STRUCTURAL STUDIES ON HUMAN TRANSFERRIN

AKADEMISK AVHANDLING

SOM MED VEDERBÖRLIGT TILLSTÅND AV KANSLERN FÖR RIKETS UNIVERSITET FÖR VINNANDE AV MEDICINE DOKTORSGRAD FRAMLÄGGS TILL OFFENTLIG GRANSKNING I FÖRELÅSNINGSSALEN Å INSTITUTIONEN FÖR MIKROBIOLOGI UMEÅ UNIVERSITET ONSDAGEN DEN 17 MAJ 1967 KL. 9 P.M.

AV

JAN-OLOF JEPPSSON

MED.KAND.
STRUCTURAL STUDIES ON HUMAN TRANSFERRIN

By

Jan-Olof Jeppsson

Department of Microbiology, University of Umeå, Umeå, Sweden
This review is a dissertation and contains a summary of the following publications:


In addition the dissertation contains some hitherto unpublished results. In the text the above mentioned papers will be referred to by the Roman figures I — V, other references are indicated by Arabic figures.

The following abbreviations have been used: Tf, transferrin; CSF, cerebrospinal fluid; CM, carboxymethyl; DEAE, diethylaminoethyl; TEAE, triethylaminoethyl; Tris, Tris(hydroxymethyl)aminoethane; β-ME, β-mercaptoethanol; CNBr, cyanogen bromide; PTC, phenylthiocarbamyl; PTH, phenylthiohydantoin.
INTRODUCTION

Twenty years ago it was independently demonstrated in two laboratories that the major part of the acid-soluble iron in plasma is bound to a specific metal-combining protein usually called transferrin, less commonly siderophilin or β1 metal-combining globulin. Holmberg and Laurell (1) showed by in vivo experiments that neither the naturally occurring serum iron nor iron added to a certain limit reacted with α, α1-dipyridyl in the presence of a reducing agent. However, all iron added in excess of 315 μg/100 ml reacted with the complexing agent. Schade and Caroline (2) observed that when small amounts of ferrous iron were added the serum color changed from yellow to yellow-red up to a definite limit, beyond which further addition of iron produced no color change.

Transferrin plays an important role in the metabolism of iron in the body. Only 3—4 mg of iron or 0.1 % of the total body iron is present as serum iron, which is almost entirely transferrin-bound iron. Hemoglobin synthesis daily requires six to ten times the amount of iron available in the serum. One function of transferrin is to transport iron from sites of absorption, liberation or storage. Under normal conditions transferrin is saturated to only 30 % of its iron binding capacity, thus its reserve capacity for iron binding provides a mechanism of protection against iron intoxication.

The molecular weight of human transferrin has until recently been assumed to be about 90,000 (ref. 3, 4). The molecule has been thought to consist of a single polypeptide chain with one N-terminal amino acid, valine (5).

Two atoms of ferric iron are bound by ionic binding to each transferrin molecule, when fully saturated. The iron complex is pH dependent and is stable in the pH range 7.0—10.0.

A transferrinemia has been reported in three cases. In two of them (6, 7) the deficiency was probably acquired, whereas the third (8) appears to have been a case of hereditary absence of transferrin.
Since Smithies ten years ago first demonstrated the existence of genetic variants of transferrin by their different mobilities in starch gel electrophoresis, seventeen such variants have been reported (9, 10). The most commonly occurring transferrin has been designated C. Variants with greater mobilities in starch gel electrophoresis are called B and those with lower mobilities D, with subscripts indicating their relative positions in the gel. Because of the low frequency of transferrin variants, the precise mode of inheritance has not yet been established. However, some data suggest an autosomal multiple-allele system without dominance. None of the transferrin variants has been associated with a clinical abnormality. No significant differences have been demonstrated among four transferrins so far studied with respect to the amount and rate of iron binding and transport (11). Smithies and Connell have suggested that the transferrins differ with respect to charge rather than molecular size and have pointed out that several amino acids would have to be interchanged to account for all the known forms (12). However, to date only a few preliminary attempts have been made to discover the chemical differences among the transferrin variants.

In the present work human transferrin CC was separated into two fractions by chromatography on cation-exchange cellulose (I). The two fractions showed similarities in their immunological, biological and physical properties. Their carbohydrate compositions were somewhat different. Evidence is presented indicating that the transferrin molecule consists of two subunits, each with a molecular weight half that of native transferrin (III). Genetic variants of transferrin were separated by various methods. Chemical differences between two variants were indicated by their peptide patterns on fingerprints (II). Some characteristic peptides were isolated from a minor fragment, obtained by cyanogen bromide degradation (V). Amino acid analyses and amino acid sequence determinations on some of these peptides showed at least one amino acid substitution, glycine replacing aspartic acid, which distinguishes transferrin D₁ from C.

PREVIOUS INVESTIGATIONS

Purification of transferrin.

Schade and Caroline showed that a protein fraction in Cohn's fraction IV — 3, 4, chelated iron, forming a salmon-pink colored complex (2). Transferrin was first isolated from swine plasma (14) and later from human plasma (15). Human transferrin was first crystallized by Koechlin from Cohn's fraction IV-7,75 % of which consists of transferrin (4). Inman et al (16) also described a large-scale purification of transferrin from Cohn fraction IV-4 giving
yields of 30—50 % with a purity of 90—95 %. When the rivanol method is used, all serum proteins are precipitated except transferrin and γ-globulin (17, 18). The γ-globulin fraction can easily be removed by precipitation with ammonium sulfate, alcohol precipitation or ion-exchange chromatography. Parker and Bearn have described a convenient small-scale method for preparing transferrin of high purity (19). The β-globulin fraction was isolated by preparative electrophoresis, concentrated and chromatographed on TEAE-cellulose. The yield was reported to be about 50 mg from 40 ml serum, corresponding to 50 % recovery. The chromatographic step also permitted a partial separation of genetic variants of transferrin. Recently several authors have published methods including combinations of gel chromatography and ion-exchange chromatography (20—23). The yields and purities of the preparations obtained were highly variable.

Physical and chemical properties of transferrin.

