The amino acid sequence of human liver apoferritin

J.M. Addison, J.E. Fitton, W.G. Lewis[+], K. May* and P.M. Harrison

The Department of Biochemistry, The University, Sheffield S10 2TN, England

Received 4 October 1983

The protein component of the iron storage molecule, ferritin, contains 24 subunits in form of a hollow shell known as apoferritin. The amino acid sequence has been determined for apoferritin subunits from human liver. The sequence comprises 174 amino acids giving an M_r of 19 900. It shows extensive homology with the primary structures of apoferritins from human and horse spleen and from rat liver. Sequence substitutions are discussed in relation to the known three-dimensional structure of horse spleen apoferritin. Evidence for a second minor sequence in human liver apoferritin is presented.

Human liver apoferritin *Amino acid sequence*

1. INTRODUCTION

Iron is an essential trace element in almost all living organisms. It is a constituent of many enzymes (catalases, peroxidases, nitrogenases, hydrogenases), electron transfer proteins (cytochromes) and oxygen carrier proteins (haemoglobin, myoglobin). However, the solution chemistry of iron has necessitated the development of specialised storage and transport systems. In aqueous media, at physiological pH, ferrous iron has a tendency to oxidise and hydrolyse forming polymeric hydrous ferric oxide or 'ferrihydrite' [1]. Intracellular iron storage therefore requires a specialised protein and this role is performed by ferritin. It consists of an iron core of ferrihydrite deposited in the cavity of a hollow protein shell (apoferritin), maintaining the iron in a stable, non-toxic and readily available store from which it can be mobilised when required.

The protein shell consists of 24 structurally equivalent subunits [2] of two species differing slightly in size or apparent size as determined by SDS–polyacrylamide gel electrophoresis; the major or light subunit (L) with an M_r of 19 000 and the minor or heavy subunit (H) of M_r 21 000 [3]. The relative proportions of the two subunits appear to be tissue specific [3]. The three-dimensional structure of horse spleen apoferritin has been determined and refined to 2.8 Å; each subunit consisting of a parallel bundle of 4 helices A–D with a smaller fifth helix perpendicular to them and an extended loop region connecting helices B and C [4].

Studies on the primary structure of ferritins are of interest for several reasons: they assist in the identification of structural and functional regions in the three-dimensional structure; form a comparison with other ferritin sequences; and may provide a molecular explanation for the heterogeneity of ferritin, i.e. the difference between H and L subunits.

We present below the complete amino acid sequence of the major subunit of human liver apoferritin and its comparison with the primary sequences of horse spleen [5], human spleen [6] and rat liver apoferritin [7]. We also report partial sequence data for a second minor sequence.

[+] Present address: Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, England

* Present address: Unilever Research Colworth Laboratory, Colworth House, Sharnbrook, Bedford MK44 1LQ, England

Published by Elsevier Science Publishers B.V.

2. MATERIALS AND METHODS

Ferritin from human liver was prepared by a modification of the procedure in [8] involving heat treatment, acidification to pH 4.8, ammonium sulphate precipitation and high speed centrifugation at 100 000 × *g* for 2 h. The purity of the preparation was monitored by SDS–polyacrylamide gel electrophoresis and amino acid analysis. Samples were hydrolysed for 24 h in vacuo with 6 N HCl (Aristar grade) at 105°C and analysis was performed on a Locarte single column amino acid analyser (Locarte, London).

Apoferritin was prepared by dialysis against 0.1 M thioglycollic acid (Sigma, Poole) in 0.1 M sodium acetate solution at pH 4.25 and 4°C. Apoferritin was reduced [9] and carboxymethylated [10] in 6 M guanidine hydrochloride prior to cleavage with trypsin, staphylococcal protease or *Armillaria mellea* protease [11]. Trypsin and staphylococcal protease were also used to cleave carboxymethylated apoferritin, after citraconylation of its lysine residues [12] in 6 M guanidine hydrochloride.

A large *A. Mellea* protease peptide (AMI) of the N-terminal region of the sequence was further digested with chymotrypsin and staphylococcal protease. A large peptide TcIIIf from the citraconylated tryptic digest was digested with trypsin after deblocking of the lysine residues by incubation with 10% (v/v) acetic acid for 3 h at 45°C.

Digests were initially separated by gel filtration on Sephadex G-50SF and the smaller peptides obtained were purified by high voltage paper electrophoresis at pH 2.0 and 6.5. Larger peptides were subjected to further gel filtration. The tryptic digest was initially fractionated by ion exchange chromatography on Dowex AG 50Wx2, followed by high voltage paper electrophoresis.

