10. H. L. Shands and G. Kimber, *Proc. 4th Int. Wheat Genet. Symp.* (1973), p. 101.
11. P. Sarkar and G. L. Stebbins, *Am. J. Bot.* **43**, 297 (1956).
12. R. Riley, J. Unrau, V. Chapman, *J. Hered.* **49**, 90 (1958).
13. G. Kimber and R. S. Athwal, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 912 (1972).
14. B. L. Johnson, *ibid.*, p. 1398.
15. ———, *Can. J. Genet. Cytol.* **17**, 21 (1975).
16. K. Chen, S. D. Kung, J. C. Gray, S. G. Wildman, *Biochem. Genet.*, in press.
17. H. Kihara, *Proc. 3rd Int. Wheat Genet. Symp.* (1968), p. 125.
18. Supported by NSF grant 75-07368 and AEC contract AT(04-3)-34, project 8. We are grateful to Drs. B. L. Johnson and H. S. Dhaliwal of the Department of Plant Sciences, University of California, Riverside, for helpful discussions and for supplying the plant material.
* Present address: Department of Biochemistry, University of Cambridge, Cambridge, England.

20 August 1975

Transferrin: Internal Homology in the Amino Acid Sequence

Abstract. *Two regions of the primary structure of human serum transferrin, of 87 and 57 residues, are reported. When these are suitably aligned by placing two gaps, 40 percent of the amino acids in corresponding positions are identical. This indicates that the doubling of an ancestral structural gene occurred during the evolution of the transferrins.*

Serum transferrin, which functions in the transport of ferric iron in the circulation, consists of a single glycopolypeptide chain of molecular weight 77,000 (*1*), within which are two similar binding sites for Fe^{3+} and other metals (*2*). Similar proteins are found in avian egg whites (ovotransferrins) and in milk and other secretions (lactoferrins), which probably have an antimicrobial function. The similar nature of the metal binding sites, both of which require concomitant binding of an anion with the metal (*3*), and the large size of the polypeptide chain have been taken by some to suggest that the protein may have originated during phylogeny by the doubling in size of the structural gene for a smaller precursor (*1, 4*). This history would be expected to be evident in the amino acid sequence of a modern transferrin as regions of internal homology, but previous structural studies of human transferrin (*5*) and human lactoferrin (*6*) have not provided any convincing evidence for internal homology. In the sequences around 34 unique cystinyl residues in ovotransferrin, Elleman and Williams (*7*) found signs of similarity only in two peptides of nine residues each, while Graham and Williams found no internal homology between sequences around the two glycosylation sites in human and other transferrins (*8*).

Studies by Williams have, however, shown that fragments of approximately half the molecular weight of the intact protein can be obtained from ovotransferrin by limited proteolysis (*9*). These fragments appear to correspond to the amino and carboxy terminal halves of the polypeptide chain, and each contains an intact Fe^{3+} binding site (*9*). This observation is consistent with the presence of two largely independent conformational domains within the transferrin molecule, each associated with one of the metal binding sites. A structure of this type might well be expected to be associated with internal homology.

In our previous structural studies with human serum transferrin, we have isolated and characterized the nine fragments obtained by CNBr cleavage at the eight methionyl residues (*10*). The sizes of these fragments (designated CN-1 to CN-9) account for the whole polypeptide chain of the protein, the two oligosaccharide prosthetic groups being present in CN-2 (molecular weight, 15,000) and CN-3 (molecular weight, 10,000) (*5*). The complete sequences of four fragments have been published: CN-6 (the amino terminal 26 residues of the protein), CN-7, CN-8, and CN-9 (*10, 11*). Partial sequences for CN-3, CN-4, and CN-5 have also been reported (*5*).

We report here the sequence of a region in CN-1, the fragment having a molecular weight of 25,000 from the carboxy terminus of the molecule (*5*) and the structure of a peptide which provides evidence for the ordering of CN-7, CN-9, and CN-4 in the polypeptide chain. The two extensive regions of sequence thus assembled provide convincing evidence for internal homology in the polypeptide chain of transferrin.

The aminoethyl (AE) and carboxamidomethyl (CAM) derivatives of CNBr fragments were prepared and isolated as

