

The Primary Structure of Hen Ovotransferrin

John WILLIAMS, Thomas C. ELLEMAN, I. Barry KINGSTON, Adrian G. WILKINS, and Katherine A. KUHN

Department of Biochemistry, University of Bristol

(Received August 26, 1981)

Peptide sequences obtained from hen ovotransferrin are compared with the complete amino acid sequence of the protein deduced from a cDNA sequence (Jeltsch and Chambon, preceding paper). Of the 705 positions of the whole protein 605 can be matched by the peptide sequences. Some possible discrepancies between the two methods are pointed out. The two halves of the chain show marked similarities in their sequences with 37% identical residues. The positions of the 15 disulphide bridges are shown; there are 6 homologous bridges in each half of the molecule and 3 extra bridges which occur only in the C-terminal half. The terminal residues of the half-molecule fragments obtained by limited proteolysis are identified. The two domains are joined by a 9-residue connecting peptide. Sequence variability has been found at 9 positions. The sequence of hen ovotransferrin is compared with the partial sequence available for human transferrin. From this some tentative conclusions about the identities of the metal-binding residues and about the evolution of transferrin are reached.

The iron-binding protein transferrin is present in all vertebrate species. The closely related protein lactoferrin (or lactotransferrin) appears to be restricted to mammals. The amino acid sequences of these proteins have been studied for many years in order to elucidate their structure, evolution and iron-binding properties. Brew and his colleagues [1, 2] have determined about 80% of the sequence of human serum transferrin and have discovered that the N-terminal and C-terminal halves of the chain have similar sequences, with about 40% identical residues. X-ray crystallography [3] and partial proteolysis experiments [4, 5] show that each half of the chain folds up into a largely independent, compact domain which carries a specific metal-binding site. Since the domains can be isolated as single-chain fragments after partial proteolysis it can be inferred that the disulphide bridges are located within the domains, none passing between them. These data suggest that the present-day bilobar protein has evolved by gene duplication and subsequent fusion from a smaller ancestral protein, which apparently no longer exists. Partial sequence data on human lactoferrin show that this protein is homologous to transferrin and that it too possesses internal duplication of structure [6, 7]. Lactoferrin may, thus, have arisen by a further duplication of the transferrin gene in the ancestors of the mammals.

In birds expression of the transferrin gene in the oviduct, which is under steroid control, leads to the synthesis of ovotransferrin (conalbumin), which differs from serum transferrin only in the nature of the attached glycan [8, 9]. The sequences of cysteine-acid-containing peptides from digests of performic-acid-oxidized ovotransferrin have been determined [10]; 34 unique cysteine acid residues were identified and some of the disulphide bridges were characterized after pepsin digestion of the unoxidized protein. A 102-residue sequence, which includes the point of attachment of the glycan, was later determined [11]. Further studies were carried out in this laboratory several years ago with the object of determining the

complete amino acid sequence but these were not concluded and have not been published. Recently, however, Jeltsch and Chambon [12] have determined the complete nucleotide sequence of hen transferrin cDNA transcribed from ovotransferrin mRNA. In this paper we will compare the structure of hen ovotransferrin as deduced from cDNA with previous studies of the protein and discuss the possible significance of certain aspects of this structure.

THE AMINO ACID SEQUENCE

Fig. 1 shows the amino acid sequence of hen ovotransferrin, in the usual one-letter code, as deduced from the cDNA sequence [12]: it will be referred to as the DNA/AA sequence. It comprises 705 residues including a 19-residue presequence. The N-terminal residue of the secreted protein is A(1). The amino acid composition of the DNA/AA sequence, excluding the pre-sequence, is shown in Table 1 where it is compared with the average of several published amino acid analyses. Discrepancies are greatest for isoleucine, serine, threonine and tryptophan where they exceed 10%. The deduced composition corresponds to a molecular weight of 77770, a figure which includes 1900 for the glycan [13], which is close to the value of 76000 obtained long ago from the iron content of the diferric protein [14].

PEPTIDE SEQUENCES (Fig. 1)

Pretransferrin Sequence (t)

Palmiter and his colleagues [9] determined the sequence of pretransferrin, synthesised *in vitro*, by automatic sequential Edman degradation from M(-19) to P(3).

N-Terminal Sequence (g)

The N-terminal sequence of ovotransferrin was found by the maleylation method to be APPKS [18].