Each transferrin molecule can combine with two atoms of ferric iron in an ionic bonding in which one bicarbonate is taken up per iron atom (13). Transferrin is colorless in the absence of iron and pink in its presence. The Fe-transferrin complex has an absorption maximum at 470 mμ and a minimum at 410 mμ. On a weight basis, the binding of iron by transferrin has been variously reported (24—26) as 1.26, 1.30 or 1.35 μg/mg. The last value is in better agreement with the earlier assumed molecular weight of about 90,000. The detailed mechanism of the binding of iron to transferrin is not yet understood, but some pertinent data exist. The iron in the transferrin complex is trivalent and is bound with essentially ionic bonds. For each Fe³⁺ bound 3 protons are released and 1 HCO₃⁻ is bound (27). It was also suggested that the binding of Fe³⁺ to transferrin may be described as coordination to two equivalent and independent sites. The electron-spin resonance spectra of the Fe-transferrin complex indicate that the metal is present as Fe³⁺. Aasa et al (27) also suggest that each chelating site involves not only three tyrosyl residues but also two imidazole groups of the protein. Experiments by Jones et al (28) support the hypothesis that tyrosyl residues are part of the iron binding sites. Heavy iodination of native transferrin destroyed the Fe-binding capacity. It was also shown that Ca²⁺, Cu²⁺, Zn²⁺ and Co²⁺ bind to transferrin at non-specific sites. Electrophoretic studies by Aisen et al on the binding of iron to transferrin indicated that the stability constants of the two binding sites are approximately equal (23). It also appeared that the binding sites are equivalent and independent, which is in good agreement with the results of Aasa et al (27).

The molecular weight of human transferrin has until recently been assumed to be about 90,000. This value was determined by the combination of sedimentation and diffusion (3), light scattering (29), osmotic pressure (30)
and iron binding studies (4). However, there is now wide variation in the molecular weights reported for human transferrin. From iron binding studies Jandl and Katz (25) have estimated a molecular weight of 86,000. From sedimentation and diffusion constants Hazen Jr. (31) reported a molecular weight of 82,500. Extreme values of 68,000 (ref. 22) and 92,000—95,000 (ref. 32) have been reported. The latter author has recently published a value of 89,000 (ref. 33). From determinations of sedimentation equilibrium, sedimentation and diffusion, osmotic pressure and maximum iron binding capacity, Roberts et al (34) have recently reported molecular weights between 73,200—76,000 for four different preparations.

Amino acid analyses of human transferrin have been performed by Parker and Bearn (19). The number of amino acid residues based on a molecular weight of 83,000 were in good agreement with the values of Hazen Jr. (31). Heimburger et al (35) have performed amino acid analyses on most of the plasma proteins, including transferrin. Their values are based on a molecular weight of 90,000 and are about 10 % higher. The results for methionine are identical, 9 residues/mole protein. 44 half cystine residues/mole were obtained (35) and no free sulfhydryl groups have been detected (3). It can therefore be assumed, that transferrin contains 22 disulfide bridges which may be interchain or intrachain. Parker et al (5) reduced and alkylated the disulfide bridges in transferrin by the method of Crestfield et al (36). They failed to demonstrate dissociation of transferrin into subunits as determined by equilibrium ultracentrifugations, gel filtration on Sephadex G-75 and urea-starch gel electrophoresis in formate (pH 2.8), phosphate-citrate (pH 5.6) and borate (pH 9.0) buffers. In addition, N-terminal amino acid determinations on transferrin by the Edman method (37), the DNP-method of Sanger (38) and the cyanate method of Stark and Smyth (39) revealed a single valine residue. By these criteria, the transferrin molecule has been regarded as consisting of a single polypeptide chain of 750—800 amino acids; this is one of the largest proteins in which subunits have not been recognized.

The glycoprotein nature of transferrin has been investigated by Schultze et al (40), who observed a total carbohydrate content of 5.5 %. This was distributed among 12 hexose residues/mole protein, 8 acetylhexosamine residues, 4 sialic acid residues and approximately one residue of fucose. Gottschalk (41) has suggested that sialic acid always is found in a terminal position in heterosaccharides of animal glycoproteins. The glycopeptide structure of human transferrin has recently been investigated by Jamieson (42). His analytical data showed that transferrin contains 4 moles of sialic acid, 8 moles of N-acetylglucosamine, 4 moles of galactose and 8 moles of mannose per 90,000 g glycoprotein. These carbohydrate residues are arranged as two identical branched chains. The terminal residue is sialic acid. Each chain is probably linked to the protein part by means of asparaginyl-sugar linkages.
Genetic variants of transferrin.

The gene locus for transferrin appears to be capable of numerous viable mutations. Each of the seventeen different molecular species of transferrin which has been described has been identified by starch gel electrophoresis. The variants migrating faster than the common C band are called B types and those migrating more slowly are called D types. Subscripts are used to distinguish different types from each other in the fast and slow moving groups. Almost all the variants have been found in heterozygous individuals in association with transferrin C. In addition to individuals homozygous for transferrin C, homozygotes of two types of B and D respectively have been described (5). Most variants occur sporadically in the populations that have been investigated but there may be some racial correlation. Transferrin B\textsubscript{2}C occurs in about 1 \% of Canadian whites and Europeans (43), CD\textsubscript{1} in about 12 \% of negroes (5). CD\textsubscript{Chi} is a slow variant observed in about 6 \% of a Chinese population (44). This variant was also observed in the Swedish Lapps (45). Transferrins B\textsubscript{1}C, B\textsubscript{2}C and CD\textsubscript{1} together occurred in a Swedish population to about 1 \%. Transferrin B\textsubscript{2}C was the most common variant (46).

The most complete study on the chemical differences of transferrin variants has been done by Parker and Bearn (19). Isolated transferrins B\textsubscript{2}, C and D\textsubscript{1} had similar sedimentation properties in the ultracentrifuge. Several B and D variants gave immunological reactions of identity with transferrin C in the Ouchterlony test. The purified transferrins B\textsubscript{2}, C, D\textsubscript{1} and D\textsubscript{3} had similar amino acid compositions.