Acid cleavage of carboxymethylated apoferritin in 70% (v/v) formic acid for 4 d at 37°C produced a single cleavage at the only Asp–Pro bond in the sequence [13] as previously seen in the horse spleen ferritin sequence [5]. Since ferritin is blocked [14] at its N-terminus, the free α-NH_2 proline peptide present could be sequenced directly by automatic liquid phase sequencing on a Beckman 890C spinning cup sequencer. One other large staphylococcal protease peptide SAcI, from the N-terminal region, was sequenced by liquid phase sequence analysis. The thiazolinones were converted to PTH amino acids and were identified by high pressure liquid chromatography.

Smaller peptides were sequenced by the manual dansyl-Edman method [15], identifying the dansyl amino acids by thin layer chromatography on polyamide sheets [16].

Purity of peptides was confirmed by dansyl N-terminal analysis [17] and amino acid analysis.

Chemicals used in sequencing, sequencer or HPLC grade, were obtained from Pierce & Warriner (Chester) or Rathburn (Scotland). All other chemicals were AnalaR grade. Trypsin and chymotrypsin were from Miles (Slough).

3. RESULTS

The primary sequence of human liver apoferritin has been determined from peptides obtained from digests of the protein by trypsin, staphylococcal protease, *A. mellea* protease and from an acid cleavage. The sequence is presented in fig.1 with the peptides used to establish the sequence displayed. Where no overlapping peptides were obtained, the homology, between the sequences of horse spleen and human spleen apoferritin, was utilised to place peptides in the sequence.

A ninhydrin-negative, arginine-positive peptide TcIVNi, sequenced by mass spectrometry was identical to the N-terminal sequences of horse and human spleen apoferritins.

Peptides obtained from a chymotryptic subdigest of the *A. mellea* protease peptide AMI (residues 1–66 by amino acid analysis) overlapped with the N-terminal peptide and allowed continuation of the sequence up to Asn 17.

The large staphylococcal protease peptide SAcI, sequenced by liquid phase sequence analysis, overlapped with AMI CH-bi and confirmed the sequence up to Phe 35. Residues 36–39 have not been identified with certainty, but the amino acid analysis of the peptide is in good agreement with the residues proposed.

Small tryptic peptides established the sequence from Asx 40 to Lys 97. Overlapping peptides were obtained for the region 47–66 from the staphylococcal protease subdigest of peptide AMI. Peptide TcIV(2), from the citraconylated tryptic digest, confirms the sequence between residues 65–72. The remaining peptides in this region were placed