Abbreviations. The one-letter system for abbreviating amino acid residues has been used, see *Eur. J. Biochem.* 5, 151–153 (1968).

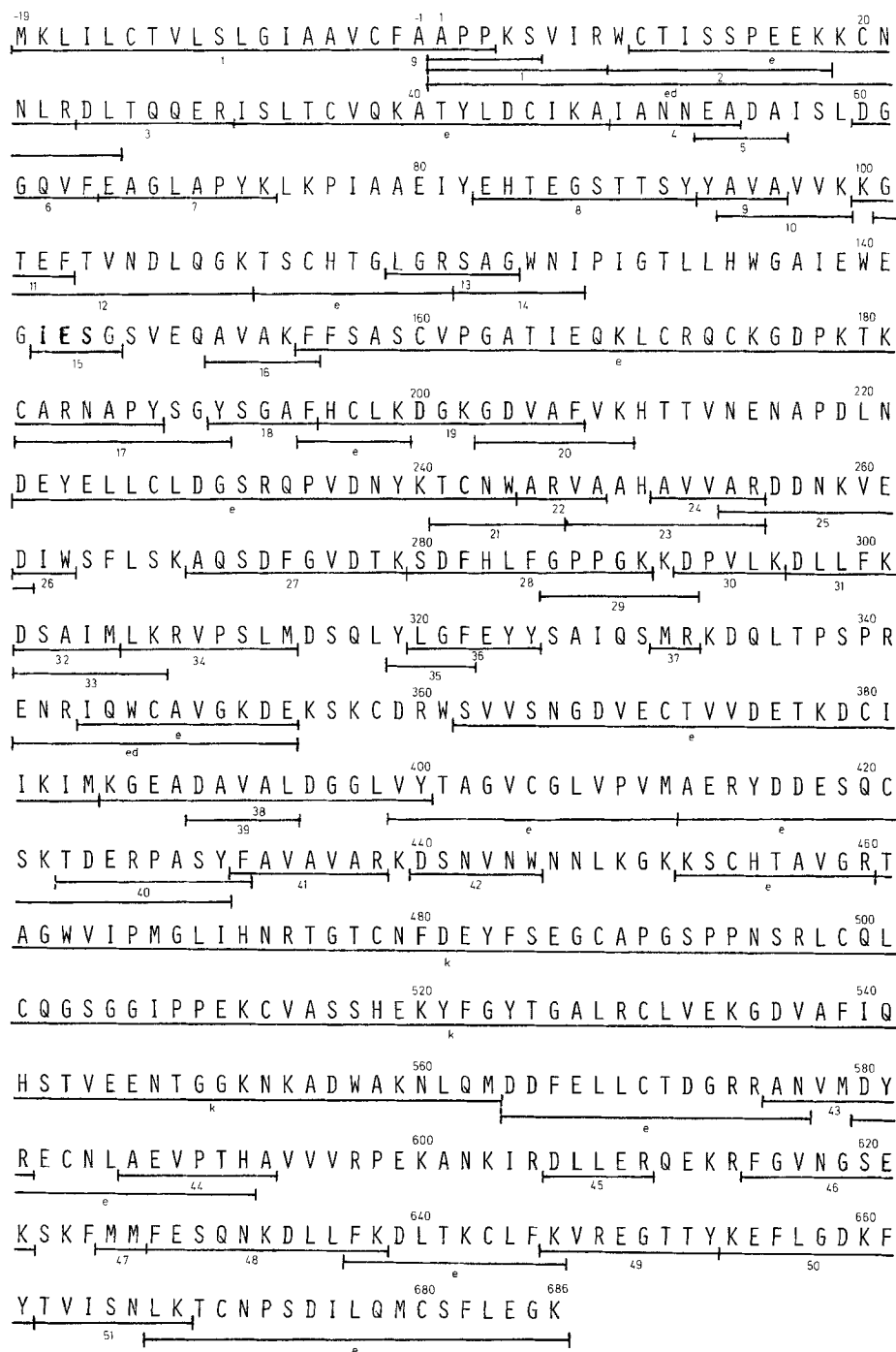


Fig. 1. The amino acid sequence of hen ovotransferrin as deduced from the cDNA sequence [12]. The residue positions are numbered above the sequence from M(-19) to K(686). The lines below the sequence indicate those parts of the structure which have been determined directly: for details see the text and Table 2. The glycan is attached to asparagine (473)

