Evolutionary Biochemistry of Proteins, pp. 151–152, Wiley-Interscience, New York
- Findlay, J. B. C. & Brew, K. (1972) Eur. J. Biochem. 27, 65-86
- Greene, F. C. & Feeney, R. E. (1968) Biochemistry 7, 1366-1371
- Jamieson, G. A. (1965) J. Biol. Chem. 240, 2914-2920
- Jeppsson, J. (1967) Biochim. Biophys. Acta 140, 447-486
- La Bar, F. E. (1965) Proc. Nat. Acad. Sci. U.S. 54, 31-36
- Mann, K. G., Fish, W. W., Cox, A. C. & Tanford, C. (1970) *Biochemistry* 9, 1348–1354
- Palmour, R. M. & Sutton, H. E. (1971) Biochemistry 10, 4026-4032
- Schade, A. L., Reinhart, R. W. & Levy, H. (1949) Arch. Biochem. 20, 170-172
- Ulmer, D. D. & Vallee, B. L. (1963) Biochemistry 2, 1335-1340
- Weber, K. & Osborn, M. (1969) J. Biol. Chem. 244, 4406-4412
- Woodworth, R. C., Morallee, K. G. & Williams, R. J. P. (1972) *Biochemistry* 9, 839–842
- Zacharius, R. M., Zell, T. E., Morrison, J. H. & Woodlock, J. J. (1969) Anal. Biochem. 30, 148-152