(1) Leu-Lys-Val-Pro-Pro-Arg-Metsulphone-Asx

CN-7 — CN-9 — CN-4

(2) Leu-Lys-Val-Pro-Pro-Arg-Metsulphone-Asx-Ala-Lys-Metsulphone-Tyr

Fig. 1. Sequences of peptides denoting the ordering of CN-7, CN-9, and CN-4. Abbreviations: Leu, leucine; Lys, lysine; Val, valine; Pro, proline; Arg, arginine; Asx, aspartic acid or asparagine; Ala, alanine; and Tyr, tyrosine.

```
                1                                  10                                  20
Line A  LEU-PHE-Ser-SER-Pro-His-Gly-LYS-Asn-LEU-LEU-PHE-Lys-ASP-Ser-Ala-His-Gly-Phe-Leu-

Line B  LEU-PHE-Arg-SER-Glu-Thr-▬-LYS-Asp-LEU-LEU-PHE-Arg-ASP-Asp-Thr-Val-Cys-Leu-Ala-

                                   30                                  40
Line A  LYS-Val-Pro-Pro-ARG-Met-Asn-Ala-Lys-Met-TYR-LEU-GLY-Tyr-GLU-TYR-VAL-Thr-ALA-Ile-

Line B  LYS-Leu-His-Asp-ARG-Asn-Thr-Tyr-Glu-Lys-TYR-LEU-GLY-Glu-GLU-TYR-VAL-Lys-ALA-Val-

                                   50                                  60
Line A  Arg-ASN-LEU-ARG-Glu-▬▬▬▬▬-Gly-Thr-Cys-Pro-GLU-ALA-Pro-THR-Asn-Glu-Cys-Lys-

Line B  Gly-ASN-LEU-ARG-Lys-Cys-Ser-Thr-Ser-Ser-Leu-Leu-GLU-ALA-Cys-THR-Phe-Arg-
```

Fig. 2. A comparison of the amino acid sequences of two regions from the polypeptide chain of human serum transferrin. Line A consists of the combined sequences of CN-7, CN-9, and CN-4; line B consists of a region from CN-1. The amino terminal region of CN-7 and the carboxy terminal region of CN-4 are not included. Identical residues in corresponding positions are denoted with capital letters, and chemically similar residues (conservative replacements) are underlined. The numbering refers to the residues within the alignment, and not to positions in the polypeptide chain of transferrin.

described for the CAM peptides (*10*), except that AE CN-3 and AE CN-4, unlike the CAM derivatives, were not resolved by gel filtration with Sephadex G-50, but were purified by ion-exchange chromatography with carboxymethyl-cellulose at 40°C. The purified fragments gave compositions in close agreement with those expected for their sequences. The amino terminal sequence of CN-4 was constructed from the sequences of overlapping peptides isolated from thermolysin and trypsin digests, and was confirmed by an automated sequencer analysis of CAM CN-4. The sequence from CN-1 is based largely on the structures of arginine cleavage fragments obtained by trypsin digestion of the citraconyl derivative of AE CN-1 and purified by ion-exchange chromatography and gel filtration. Peptides isolated from trypsin, chymotrypsin, or thermolysin digests of CAM CN-1 were also used in constructing the sequence. All peptides were sequenced by the combined dansyl chloride–Edman degradation procedure (*12*) or by the subtractive Edman degradation procedure. Amides were determined by amino acid analysis of Pronase aminopeptidase digests of appropriate peptides (*10*).

The peptides denoting the ordering of CN-7, CN-9, and CN-4 were isolated from a thermolysin digest of performic acid–oxidized transferrin by ion-exchange chromatography and gel filtration. Their sequences are shown in Fig. 1. Although only one residue of CN-4 is contained in this overlap (tyrosine), there is no ambiguity as CN-4 is unique in possessing amino terminal tyrosine. The two regions of sequence are shown in Fig. 2, having been aligned by the insertion of one gap of three residues in the CN-7,9,4 sequence and a gap of one residue in the CN-1 sequence. Of the 57 positions being compared, 23 residues in corresponding positions in the two sequences are identical. Many other amino acids in corresponding positions are chemically similar. The extent of homology based on identical residues (40 percent) is comparable to that seen in comparisons of pairs of proteins whose common ancestry is well established, for example, α-lactalbumin and lysozyme (*12*). As the two regions of sequence are well separated in the molecule, the region from CN-1 being at least 22 residues removed from the amino terminus of this fragment, this sequence alignment indicates the existence of internal homology in the polypeptide chain of transferrin, reflecting an ancestral gene doubling during the phylogenetic development of vertebrate transferrins.

We conclude from this study and the work of Williams (*9*) that transferrin is composed of two homologous regions, each having a single metal binding site and a degree of conformational independence, including an absence of disulfide cross-links between the domains (implicit in the limited proteolysis studies).

The necessity for conserving the metal binding sites and other functional areas in the two homologous regions of the molecule would be expected to restrict the sequence variation around such areas. These areas should be apparent in the internal homology as regions of strongly conserved sequence. The highly conserved region in the above sequence alignment (positions 31 to 44) is therefore a possible candidate for an area of functional significance. It is interesting that each metal binding site in transferrin contains two tyrosyl residues (*13*), which chelate with the metal, and a positively charged group, associated with anion binding. Such residues (two tyrosines and an arginine) are present in the conserved area in Fig. 2. When our current sequence studies of transferrin are completed, we hope to clarify the functional significance of this and other areas in the molecule by chemical modification.