Half-Cysteine-Containing Peptides (e)

Elleman and Williams [10] determined the sequences around cysteine acid residues and found 34 apparently unique cysteine acid residues, an excess of 4 over the actual number. Four of these sequences (CVK, CGK, CDI and CMNSF) cannot be fitted to the DNA/AA sequence and two of their sequences (those containing cysteine acid residues numbers 7 and 32 in [10]) are almost certainly the same, as will be shown below. Thus, 29 of the 30 half-cystines are verified by their work and one, C(358), was missed.

Residue C(672) appears in two of the cysteine acid peptides studied by Elleman and Williams [10]. In peptide C₄ the sequence CNPSDIL, which agrees with the DNA/AA sequence, was obtained, but in peptide T₃C₄ the sequence was TCNSDPIL and Williams [19] had earlier obtained the sequence KTCNSDPIL. In both cases the evidence on which the sequence was based was completely convincing.

A further problem concerns residue P(490) in the DNA/AA sequence. In our earlier work [10, 11] the existence of a third proline near to the two located at positions 493 and 494 was evident from amino acid analyses. Nevertheless its location

Table 1. Amino acid composition of hen ovotransferrin

The amino acid analysis given is the average of published analyses [4, 8, 15–17]

Amino acid	cDNA	Analysis
	mol amino acid/mol protein	
Lysine	58	59
Histidine	12	12
Arginine	29	31
Aspartic acid	46	72
Asparagine	32	
Threonine	37	33
Serine	47	39
Glutamic acid	45	
Glutamine	22	66
Proline	28	27
Glycine	52	51
Alanine	53	54
Valine	49	47
Methionine	11	10
Isoleucine	27	24
Leucine	50	47
Tyrosine	21	20
Phenylalanine	26	26
Cysteine	30	28
Tryptophan	11	15

in the sequence C(488)APG is surprising since the dansyl Edman technique gave FSEGCAGSPNNS in the two independently isolated peptides T₂C₁ [10] and 1T₁C₁ [11]. This could be an artefact due to diketopiperazine formation leading to simultaneous removal of the alanine and proline residues during Edman degradation [20]. In both these cases the proline is misplaced by two positions but we do not know whether this reflects sequence variability or incorrect sequence determination.

The discrepancy between the DNA/AA sequence N(366)-GDV and the sequence DDG given as part of peptide C₃ by us [10], is probably due to the occasional failure of the Edman degradation to cleave asparaginyl-glycine bonds [21]. Thus the peptide sequence was deduced from amino acid analyses of subtilisin subpeptides and was not rigorously determined. The omission of Y(400) from peptide P_{3a} [10] is also erroneous. Tyrosine was detected (0.4 mol/mol peptide) by amino acid analysis but was not found during sequencing. It was, therefore, attributed to contamination. The presence of tyrosine at position 400 would explain the chymotryptic cleavage which gave rise to peptide C₆.

Glycan-Containing CNBr Fragment (k)

Kingston and Williams [11] isolated a CNBr fragment (BC_a) and determined the sequence. This matches the DNA/AA sequence from T(461) to M(563) with the exception of P(490) which has been discussed above.

Automatic Sequential Edman Degradations (ed)

A CNBr fragment, which represents the N-terminal half of ovotransferrin, designated BC_a by Kingston and Williams [11], gave the sequence APPKSVI-WC-I--E--KCNL-DL both with the Beckman liquid-phase sequencer and the Anachem solid-phase sequencer. The N-terminal sequence of the proteolytic fragment, which represents the C-terminal

domain of the protein (CT), gave the sequence EN-IQWCA-VGKDE, which corresponds to the DNA/AA sequence from 342 to 354.

Additional Peptides (Numbered)

In this section we list 51 peptide sequences which have not been published before. The peptides were obtained from enzymic digests of ovotransferrin or of CNBr fragments of ovotransferrin. The following fragments were utilised: BC_b, which is derived from the N-terminal half of the protein BC_e which contains two sequences from the C-terminal half, and F which comprises the low-molecular-weight products of the CNBr treatment [11].