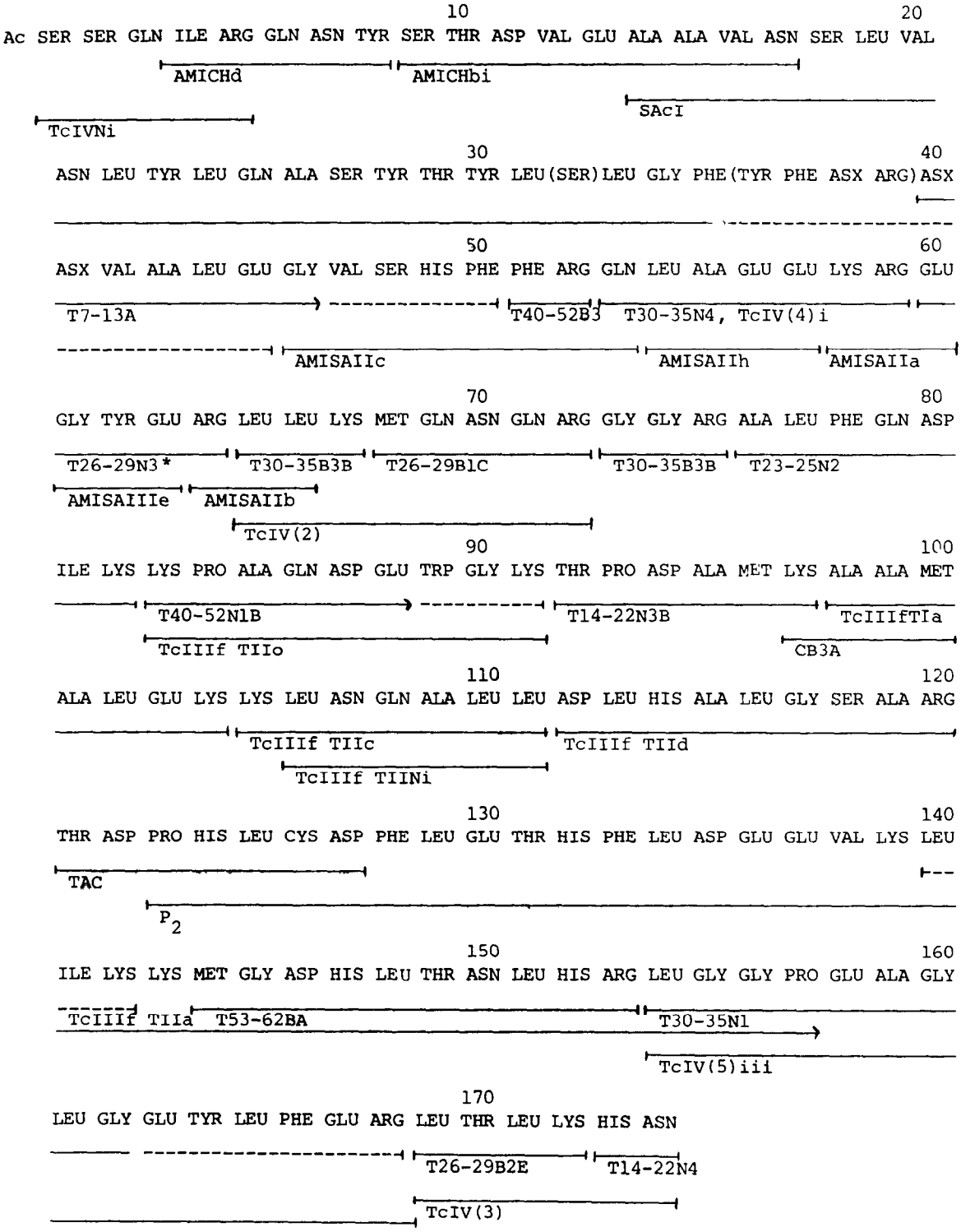

Fig.1. Amino acid sequence of human liver apoferritin. Key to peptide naming: T, tryptic digest; Tc, citraconylated tryptic digest; AM, *A. mellea* protease digest; CH, chymotrypsin subdigest; SA, staphylococcal protease digest; CB, cyanogen bromide cleavage; P_2, acid cleavage product; (——) sequence and amino acid analysis information; (------) amino acid analysis only.

by homology with the human and horse spleen sequences.

Tryptic peptides extracted from a large citraconylated tryptic fragment TcIIIf provided additional sequence information for residues 100–120, with some overlaps being provided by a cyanogen bromide peptide CB3A and tryptic peptide TcIIIf·TIINi.

The sequence of the acid cleavage peptide P_2, determined by liquid phase sequence analysis, established the sequence from residues 123–157. This region is confirmed by several tryptic peptides.

The C-terminal sequence is provided by TcIV(5)-iii, from the citraconylated tryptic digest, from Leu 154 to Arg 168 and also overlaps with P_2. Two small tryptic peptides T 26-29 B2E and T 14-22 N4 complete the sequence and are confirmed by peptide TcIV(3) from the citraconylated tryptic digest.

4. DISCUSSION

The M_r of the apoferritin subunit, calculated from the sequence, is 19900, marginally lower than the value of 20017 for human spleen and slightly greater than that of 19824 for horse spleen. All these sequences contain a total of 174 amino acids. Table 1 shows that the amino acid composition calculated from the sequence is in good agreement with values from amino acid analysis of apoferritin used in the sequence determination.

Fig.2 compares the apoferritin sequences of human liver, human spleen, horse spleen and rat liver and indicates the amino acid substitutions found. There are 23 substitutions between human liver and horse spleen apoferritins and 9 changes between human liver and human spleen. The rat liver sequence has 29 changes, plus an insertion of 8 residues, when compared with the human liver sequence.

Location of these changes on the three-dimensional structure of horse spleen apoferritin [18] shows the majority, including the 8 residue insertion, to be on the external surface of the molecule, i.e. on helices A (10–39), C (92–120) and in the loop region (73–91) and are therefore easily accommodated within the tertiary structure. There are few significant changes in the tightly packed hydrophobic regions of the helices, or in the critical structural regions, involved in assembly or function of the molecule.

Table 1

Amino acid composition of human liver apoferritin

Amino acid	Number of residues/subunit	
	1	2
Cys	1	1.8*
Asx	19	20.9
Thr	7	7.3
Ser	8	9.1
Glx	23	24.5
Pro	4	4.7
Gly	12	11.2
Ala	15	14.3
Val	6	7.8
Met	4	3.8
Ile	3	3.4
Leu	27	26.0
Tyr	7	7.0
Phe	8	7.0
His	7	7.2
Trp	1	0.8
Lys	12	11.1
Arg	10	8.9

The amino acid composition is expressed as residues of each amino acid per subunit of M_r 19900. Column 1 is calculated from the amino acid sequence and column 2 from amino acid analysis of apoferritin used in the sequence determination.

* Determination as carboxymethyl cysteine

Only 3 substitutions (48, 53 and 174) have occurred on the internal surface of the molecule in human liver and cause no alteration to the overall structure. Two additional proline residues (93 and 157) are found in the human liver sequence when compared with horse spleen. These changes occur at the beginning of helix C and in the DE turn respectively, positions where proline residues can be accommodated.