Enzymatic removal of sialic acid by neuraminidase produced stepwise four additional slower moving components, whose relative intensities depended on the concentration and duration of the enzyme activity. Since the same kind of five-band pattern was obtained for each of the variants tested, it appeared that differences in sialic acid content were not involved in the characteristic mobility differences of the untreated transferrin variants (19). The starch gel experiments indicated that almost all of the sialic acid could be removed from human transferrin.

The detection of structural differences among hemoglobin variants achieved by fingerprinting suggested an obvious approach to the transferrin variants. Several laboratories have attacked this problem. Fingerprints of trypsin digested material from Tf C, B\textsubscript{1}, D\textsubscript{1} and D\textsubscript{3} revealed three peptides in D\textsubscript{3} that could not be detected in other types tested, whereas fingerprints of B\textsubscript{1} and D\textsubscript{1} could not be differentiated from Tf C (47). Roop found a characteristic peptide spot on a fingerprint of Tf D\textsubscript{1} that was missing in Tf C (48). However, he did not observe the absence of a peptide spot in Tf D\textsubscript{1} that was present in Tf C. Comparative studies on chymotryptic digests of Tf CC and homozygous Tf D\textsubscript{1}D\textsubscript{1} indicated that an aspartic residue in Tf C was probably
replaced by a glycine residue in Tf D1 (49). Chymotryptic peptides from the whole molecule were isolated directly from fingerprints after faint ninhydrin staining.

**Transferrin in cerebrospinal fluid (CSF)**

Starch gel electrophoresis and immunoelectrophoresis of normal CSF show two distinct proteins which have the electrophoretic and immunological characteristics of transferrin (19). In the starch gel these two bands are of equal intensity. The band with the highest mobility occupies the position of normal serum transferrin and the second migrates close to the position of serum transferrin when all four residues of sialic acid have been removed (19). The slower moving transferrin corresponds to a previously identified ψ or τ-fraction (50) which migrates between the β and γ fractions in paper electrophoresis and which is not detectable in serum. Burtin (51) identified the double arc precipitation line in immunoelectrophoresis as a protein with the antigenic properties of serum transferrin and it was later shown that both arcs of the double bow had the capacity to bind radioactive iron (52). Pette and Stupp (53) showed that after neuraminidase treatment of CSF, the faster migrating part of the double bow disappeared and the remaining transferrin arc of treated CSF coincided with the arc of treated serum. The presence of four transferrins in the CSF of an individual heterozygous for transferrin CD1 has recently been reported (54). Starch gel electrophoresis separated two fast and two slow moving fractions, the latter presumably corresponding to transferrins C and D1 devoid of sialic acid. Morgan and Laurell (55) have presented evidence for the existence of neuraminidase in the brain tissue. However, the presence of neuraminidase cannot explain the electrophoretic pattern of CSF transferrins since the slower fraction does not occupy the position of serum transferrin treated by neuraminidase to remove four sialic acid residues. These observations have been regarded as consistent with the absence of a carbohydrate prosthetic group in the slower moving transferrin in CSF (54).

**PRESENT INVESTIGATIONS**

*Isolation and identification of two normal transferrins (I).*

Ethanol fractionated human transferrin CC has been shown to separate into two fractions when subjected to chromatography on CM-cellulose. The first and second peaks emerging from the column are referred to as TfI and TfII. The CM-cellulose was equilibrated with 0.005 M phosphate buffer, pH 6.0, in which TfI emerged, and TfII was recovered by gradient elution with 0.1 M
phosphate, pH 7.5. Rechromatography of each fraction resulted in single peaks whose \( R_f \) values remained unchanged. Two fractions were obtained by column chromatography whether or not transferrin was iron saturated or iron depleted. No difference in the total iron binding capacity (TIBC) of the two fractions was observed. Under optimal conditions 1.35 \( \mu \text{g } \text{Fe}^{3+} \) per mg of each fraction was bound. After separation TFI had a more intense pink color than TFIiI presumably due to the higher ionic strength of the phosphate buffer required to elute this component. There was no difference in mobility of the fractions on starch gel electrophoresis or paper electrophoresis at pH 8.6. Sedimentation velocity experiments at 0.8 \%\text{v} protein concentration yielded homogeneous peaks and an \( S_{20,w} \) value of 5.0 for both TFI and TFIiI. This is in good agreement with published data (34).

N-terminal amino acid determinations yielded 1 mole of valine per 90,000 g protein for both TFI and TFIiI. Amino acid analysis of the two fractions showed no difference in amino acid composition. However, considering the high molecular weight only very large differences in amino acid content would be detectable.

Immunelectrophoresis of TFI, TFIiI and TFI plus TFIiI gave a single smooth precipitation line when tested against either rabbit anti-human serum or rabbit anti-human transferrin serum. Rabbit anti-TFI and anti-TFIiI sera were prepared. In immunelectrophoresis TFI and TFIiI gave similar precipitation patterns whether the homologous or heterologous antiserum was used. If precipitation in agar gel was performed with TFI against TFIiI adsorbed anti-TFI and with TFIiI against TFI adsorbed anti-TFIiI no precipitation line was formed. Thus this immunological technique did not reveal any differences in the chemical or physical structure of TFI and TFIiI.

The possibility was considered that one of the transferrin fractions represents a denatured form of the other. This was tested in several ways. Acid-base titration curves of the fractions revealed no differences. In addition, the ultraviolet absorption spectra of TFI and TFIiI were identical. The only assay of the biological activity of transferrin is based on its capacity to incorporate iron into reticulocytes (56). The assay has been shown to be a sensitive indicator of the extent of protein denaturation (57). Activity is expressed as \( \mu \text{g } \text{Fe} \) incorporated per min per ml reticulocyte suspension under defined conditions. Both TFI and TFIiI gave similar iron incorporation values which were close to the value obtained with a control consisting of fresh native serum transferrin. Thus by many criteria the two fractions appeared to be very similar.

Carbohydrate analyses showed that TFI has a slightly higher content of hexose, hexosamine and sialic acid than TFII. Based on a molecular weight of 90,000 for transferrin, the sialic acid values correspond to 5 and 4 moles of sialic acid per mole of TFI and TFII respectively. The latter value has been published as representing the sialic acid content of transferrin (40, 42). The difference
in sialic acid content should increase the negative charge of TfI over that of TfII and one would expect to find a difference in their mobility on starch gel electrophoresis. It seemed possible that differences in carbohydrate composition could account for the chromatographic separation of the two fractions. However, contamination of TfI with mucoprotein could not be excluded as the source of its higher carbohydrate content. More physical-chemical information is still required to explain the separation of the two proteins on CM-cellulose.