Digestions were performed with trypsin, chymotrypsin, pepsin, thermolysin or subtilisin. Peptides were purified by ion-exchange chromatography on Dowex AG50X4 (mesh size 400) with a buffer gradient of pyridine/acetic acid from 0.2 M at pH 3.1 to 2.0 M at pH 5.0 [20]. Further purification was carried out by paper electrophoresis and paper chromatography and sequences were determined by the dansyl Edman method [21, 22] or by mass spectrometry.

Many of the peptides studied merely confirmed sequences which were already published. Those which give new information and their positions in the DNA/AA sequence are shown in Table 2. It will be seen that several of the cleavages are unconventional; especially common are chymotryptic cleavages in trypsin digestions. It is of interest that, although many enzyme digests were studied, some parts of the polypeptide chain were not recovered in any of the digests and taking all the peptide sequences together, published and unpublished, 605 residues can be matched to the DNA/AA sequence (86%).

Treatment of partially iron-saturated ovotransferrin with trypsin enables the N-terminal and C-terminal domains to be separated as fragments NT and CT [4, 5]. From the sequence data the boundaries of the proteolytic fragments NT and CT can be defined. Treatment of the N-terminal fragment NT with CNBr released free arginine indicating the C-terminal sequence MR. CNBr treatment of the whole protein does not release arginine. Thus, the C-terminal residue of fragment NT is R(332). The N-terminal residue of the C-terminal fragment CT is E(342) by the sequencer runs on this fragment. The two domains are thus joined by a 9-residue connecting peptide.

The ability of the cDNA sequence to predict molecular weight, amino acid composition and peptide sequences which are in close agreement with the observations made on the protein itself leaves no doubt that the DNA/AA sequence is very accurate.

SEQUENCE VARIABILITY

Intraspecific and intra-individual heterogeneity of ovotransferrin is known from electrophoresis studies [23–25]. The ovotransferrin used in our peptide studies was of the common B electrophoretic type. Nevertheless there was clear evidence of sequence variation at nine positions (Table 3). This list is clearly not complete. Further, there are some positions where asparagine or glutamine in the DNA/AA sequence appeared as aspartic acid or glutamic acid in the peptide sequences. We have not included these as evidence of sequence variation because of the ease with which deamidation can occur during peptide isolation. The DNA/AA sequence is only one of the sequences of hen ovotransferrin.

Table 2. *Unpublished peptide structures from hen ovotransferrin*

Peptide numbers and positions refer to Fig. 1. The source of each peptide is indicated. OT indicates ovotransferrin. BCb, BCE and D/E are CNBr fragments [11]. BCb is from the N-terminal domain; BCE and D/E from the C-terminal domain; NT is the N-terminal half-molecule obtained by partial proteolysis