R. T. A. MacGillivray
K. Brew

Department of Biochemistry, School of Medicine, University of Miami, Miami, Florida 33152

References and Notes

1. F. C. Greene and R. E. Feeney, *Biochemistry* **7**, 1366 (1968); K. G. Mann, W. W. Fish, A. C. Cox, C. Tanford, *ibid.* **9**, 1348 (1970); R. M. Palmour and H. E. Sutton *ibid.* **10**, 4026 (1971).
2. A. L. Schade, R. W. Reinhart, H. Levy, *Arch. Biochem.* **20**, 170 (1949); D. D. Ulmer and B. L. Vallee, *Biochemistry* **2**, 1335 (1963); R. C. Woodworth, K. G. Morallee, R. J. P. Williams, *ibid.* **9**, 839 (1972).
3. R. C. Woodworth *et al.*, in *Proteins of Iron Storage and Transport in Biochemistry and Medicine*, R. R. Crichton, Ed. (North-Holland, Amsterdam, 1975), p. 39.
4. R. E. Feeney and R. G. Allison, in *Evolutionary Biochemistry of Proteins* (Interscience, New York, 1969), p. 151.
5. M. R. Sutton, R. T. A. MacGillivray, K. Brew, in *Proteins of Iron Storage and Transport in Biochemistry and Medicine*, R. R. Crichton, Ed. (North-Holland, Amsterdam, 1975), p. 19; J. Jollès, P. Charet, P. Jollès, J. Montreuil, *FEBS Lett.* **46**, 276 (1974).
6. P. Montreuil and G. Spik, in *Proteins of Iron Storage and Transport in Biochemistry and Medicine*, R. R. Crichton, Ed. (North-Holland, Amsterdam, 1975), p. 27.
7. T. C. Elleman and J. Williams, *Biochem. J.* **116**, 515 (1970).
8. I. Graham and J. Williams, *ibid.* **145**, 263 (1975).
9. J. Williams, *ibid.* **141**, 745 (1974); *ibid.* **149**, 237 (1975).
10. M. R. Sutton and K. Brew, *ibid.* **139**, 163 (1974). M. R. Sutton, R. T. A. MacGillivray, K. Brew, *Eur. J. Biochem.* **51**, 43 (1975).
11. M. R. Sutton and K. Brew, *FEBS Lett.* **40**, 146 (1974).
12. J. B. C. Findlay and K. Brew, *Eur. J. Biochem.* **27**, 65 (1972).
13. J. L. Phillips and P. Azari, *Arch. Biochem. Biophys.* **151**, 445 (1972).
14. Supported by NIH research grant GM21363, and an NIH research development award to K.B. We thank Drs. M. R. Sutton and R. A. Bradshaw of the Washington University School of Medicine for the sequencer analysis of CN-4.

12 September 1975; revised 14 October 1975

Estrogen Target Sites in the Brain of the Chick Embryo

Abstract. *Autoradiograms prepared from brains of chick embryos after injection of [^{3}H]estradiol demonstrate the existence of target cells for estrogen in the medial preoptic and ventral hypothalamic regions as early as day 10 of incubation. Target cells also appear in telencephalic, diencephalic, and mesencephalic locations during later stages of embryonic development. These hormone-concentrating cells probably are the anatomical substrate for the formative action of sex steroids during embryonic life on certain brain functions, such as control of sexual and aggressive behavior and gonadotropin secretion.*

The role of gonadal steroids in influencing the development of the brain and determining differences between males and females in behavior and hormonal regulation has been examined in numerous species, such as the rat, mouse, guinea pig, dog, rhesus monkey, chicken, and Japanese quail (*1–5*). It has been suggested that early gonadal secretions also affect the development of the brain in man (*6*). In some species, such as the rat, the critical period for sexual differentiation of the brain occurs perinatally or in the early postnatal period (*4*). In many species, the effect of gonadal steroids on brain development is initiated prenatally or restricted to a period during embryonic life, as is likely to be the case in primates (*1, 7*). Efforts to demonstrate steroid hormone "receptors" in the neonatal brain with biochemical techniques have been largely unsuccessful (*8*), and hitherto no report exists in the literature demonstrating steroid hormone receptors in embryonic nervous tissue. In contrast, the presence of target cells for estrogen and androgen in the brain of the 2-day-old neonatal rat has been demonstrated with autoradiography (*9*). In our study, we used autoradiography to investigate the ontogeny of estrogen target cells. The chick embryo was selected, since the presence of maternal and placental steroids would be expected to impair demonstration of binding sites in placental animals.

Chick embryos of the White Rock (Arbor Acre) strain were incubated at 38°C in 60 to 90 percent relative humidity in a still air incubator. Prior to incubation the eggs had been stored at 18°C. Nine chick embryos at different stages of development and two newly hatched chicks received 0.2 μg of [2,4,6,7-^{3}H]17β-estradiol (specific activity, 105 mc/mmole) in a mixture of 10 percent ethanol and saline. One additional