The separation obtained could have been due in part to the ethanol fractionation or the fact that the starting material was fractionated from pooled human sera or plasma. Therefore whole human serum from individual donors was chromatographed on CM-cellulose columns using exactly the same procedure as before. The first two peaks to emerge were eluted in approximately the same positions as TfI and TfII and both peaks contained transferrin as judged by agar gel precipitation tests. These fractions were also capable of binding Fe³⁺. Sera from ten adults and from four umbilical cords were examined by the same chromatography technique. In all these sera the two transferrin fractions were found, although the relative amounts of each seemed to vary among the donors. Woodworth has recently reported that two transferrin fractions can be separated by gradient elution on DEAE-Sephadex (58). One of the fractions contained small amounts of hemopexin. These results support our findings that two kinds of transferrin exist in normal human serum.

**Purification of transferrin (II)**

The first peak to emerge from a DEAE-Sephadex column equilibrated with a weak neutral buffer is pure γ-globulin. The second peak, appearing during elution with increasing ionic strength, contains all the serum transferrin, as well as some α₂-lipoprotein, γ-globulin and hemopexin. Gel chromatography of this peak on Sephadex G-200 separated α₂-lipoprotein and γ-globulin from the transferrin fraction which was still contaminated with hemopexin. When this peak was subjected to recycling chromatography (59) on the same column in 0.1 M Tris-HCl, 0.2 M NaCl, pH 8.0, pure transferrin was obtained in a single peak in the third cycle. Although hemopexin has about the same molecular weight as transferrin, it emerged in front of transferrin in this system. There was a striking difference in separation depending on whether the recylization buffer contained 1 M NaCl or 0.2 M NaCl. In the first case only one peak was obtained, containing both transferrin and hemopexin. The emergence of hemopexin in front of transferrin in the latter case was interpreted as being due to formation of hemopexin aggregates. During preparation of transferrin it was also observed that plasma saturated with iron gave higher yields of transferrin than untreated plasma. This was interpreted
as being due to stabilization of the transferrin molecule in its native form by the iron atoms. Polyacrylamide gel electrophoresis of the transferrin preparations showed no visible contamination. Spectral analysis of an iron-saturated sample confirmed that the single peak in the third cycle was pure transferrin with the characteristic absorption maximum of 469 mµ. The spectrum of samples earlier in the purification procedure showed a characteristic absorption maximum of hemopexin at 415 mµ. The recovery of transferrin was calculated from the total iron binding capacity (TIBC) of the serum used as starting material assuming that transferrin binds 1.3 µg Fe/mg. With this method the yield of transferrin was estimated to 83—89 %.

Separation of individual transferrins from genetic variants (II)

The individual proteins from transferrin variants were successfully separated by two techniques: starch gel electrophoresis and isoelectric separation. For structural studies other authors have used normal transferrin CC, transferrin from individuals homozygous for a transferrin variant or components of heterozygous variants separated on TEAE-cellulose (19, 49). Neither TEAE-cellulose nor DEAE-Sephadex chromatography gives complete separation of individual transferrins. Preparative electrophoresis on polyacrylamide gel (60) at pH 8.9 gave very low recovery. The best results in earlier experiments were obtained with preparative starch gel electrophoresis. Up to 60 mg of isolated Tf CD1 and B1C or 40 mg of B2C could be separated. The separation was visualized by the intense pink color of the two finally separated bands. The difficulty with this method is in the recovery of the protein from the gel. The colored bands were cut out and eluted by electrophoresis in Pevikon block (61). By special arrangements the transferrin fractions were concentrated in a narrow zone, from which the protein could be recovered by washing with small volumes of buffer. The transferrin fractions also contained small amounts of starch, which was removed by DEAE-Sephadex chromatography. This technique gave about 70 % recovery of each protein.

In recent experiments the individual transferrins from a heterozygote have been separated by isoelectric separation as described by Svensson (62) and Vesterberg and Svensson (63). The working principle of this method is the focusing of the proteins at their respective isoelectric points in an ampholyte medium at low ionic strength. Synthetic ampholytes (poly-amino-poly-carboxylic acids) giving a pH gradient between pH 5.5 and 6.5 were used for the final separation of the components of a transferrin preparation obtained by DEAE-Sephadex chromatography and gel chromatography. There was no resolution unless 4M urea was incorporated into the ampholyte-sucrose gradient. Three main eluted peaks were obtained after separation of 25 mg Tf CD1 on a 110 ml column. One which focused at pH 6.36 was shown by starch gel electrophoresis to contain unseparated material. The other two
fractions focused into two clearly separated bands at pH 6.01 and 5.98 and were shown to contain Tf D₁ and Tf C respectively. The recovery of each protein was 72 %. Starch gel electrophoresis of ⁵⁹Fe-labelled fractions showed pure preparations of the individual transferrins with retention of the capacity to bind iron. The lack of separation in the absence of urea may be due to the formation of transferrin aggregates in the low ionic strength ampholyte system.

One advantage of isoelectric separation over preparative starch gel electrophoresis is that the carrier ampholytes are of low molecular weight and can be efficiently removed by dialysis. Although electrophoresis and ion-exchange chromatography were used to remove starch after electrophoretic separation this step will always be critical and small amounts of starch may remain and interfere with analyses of the carbohydrate content of the preparation.

*Ultracentrifugation and gel chromatography studies of transferrin subunits (III).*

Transferrin has been assumed to consist of a single polypeptide chain (5). However, the molecule binds two atoms of iron and structural data from similar metal binding proteins suggest by analogy that transferrin may consist of a pair (or pairs) of subunits. The finding of two identical heterosaccharide chains (42) supports the theory of a single glycopeptide for each subunit. The presence of two polypeptide chains was also suggested by the fact that fingerprints of performic acid oxidized transferrin contained only half the expected number of tryptic peptides (III).