No.	Structure	Position	Source of peptide
1	APPKSVIR	1–8	trypsin digest of tri-fluoroacetylated BCb
2	WCTISSPEEK	9–18	trypsin digest of BCb
3	DLTQQR	25–31	trypsin digest of BCb
4	IANNEA	49–54	pepsin digest of BCb
5	EADA	53–56	pepsin digest of BCb
6	DGGQVF	60–65	pepsin digest of BCb
7	EAGLAPYK	66–73	trypsin digest of BCb
8	EHTEGSTTSY	83–92	trypsin digest of BCb
9	YAVA	93–96	pepsin digest of BCb
10	AVAVVK	94–99	trypsin digest of BCb
11	KGTEF	100–104	trypsin digest of BCb
12	GTEFTVNDLQ GK	101–112	trypsin digest of BCb
13	LGRSAG	119–124	thermolysin digest of BCb
14	SAGWNI	122–127	sequencer run on thrombin fragment
15	IESG	143–146	pepsin digest of BCb
16	AVAKF	151–155	pepsin digest of BCb
17	CARNAPYSGY	182–191	pepsin digest of BCb
18	YSGAF	191–195	pepsin digest of BCb
19	HCLKDGKGDVAF	196–207	pepsin digest of BCb
20	GDVAFVK	203–209	trypsin digest of BCb
21	TCNWAR	241–246	trypsin digest of BCb
22	ARVA	245–248	pepsin digest of BCb
23	VAAH(AVVAR)	247–255	trypsin digest of BCb
24	AVVAR	251–255	trypsin digest of BCb
25	ARDD(NKVED)	254–262	pepsin digest of BCb
26	DIW	262–264	subtilisin digest of BCb
27	AQSDFGVDTK	270–279	trypsin digest of BCb
28	SDFHLFGPPGK	280–290	trypsin digest of BCb
29	GPPGKKN	286–292	chymotrypsin digest of BCb
30	NPVLK	292–296	trypsin digest of BCb
31	DLLFK	297–301	trypsin digest of BCb
32	DSAIM	302–306	trypsin digest of BCb (C-terminal of BCb)
33	DSAIMLK	302–308	trypsin digest of OT
34	LKRVP SLM	307–314	CNBr digest of ovotransferrin (OT)
35	YLGF	319–322	pepsin digest of OT
36	LGFEYY	320–325	trypsin digest of OT
37	MR	331–332	CNBr digest of NT. R351 terminal
38	KGEADA(VALDGGLVY)	386–400	trypsin digest of D/E
39	DAVAL	399–394	pepsin digest of D/E
40	TDERPASYF	424–432	trypsin digest of BCE
41	FAVAVAR	432–438	trypsin digest of BCE; also obtained with terminal K
42	DSNVNW	440–445	trypsin digest of OT
43	ANVMDYR	576–582	trypsin digest of OT
44	AEVPTHA	587–593	pepsin digest of OT
45	DLLER	606–610	trypsin digest of D/E
45a	KDLL	605–608	trypsin digest of D/E
46*	FGVNGSQK	615–622	trypsin digest of D/E
47	MM	626–627	CNBr cleavage → HSL
48	FE(SQNK)DLLFK	628–638	chymotrypsin digest of BCE
49	KVR(EGTTY)	646–653	chymotryptic digest of BCE

Table 2 (Continued)

No.	Structure	Position	Source of peptide
50	KEFLGDKFY	654–662	chymotrypsin digest of BCE
51	TVI(SSL)K	663–669	trypsin digest of BCE

* In peptide 46 Edman degradation failed to cleave the asparaginyl-glycine bond, the sequence was established by mass spectrometry.

Table 3. *Variable sites in ovotransferrin*

Position	DNA/AA	Peptide	Comments
33	serine	serine or alanine	[10]
292	aspartic acid	asparagine	peptides 29 and 30
438	arginine	arginine or lysine	peptide 41
512	lysine	lysine or arginine	[10,11]
519	glutamic acid	glutamic acid or glutamine	glutamic acid in [10] and [11] but an unpubl. peptic peptide HQKY
540	isoleucine	isoleucine or valine	[11]
605	arginine	lysine	peptide 45a
621	glutamic acid	glutamine	peptide 46
667	asparagine	serine	peptide 51

INTERNAL DUPLICATION OF STRUCTURE

A clear visual demonstration of the sequence homology of the N-terminal and C-terminal halves of the ovotransferrin chain is given (Fig. 2) by the diagonal plot method of Gibbs and McIntyre [26]. Fig. 3 shows an alignment of residues 1–332 and 342–686, these sections corresponding to the N-terminal and C-terminal proteolytic fragments. Gaps have been inserted to give as many identical matches as possible. Out of the total 350 positions thus created there are 131 identical residues (37.4%), each gap position being counted as one difference. The diagonal plot gives no strong visual evidence of the quadruplication suggested earlier [1] nor were attempts to align quarter-molecules convincing. However, Jeltsch and Chambon [12] have discussed some peptide similarities which may reflect an ancient quadruplication.

The conservation of residues in matching positions in the two domains probably indicates that they have common structural and functional roles. As Fig. 3 shows, the conserved residues are distributed in an irregular manner and are not collected in any particular region of the chain. The most highly conserved type of amino acid is half-cystine since 86% (24/28) of its occurrences are in conserved positions. The next most highly conserved amino acids are histidine (8/12), glycine (30/49), tyrosine (12/21) and leucine (26/47).

THE ARRANGEMENT OF DISULPHIDE BRIDGES

The structural role of the half-cystine residues is well shown in the arrangement of disulphide bridges in the two halves of ovotransferrin. By sequencing cystine-containing peptides from pepsin digests Elleman and Williams [10] identified 12 of the 15 bridges. The half-cystine residues linked by these bridges are as follows: 10 to 45, 115 to 197, 160 to 174, 171 to 182, 228 to 242, 348 to 380, 405 to 680, 421 to 643, 454

to 530, 478 to 671, 488 to 499 or 502 and 513 to 502 or 499. We can resolve the ambiguity in the last two bridges given in the list and also give probable identifications of the remaining three bridges by using the evident homology between the domains and the fact that no bridges pass from one domain to the other. Thus, we propose the following connections 488

to 502, 513 to 499, 20 to 36, 358 to 371 and 570 to 584. This arrangement is shown in Fig. 4.

It is clear that there are six homologous bridges in the two domains and an additional three bridges which are present only in the C-terminal domain. The extra bridges are of much longer range than the homologous bridges, since they connect parts of the chain separated by 193–275 residues.

In the C-terminal domain the disulphide bridges numbered 3, 4, 5, 7, 8 and 9 (Fig. 4) probably impose severe constraints on the conformation of the polypeptide chain, as a 'string model' shows (Fig. 5). The central region of 125 residues from positions 405 to 530 forms an intricately cross-linked structure, which probably possesses considerable stability. The outlying loops determined by bridges 1, 2 and 6 may be packed around this central region. Although not possessing as many cross-links it is likely that the overall conformation of the N-terminal half of the chain is similar to that of the C-terminal half, as is suggested in Fig. 5. The results of X-ray analysis will be of considerable interest in confirming or disproving these ideas.

COMPARISON OF HEN AND HUMAN TRANSFERRINS

Partial sequence data on human transferrin have been obtained by Brew and his colleagues [1,2]. Although these data are incomplete it is of great interest to compare them with the new picture of hen transferrin.

21 half-cystine residues in human transferrin are clearly homologous with half-cystines in the hen sequence (Fig. 6) and it is likely that the still undetermined regions of the human protein contain further half-cystines. The disulphide bridges of human transferrin have apparently not been characterized yet but by assuming homology with hen transferrin the following conclusions seem likely: the N-terminal domain possesses bridges numbered 1, 2, 3 and 6 (the region

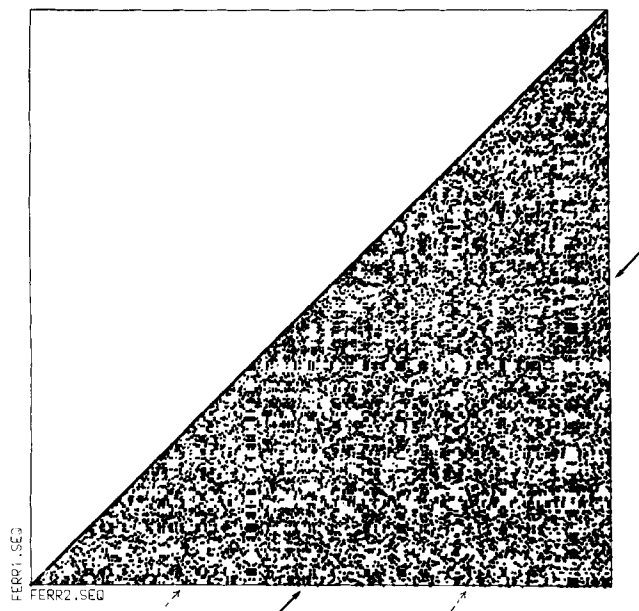


Fig. 2. Diagonal plot of the ovotransferrin amino acid sequence against itself [26]. The similarity of the N-terminal and C-terminal halves is shown by the strong diagonal clustering of points (arrows). If quadruplication had occurred it would give two more diagonal clusters in the positions shown by the dashed arrows



Fig. 3. An alignment of the sequences of the N-terminal and C-terminal halves of hen ovotransferrin. Gaps have been introduced in order to maximise the matching identical residues. Identical residues are enclosed in boxes. The connecting peptide KDQLTPSPR is not shown. 0 marks those tyrosyl, histidyl and arginyl residues which are conserved in the N-terminal and C-terminal domains of both human and hen transferrins

Table 4. Percentages of identical residues in comparisons of N-terminal and C-terminal halves of hen and human transferrins