Earlier work showed a radical change in the sedimentation coefficient of native human transferrin after reduction and alkylation in 8M urea (64). This procedure was performed essentially according to Crestfield et al (36). Iron-free transferrin was reduced with β-mercaptoethanol in a Tris-HCl buffer, pH 8.6, 8M in fresh deionized crystalline urea. Alkylation was performed with iodoacetic acid. The reduced and alkylated material was either dialyzed against water and freeze-dried or desalted by gel chromatography on Sephadex G-25. Material prepared with the former technique was homogeneous, while desalting on Sephadex G-25 always gave a polydisperse product as shown by sedimentation velocity experiments. The dialyzing technique was therefore preferred for preparing reduced and alkylated transferrin.

Starch gel electrophoresis of reduced and alkylated material in 8M urea-glycine buffer (65), pH 7—8, or in 8M urea-formate buffer (66), pH 3.9, did not show any separation of the suggested subunits. Gradient elution of the same material on a CM-cellulose chromatography column equilibrated with a phosphate buffer containing 8M urea yielded two well separated peaks. These two fractions had the same mobility on starch gel electrophoresis in both types of buffer. N-terminal amino acid determinations (37) showed that
the main N-terminal, valine, occurred in the same amount in both fractions. The two peaks obtained were presumed to be identical with the two fractions of normal transferrin (TfI and TfII) observed earlier (I).

Sedimentation velocity studies of reduced and alkylated transferrin desalted by dialysis and freeze-dried gave homogeneous peaks. The sedimentation coefficients varied but molecular weight determinations (67) gave values corresponding to half the molecular weight of native transferrin. Ultracentrifugation of the same material in 0.1 M Tris-HCl, pH 8.6, at four different protein concentrations yielded an $S_{20,w}^0$ value of 1.24 S when extrapolated to zero concentration. After the material had been dialyzed extensively against the same buffer for 72 hours, four diffusion coefficients were determined at different protein concentrations. From this study the $D_{20,w}^0$ value was calculated to be $2.80 \times 10^{-7}$ cm$^2$/sec. This value, when substituted into the Svedberg equation (68) along with the $S_{20,w}^0$ value, gave a molecular weight of 39,400. Molecular weights were also determined at constant protein concentrations by use of the Archibald method. Small variations were obtained, 41,500 by the method of Trautman (69) and 41,000—44,000 by the method of Ehrenberg (67). These results indicate that normal human transferrin, with a molecular weight close to 80,000, consists of two subunits, each having a molecular weight of about 40,000. Transferrin appears to contain one free N-terminal amino acid, valine, per mole. This suggests the presence of a blocked N-terminal amino acid in one of the subunits.

Preliminary ultracentrifugation studies showed that treatment with 8M urea alone brought about a decrease in the sedimentation coefficient to 3 S (64). The molecular weight determination on this material in a Tris-HCl buffer gave a value of 78,000—79,000, i.e. the same as that reported for native transferrin (34). Schwick recently reported that half the molecular weight of transferrin was obtained when ultracentrifugations were performed in 8M urea (70). This suggests that S-S bonds are not involved in holding the subunits together. Therefore noncovalent bonds can be presumed to have that function in the native transferrin molecule.

Attempts were made to correlate the molecular weight of the transferrin subunits to the elution volume after gel chromatography on Sephadex G-200 in 0.1 M Tris-HCl buffer, pH 8.6. A column was calibrated by determining elution volumes for 6 isotopically labelled proteins: human $\gamma$-globulin, transferrin, albumin, ovalbumin, pepsin and ribonuclease. Three labelled proteins could be chromatographed simultaneously with reduced and alkylated transferrin, the latter being registered continuously at 280 m$\mu$ (71). However, gel chromatography of the subunits of transferrin gave an unexpected result in that a symmetrical peak preceded those of native transferrin and $\gamma$-G globulin. The elution volume for each of the six isotopically labelled proteins was then plotted according to Andrews (72) on a graph against the logarithm of the
molecular weight for each protein. The curve was consistent with the method of Andrews, who stated that elution volumes of most globular proteins are directly proportional to the logarithms of their molecular weights. However, when the molecular weight of reduced and alkylated transferrin was calculated from the elution volume by means of the graph, an apparent molecular weight more than twice that of native transferrin was obtained.

Laurent and Killander (73) have postulated that the separation by gel chromatography of molecules varying in size can be explained as a steric exclusion of solutes from the gel phase. Their parameter $K_{av}$, which is the fraction of the gel phase volume available for the solute, was determined from the elution volume, the void volume and total volume of the column. The radius, $r_s$, of an equivalent sphere was calculated from the diffusion constants of native transferrin and reduced and alkylated transferrin using Stoke's formula (74). The calculations give an estimation of the molecular size of a protein.

The $K_{av}$ values of the test proteins, including native transferrin, were almost identical with those calculated by Laurent and Killander. The $K_{av}$ and $r_s$ values calculated for native and for reduced and alkylated transferrin coincided almost exactly with the theoretical curve of the gel chromatography behaviour given and tested by the authors.

The transferrin subunits tested were found to have a larger molecular size than native transferrin, i.e. the Stokes radius was $76.7 \times 10^{-8}$ cm versus $36.7 \times 10^{-8}$ cm for the native protein. The gel chromatography profile of the transferrin subunits was thus in good agreement with one of the recently published theories of gel chromatography (73).

The above results show that the transferrin molecule has dissociated into two subunits which have swelled to a larger molecular size. The swelling may be a secondary phenomenon to disruption of S-S bonds, hydrogen bonds and/or hydrophobic bonds which causes the molecule to unfold and become more hydrophilic than the compactly folded native molecule.

**Cyanogen bromide degradation of transferrins (V)**

Earlier attempts to show differences in the primary structure of genetic variants of transferrin were made using the whole transferrin molecule. To investigate whether amino acid substitutions have occurred among genetic variants, the large molecule was degraded to smaller fragments by cyanogen bromide (CNBr). Transferrin contains nine methionine residues (35). Assuming that CNBr specifically attacks the sites of methionine (75), ten fragments should be obtained from the transferrin molecule under optimal conditions. The material degraded in 0.1 M HCl was non-reduced, thus the S-S bonds were intact. Gel chromatography of the split products on Sephadex G-75 in 0.1 M formic acid resulted in three peaks, which in order of elution, were
designated CNBr I, II and III. The first peak, eluted with the void volume, was broad and asymmetric. It was assumed to contain intact transferrin or aggregated fragments. Rechromatography of the material on Sephadex G-75 in an ammonium bicarbonate buffer yielded two fractions, Ia and Ib. Rechromatography of CNBr II and CNBr III on Sephadex G-75 in ammonium bicarbonate buffer resulted in single symmetric peaks with Rf-values corresponding to those in the first gel filtration in formic acid. Almost 100% of the applied material was recovered from the column. The chromatographic profiles of the separated fractions were similar for degraded Tf CC, Tf CD\textsubscript{1}, Tf B\textsubscript{1}C, Tf B\textsubscript{2}C or their isolated individual phenotypes. These results suggested structural identities between normal transferrin CC and its genetic variants.

In starch gel electrophoresis in urea-glycine buffer, (65) the fragments migrated with different mobilities towards the anode. CNBr I produced a smear and CNBr II and III each produced two bands. If the gel contained β-ME, CNBr I separated into several bands whereas CNBr II and III still produced two bands. Of the bands from CNBr I in the β-ME gel, two bands were in positions corresponding to those of CNBr III. CNBr II and CNBr III also yielded two N-terminal residues each, which may correspond to the two bands of each fragment seen on starch gel electrophoresis. From these results it can be assumed that the two bands of CNBr II and CNBr III represent corresponding parts of the two polypeptide chains. (Fig. 1.) The two band patterns were obtained whether or not the gel contained β-ME as reducing agent.

\begin{center}
\begin{tabular}{c|c|c}
  & I & II & III \\
  Val & Lys & Gly & /
  Val & Gly & /
  Site of CNBr Cleavage
\end{tabular}
\end{center}

\textit{Fig. 1.} A schematic representation of one possible arrangement of the CNBr fragments in the intact transferrin molecule. The order of fragments I—III is not known.
Sedimentation velocity experiments on the degradation products CNBr Ib, II and III in 0.1 M Tris-HCl, pH 9.0, indicated that the preparations were homogeneous with respect to molecular weight. Each fraction gave a symmetrical peak, the $S_{20,w}$ values being 3.9 S, 1.9 S and 1.2 S respectively. The molecular weights of CNBr Ib, II and III, determined by the approach to sedimentation equilibrium method (67), were calculated to be 61,000, 26,000 and 17,400 respectively. The sum of the molecular weights exceeds that of native transferrin. This indicates that cyanogen bromide degradation of the molecule was incomplete or that there were aggregates of split products. Iron incorporation studies showed that CNBr Ib retained the iron binding capacity.

N-terminal amino acid analyses (III, IV, V)

The N-terminal amino acids were determined as phenylthiohydantoin derivatives (37) and analyzed by thin-layer (IV) or paper (76) chromatography. Determinations were performed on five different intact normal transferrin preparations, on reduced and alkylated transferrin and on CNBr Ib, CNBr II and CNBr III (V). Assuming a molecular weight of 88,000 for transferrin the amount of N-terminal valine determined on different preparations was 0.8—0.9 residues per molecule. In addition small amounts (0.1—0.2 moles/mole) of N-terminal aspartic acid, glycine and serine were found in our own preparations and in a preparation obtained from Kabi (Stockholm, Sweden). One of the two preparations from Behringwerke (Marburg/Lahn, Germany) contained these three N-terminal amino acids in addition to valine, whereas the other preparation contained only N-terminal valine and a small amount of aspartic acid. A sample recently obtained from Dr. P. Aisen (23), prepared from Cohn fraction IV-7, yielded 0.91 mole valine and 0.15 mole aspartic acid per mole transferrin. N-terminal determinations on reduced and alkylated transferrin yielded 0.9 mole valine/mole and 0.3 moles of each of the additional amino acids (III). The source of the additional N-terminals is obscure. All preparations were pure judging from immunoelectrophoresis, polyacrylamide gel electrophoresis and absorption spectra.

The extraneous N-terminal amino acids may have been due to contamination, partial hydrolysis of susceptible peptide bonds during preparation or to the existence of a four chain form of transferrin. The analytical differences between our preparations, Behringwerke’s, and that from Dr. P. Aisen argue against the interpretation that the additional N-terminals represent parts of a multichain structure.

The N-terminal amino acids of CNBr Ib were valine, tyrosine, glycine, aspartic acid and serine (V). Based on a molecular weight of 61,000, the preparation contained approximately half a mole each of valine and tyrosine, the latter being a new N-terminal not present in the original transferrin
preparation. In addition 0.8 mole of glycine, 0.3 mole of aspartic acid and 0.3 mole of serine were found. At least some of the N-terminal aspartic acid, glycine and serine must have been present as free N-terminals in the original preparation since none of them appears in the other two fragments. CNBr II had valine and lysine as N-terminals in almost equal proportions, one mole each per mole of the fragment. CNBr III yielded N-terminal glycine 1.7 mole/mole protein. The two N-terminal residues found in each of CNBr II and III may correspond to the two bands formed by these preparations in starch gel electrophoresis and thus to the segments of the two polypeptide chains. (Fig. 1)

A modified method (77, 78) for isolating the characteristic peptides from fingerprints as phenylthiocarbamyl (PTC) derivatives provided sufficient material for amino acid analyses and sequence determinations on the isolated peptides after trypsin digestion of CNBr III (V). After electrophoresis and chromatography the N-terminal amino acid in each peptide on the paper is allowed to react with phenylisothiocyanate according to the Edman procedure. This coupling step permits location of the peptides on the paper under ultraviolet light, eliminating the use of ninhydrin. After eluting the peptides one can complete the formation of the N-terminal phenylthiohydantoin derivative in each, isolate the N-terminal amino acid and continue with the sequence determination. Or, following the isolation of the N-terminal amino acid, the remainder of the peptide can be hydrolyzed for amino acid analysis. The yields were 0.1—0.15 μmoles of each isolated peptide from 4 fingerprints of CNBr III.