Transferrin	Hen C	Human N	Human C
Hen N	37	53	41
Hen C	—	45	50
Human N	—	—	40

Those residues which are critically involved in iron binding would be expected to be constantly present in all N-terminal and C-terminal domains. It is generally thought that each iron-binding site contains three tyrosyl residues and two histidyl residues [28]. The concomitantly bound anion may be held electrostatically to an arginyl side group. There are 6 conserved pairs of tyrosine residues in hen transferrin (Fig. 3); five of them are also conserved in human transferrin. These are residues 82, 92, 188, 191 and 239 in the N-terminal domain and residues 415, 431, 521, 524 and 581 in the C-terminal domain. Similarly conserved histidyl residues are 116, 210 and 250 in the N-terminal domain and 455, 542 and 592 in the C-terminal domain. Conserved arginyl residues are 121, 233 and 255 in the N-terminal domain and 460, 574 and 597 in the C-terminal domain. The positions of these residues are indicated in Fig. 3. However, when more sequences are known there may be fewer constant residues. Chemical modification experiments may enable the residues involved in iron binding to be identified more precisely. The fact that in both hen and human transferrins there are more disulphide bridges in the C-terminal domain than in the N-terminal domain may explain the resolution of the N-terminal and C-terminal monoferric protein complexes upon polyacrylamide gel electrophoresis in the presence of 6 M urea [29]. In the absence of bound metal 6 M urea causes partial unfolding of the protein and a consequent reduction in electrophoretic mobility. The unfolding of the N-terminal domain causes a greater decrease in mobility than does the unfolding of the C-terminal domain [30] and this may be due to the greater number of disulphide bridges in the latter.

Clearly an early event in transferrin evolution was the gene duplication which produced the two-domain structure. Presumably this was necessary to increase the molecular weight

and avoid loss of the protein from the circulation by diffusion through the glomerular membrane [31].

Since the available sequence data are very limited, only tentative suggestions can be made about the subsequent evolution of the protein. Nevertheless, from the pattern of disulphide bridges and the percentage of identical residues between N-terminal and C-terminal domains it is likely that the structural differentiation of the initially identical domains had already occurred in the reptilian ancestors of birds and mammals. It would be of great interest to know whether this differentiation could be traced further back and the sequences of reptile, amphibian and fish transferrins may throw light on this. These structural differences are associated with differences in the iron-binding properties of the two domains, so that the affinity of the two domains for iron is different [32] and they are clearly distinguished by their response to changes in pH and salt concentration [33, 34]. Whether this structural differentiation of the domains is accompanied by physiological functional differences is an open question. Studies on the ability of the diferric protein to donate iron to primitive red blood cells show no differences between the two sites [35] but other possible physiological functions of transferrin have received little study. The still incomplete data on human transferrin suggests that a further differentiation of the domains has occurred in mammals as shown by the acquisition of extra non-homologous disulphide bridges in both domains. In order to obtain a fuller picture of transferrin evolution it is to be hoped that the amino acid sequences and disulphide bridge patterns of a wide range of transferrins and lactoferrins will be studied.

We are grateful to Dr Vincenzo Starace (Unit of Experimental Medicine, UCL-ICP 7430, Avenue Hippocrate 75, B-1200 Brussels, Belgium) and to Dr Chris Brock of this laboratory for carrying out the automatic sequencer runs. We thank James Warwicker for making the diagonal plot of transferrin.