A thin-layer chromatography technique has been developed whereby minute amounts of almost all PTH amino acids can be identified on the same sheet chromatographed in two systems (IV). The method described is especially designed for identification of N-terminal amino acids released by the Edman degradation method on small amounts of peptide material. The lower limit at which the derivatives can easily be detected is $5 \times 10^{-3}$ μmoles. The method is about 2—3 times more sensitive than the paper chromatographic technique and entails a single application of the sample in contrast to the four applications required for paper chromatography as developed by Sjöquist (76).

The thin-layer chromatography is performed on precoated, flexible Eastman Chromagram Sheets with an incorporated fluorescent indicator. The spots can therefore be located under UV light. Six groups of PTH amino acids are obtained. If the sample contains not more than one member of each group, these can be identified with certainty. Furthermore, the characteristic color development at 100° of some of the derivatives aids in their identification. The advantage of thin-layer chromatography lies in its rapidity and the small amounts of material required for qualitative identification of the PTH amino acids.
Fingerprinting of native and degraded transferrins (II, III, V)

Fingerprints were always performed on trypsin digested material. Two samples were applied, one on either side of the midline of a sheet of whole Whatman 3MC paper. High voltage electrophoresis was performed at pH 6.1 on a water cooled plate. The paper was divided and each half was sewn to a new Whatman paper and chromatography was performed in a second dimension (79). The fingerprints of a transferrin sample were identical whether the sample was digested with trypsin in a pH stat, in 0.01 M NH$_4$OH, pH 8.4, or in 1 % NH$_4$HCO$_3$:NH$_3$, pH 8.2. During digestion in NH$_4$OH for 5 hours the pH dropped to 7.9. Digestion in this solution was preferable because the sample was easily rendered salt-free by freeze-drying. Fingerprints of material so digested were always very distinct.

The presence of two polypeptide chains was suggested by the fact that fingerprints of reduced and alkylated or performic acid oxidized transferrin, developed by ninhydrin, contained less than half the number of tryptic peptides predicted from the arginine and lysine contents (III). Furthermore, specific staining for arginine and histidine showed about one half the number of positive peptides predicted by amino acid analyses.

If a duplicate fingerprint was developed by the chlorine-iodine reaction one additional spot was obtained. Transferrin appears to contain only one free N-terminal amino acid, valine, per mole if one ignores the small amount of additional amino acids. Ninhydrin failed to react with a tryptic peptide spot which was positive in the chlorine-iodine reaction. This suggests the presence of a blocked N-terminal amino acid in one of the subunits if that peptide is one of the N-terminal peptides. Such a block may be due to the presence of an acyl, a pyroglutamic acid residue or a cystine. Since preliminary amino acid analyses of the peptide were negative for glutamic acid and cysteic acid, pyroglutamic acid or cystine can be excluded as N-terminal amino acids. These results suggest that the primary structure of the subunits must be similar except for their N-terminal amino acids.

Analyses of transferrins C, B$_2$, D$_1$ and D$_3$ showed no differences in their amino acid compositions (19). Substitutions of only a few amino acid residues would not be detectable by this type of analysis of a molecule approximately 90,000 in molecular weight. The analytical error of 3 % would not permit the detection of real differences in the content of any of the following amino acids, replacement of which is most likely to lead to a change in charge: 86 lysine residues, 22 histidine residues, 30 arginine residues, 89 aspartic acid residues and 79 glutamic acid residues (35). Asparagine and glutamine are included in the values for aspartic and glutamic acids respectively. The whole transferrin molecule contains approximately 827 amino acid residues. It was hoped that peptide patterns as seen on fingerprints would permit correlation
with the genetic variant and its electrophoretic mobility on starch gel electrophoresis.

Fingerprints were made of isolated individual transferrins from two blood donors of Tf CD$_1$, two of Tf B$_1$C and one of Tf B$_2$C. The gross structure of the individual proteins as judged from the fingerprints was similar. A few peptides were found, however, which were characteristic for each individual protein (II). This may indicate that amino acid substitutions have occurred during evolution.

Further experiments showed that one of the substituted peptides of Tf CD$_1$ is included in a low molecular weight fragment (CNBr III) resulting from cyanogen bromide degradation of isolated individual transferrin D$_1$ (V). Fingerprints of transferrins CC and CD$_1$ are shown in Fig. 2. One difference was observed, i.e. a peptide designated T10 was found in CNBr III of Tf CD$_1$

![Fig. 2. Fingerprint tracings of CNBr III-fragments from Tf CC, Tf CD$_1$, Tf CCD$_1$, Tf D$_1$CD$_1$, Tf B$_1$C and Tf B$_1$B$_2$C. The numbered spots were isolated and further characterized.](image)
but not in the corresponding fragment of Tf CC. Neither was this peptide found in the CNBr III fragment of Tf B1C or Tf B2C. Peptide T10 may correspond to the characteristic peptide found in fingerprints of native Tf D1. This peptide had the same electrophoretic and chromatographic characteristics as T10 in the CNBr III fragment. The fingerprints of isolated Tf CCD1 contain peptides T7 and T9 which are not found in Tf D1CD1. The latter has the characteristic peptides T6, T8 and T10. Peptide T10 is specific for CNBr III of Tf CD1 or Tf D1CD1, whereas peptides T5, T6, T7, T8 and T9 are found in both Tf CC and Tf CD1.

Peptides T5 — T10 were isolated as phenylthiocarbamyl (PTC) peptides directly from fingerprints of CNBr III from Tf CD1 (77, 78). T5 and T10 are heptapeptides with identical amino acid compositions except for the N-terminal amino acids. Stepwise degradation of the peptides showed the following sequences; T5: Asp-Leu-Leu-Phe-Leu-Phe-Lys, T10: Gly-Leu-Leu-Phe-Leu-Phe-Lys. Thus, T5 and T10 differ only in the N-terminal position. The difference is compatible with a single point mutation in the trinucleotide sequence for aspartic acid (GAU or GAC) resulting in a replacement of this amino acid by glycine (GGU or GGC), (80). This amino acid substitution may explain the difference in mobility of transferrins CC and CD1 in starch gel electrophoresis.

The two chain structure of transferrin (III) is further supported by the results summarized in Fig. 3. It has been shown that the amount of each of the individual fractions in a genetic variant corresponds to half the amount

Fig. 3. Schematic structure of the two molecular species isolated from transferrin CD1. Each molecule is represented as consisting of two polypeptide chains with different N-terminal amino acids, valine and a blocked one. In one of the chains of species D1 an aspartic acid residue has been replaced by glycine. Indicated numbers (5 and 10) refer to isolated tryptic peptides T5 and T10.
of normal transferrin CC. On N-terminal analyses of peptides T5 and T10 from the fingerprints of fragment CNBr III of isolated Tf D1 equal amounts of aspartic acid and glycine respectively were obtained. The same analyses of CNBr III from the heterozygote Tf CD1 showed 2.5—3 times as much aspartic acid in T5 as there was glycine in T10. The results are compatible with the structure of transferrin shown in Fig. 3. The heterozygote Tf CD1 is represented as consisting of one mole Tf C per mole Tf D1, each protein consisting of two polypeptide chains. Each chain of Tf C contains T5, and in Tf D1 one chain contains T5 and the other the amino acid substituted T5, i.e. T10 (ref. 81).

If corresponding fingerprints of B1C and B2C were analyzed for the peptides T6—T9 more information was obtained (Fig. 2). Fingerprints of Tf B1B1C and B2B2C both contained peptides T7 and T9, whereas T6 and T8 were missing. Fingerprints of Tf C1C and C2C both contained peptides T6 and T8, but lacked T7 and T9. Thus the peptide pairs T6, T8 and T7, T9 are represented in another pattern, if one compares corresponding fingerprints of a transferrin C fraction and a substituted fraction from a slow moving variant. Thus transferrin C fractions from heterozygotes with slow or fast moving fractions differ in their primary structure. All C-fractions have the same mobility on starch gel electrophoresis as the normal Tf CC. The difference in the primary structure can be interpreted as indicating that there are two genes responsible for synthesis of two normal transferrins. Tf CC would then be heterogeneous and can be written Tf C'C". Thus B1C could instead be written as B1C" and CD1 as C'D1. Mutations in the C' and/or the C" genes would give the variants in the B-group and D-group. These two different molecules, C' and C", may be identical with TfI and TfII, whose relation and characterization have previously been described (I).

GENERAL SUMMARY

1. Ethanol fractionated normal human transferrin separated into two fractions when chromatographed on cation-exchange cellulose. The two fractions showed similarities in their immunological, biological and physical properties. There were some differences in their carbohydrate compositions. The possibility that one of the two fractions resulted from the denaturation of the other was excluded by several methods. The same separation of transferrin occurred if whole sera from individual donors were subjected to the same chromatographic procedure (I).

2. Human transferrins were isolated by chromatography on DEAE-Sephadex followed by recycling chromatography on Sephadex G-200. Individual
genetic variants of human transferrin were then completely separated by preparative starch gel electrophoresis or by isoelectric separation in an ampholyte-urea sucrose gradient. The over-all recovery was 70% and the preparations were pure by the criteria of immunoelectrophoresis polyacrylamide gel electrophoresis and absorption spectra. Fingerprints of tryptic peptides from some individual transferrins showed each of them to have a characteristic peptide map, suggesting differences in their primary structures (II).

3. Evidence that human transferrin consists of at least two subunits was obtained by physical and chemical studies of reduced and alkylated transferrin. The molecular weights of the subunits were equal and were estimated to be half the molecular weight of native transferrin by sedimentation and diffusion and by the approach to sedimentation equilibrium methods. The sedimentation coefficient was found to decrease from 5S to 1.2S after reduction and alkylation in 8M urea. Fingerprint patterns of trypsin digested, reduced and alkylated material showed approximately half the number of peptides predicted from the arginine and lysine content of native transferrin (III).

4. Normal transferrin CC and a transferrin variant CD1 were degraded by cyanogen bromide, which specifically splits peptide chains at the sites of methionine residues. Gel chromatography of the reaction mixtures separated in each case three fragments with molecular weights of 61,000, 26,000 and 17,400. Fingerprints of tryptic peptides from fragments of isolated transferrins C and D1 showed a difference in the peptide pattern of the smallest fragment (V).

5. A thin-layer chromatography method has been developed for identifying small amounts of phenylthiohydantoin derivatives of amino acids. The technique has been used in the analyses of products obtained by Edman degradation of small amounts of peptides isolated from fingerprints (IV).

6. Amino acid analyses and amino acid sequences of some of the peptides characteristic for a genetic variant showed at least one amino acid substitution, glycine replacing aspartic acid, which distinguishes transferrin D1 from transferrin C (V).
ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to Dr. John Sjöquist for all valuable advice and stimulating discussions throughout this investigation.

Thanks are also due to Professors S. Bergman, H. G. Boman, B. Borgström, T. C. Laurent and K. G. Paul for their interest and for the facilities they have placed at my disposal.

The skilful technical assistance of Miss R. Högberg, Miss I. M. Jönsson and Mr. T. Svedberg is gratefully acknowledged. I am very much indebted to Dr. B. Steele for revision of the English text and for helpful discussions and to Mrs. K. Gladh and Miss R.-M. Ringlöw for their skilful secretarial assistance.

I also extend thanks to Dr. H. Pertøft for carrying out the ultracentrifugations and to Mrs. I. Sjöquist for her valuable help with amino acid and N-terminal analyses.

This investigation was supported by grants from the Swedish Medical Research Council (project no 13x-620-01 and 13x-620-02), The Medical Faculty, University of Lund, The Medical Faculty, University of Umeå, Svenska Sällskapet för Medicinsk Forskning, Stockholm, and J. C. Kempes Minne, Umeå.

REFERENCES


25
32. A. Bezkorovainy and M. E. Rafelson Jr., *Arch. Biochem.*, 107 (1964) 302
43. O. Smithies, *Nature*, 181 (1958) 1203
58. R. C. Woodworth, personal communication

26
73. T. C. Laurent and J. Killander, *J. Chrom.*, 14 (1964) 317
78. J. Sjöquist, I. Sjöquist and J. O. Jeppsson, to be published