REFERENCES

- MacGillivray, R. T. A., Mendez, E. & Brew, K. (1977) in *Proteins of Iron Metabolism* (Brown, E. B., Aisen, P., Fielding, J. & Crichton, R. R., eds) pp. 133–141, Grune & Stratton, New York.
- Lineback-Zins, J. & Brew, K. (1980) *J. Biol. Chem.* **255**, 708.
- Gorinsky, B., Horsburgh, C., Lindley, P. F., Moss, D. S., Parkar, M. & Watson, J. L. (1978) *Nature (Lond.)* **281**, 157–158.
- Williams, J. (1974) *Biochem. J.* **141**, 745–752.
- Williams, J. (1975) *Biochem. J.* **149**, 237–244.
- Metz-Boutigue, M. H., Jolles, J., Mazurier, J., Spik, G., Montreuil, J. & Jolles, P. (1978) *Biochimie (Paris)* **60**, 557–561.
- Metz-Boutigue, M. H., Jolles, J. P., Mazurier, J. J., Spik, G. & Montreuil, J. (1980) *Biochim. Biophys. Acta*, **622**, 308–314.
- Williams, J. (1962) *Biochem. J.* **83**, 355–364.
- Thibodeau, S. N., Lee, D. C. & Palmiter, R. D. (1978) *J. Biol. Chem.* **253**, 3771–3774.
- Elleman, T. C. & Williams, J. (1970) *Biochem. J.* **116**, 515–535.
- Kingston, I. B. & Williams, J. (1975) *Biochem. J.* **147**, 463–472.
- Jeltsch, J. M. & Chambon, P. (1982) *Eur. J. Biochem.* **122**, 291–295.
- Spik, G., Fournet, B. & Montreuil, J. (1979) *C. R. Hebd. Seances Acad. Sci. Ser. D. Sci. Nat.* **288**, 967–970.
- Warner, R. C. & Weber, I. (1951) *J. Biol. Chem.* **191**, 173–180.
- Azari, P. R. & Baugh, R. F. (1967) *Arch. Biochem. Biophys.* **118**, 138–144.
- Tsao, D., Azari, P. R. & Phillips, J. L. (1974) *Biochemistry*, **13**, 400–413.
- Bezkorovainy, A., Gerbeck, C. M. & Grohlich, D. (1968) *Biochem. J.* **110**, 765–770.
- Graham, I. & Williams, J. (1975) *Biochem. J.* **145**, 263–279.
- Williams, J. (1968) *Biochem. J.* **108**, 57–67.
- Wünsch, E. (1963) *5th Eur. Peptide Symp. (1962)* pp. 89–91, Pergamon, Oxford.
- Smyth, D. G., Stein, W. H. & Moore, S. (1963) *J. Biol. Chem.* **238**, 227–234.
- Schroeder, W. A., Jones, R. I., Cormick, J. & McCalla, K. (1962) *Anal. Chem.* **34**, 1570–1575.
- Gray, W. E. (1967) *Methods Enzymol.* **11**, 139–151.
- Gray, W. R. (1967) *Methods Enzymol.* **11**, 469–475.
- Ogden, A. L., Morton, J. R., Gilmour, D. G. & McDermid, E. M. (1962) *Nature (Lond.)* **195**, 1026–1028.
- Baker, C. M. A. (1967) *Comp. Biochem. Physiol.* **20**, 949–973.
- Stratil, A. (1967) *Comp. Biochem. Physiol.* **22**, 227–233.
- Gibbs, A. J. & McIntyre, G. A. (1970) *Eur. J. Biochem.* **16**, 1–11.
- Bezkorovainy, A. (1980) *Biochemistry of Nonheme Iron*, Plenum Press, New York, London.
- Chasteen, N. D. (1977) *Coord. Chem. Rev.* **22**, 1–36.
- Makey, D. G. & Seal, U. S. (1976) *Biochim. Biophys. Acta*, **453**, 250–256.
- Evans, R. W. & Williams, J. (1980) *Biochem. J.* **189**, 541–546.
- Schultze, H. E. & Heremans, J. F. (1966) *Molecular Biology of Human Proteins*, vol. 1, p. 706, Elsevier Publishing Co., Amsterdam.
- Aisen, P. A., Leibman, A. & Zweier, J. (1978) *J. Biol. Chem.* **253**, 1930–1937.
- Chasteen, N. D. & Williams, J. (1981) *Biochem. J.* **193**, 717–727.
- Williams, J., Chasteen, N. D. & Moreton, K. (1981) *Biochem. J.* in the press.
- Harris, D. C. (1977) *Biochim. Biophys. Acta*, **496**, 563–565.

J. Williams, Department of Biochemistry, University of Bristol, University Walk, Bristol, Avon, Great Britain, BS8 1TD

T. C. Elleman, Division of Protein Chemistry, Commonwealth Scientific and Industrial Research Organization, 343 Royal Parade, Parkville, Victoria, Australia 3052

I. B. Kingston, Laboratory of Molecular Biology of the Medical Research Council, Postgraduate Medical School, University of Cambridge, Great Britain, CB2 2QH

K. A. Kuhn, Home Office Forensic Laboratory, Central Research Establishment, Aldermaston, Reading, Berkshire, Great Britain, RG7 4PN

A. G. Wilkins, Systime Ltd, Concourse Computer Centre, 432, Dewsbury Road, Leeds, Great Britain LS11 7LJ