Strongly conserved regions of the ferritin molecule are those around the 2-, 3- and 4-fold symmetry axes. All the residues lining the 3- and 4-fold channels are conserved, with one exception Ser 131 to Thr and this may be regarded as a conservative change. Many polar residues on the B helix near the 2-fold axis, which are disordered in the structure, are conserved in all four sequences, i.e.

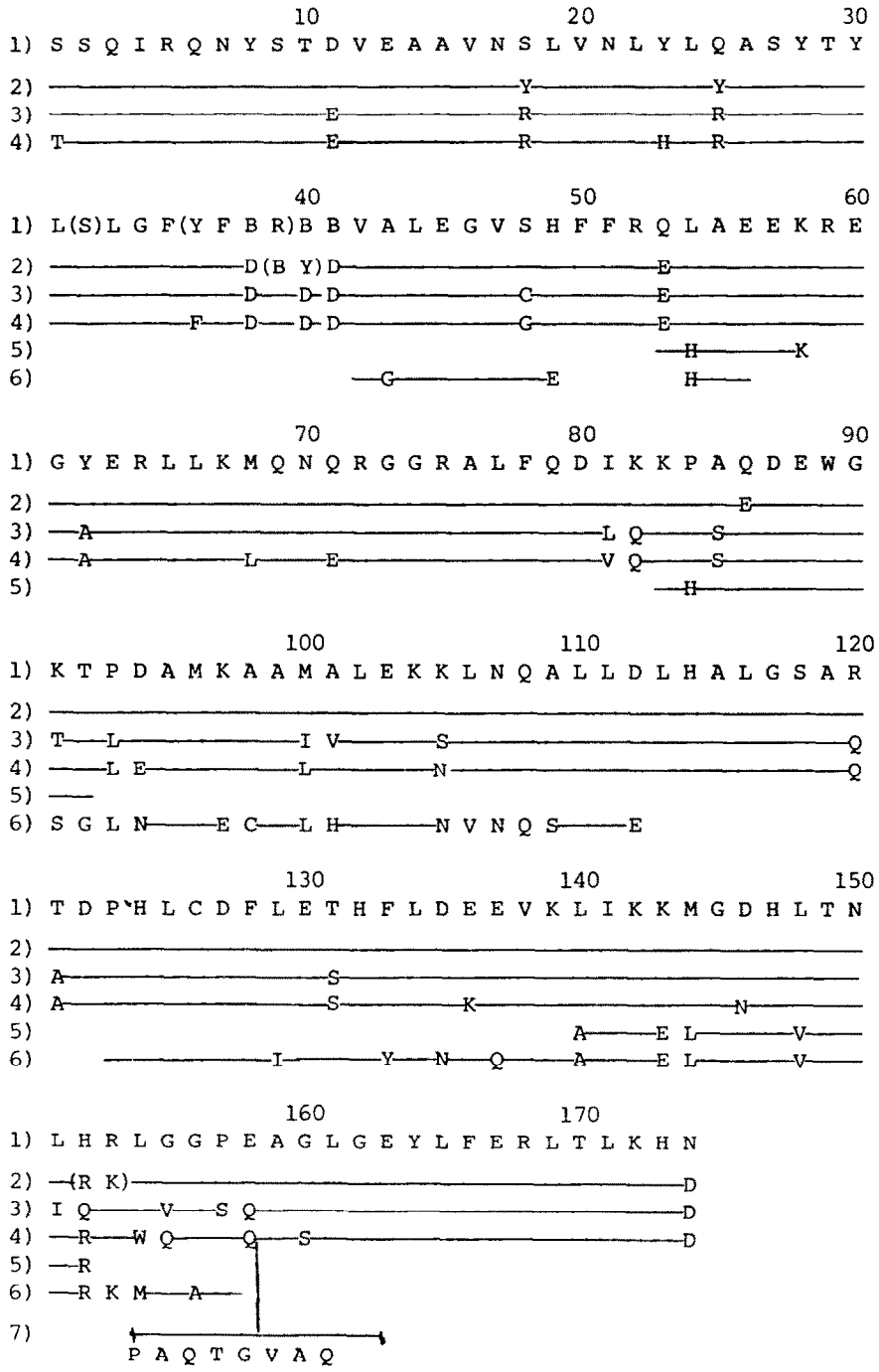

Fig.2. Comparison of ferritin sequences. (1) Sequence of human liver apoferritin. (2) Sequence of human spleen apoferritin [6]. (3) Sequence of horse spleen apoferritin [5]. (4) Sequence of rat liver apoferritin [7]. (5) Partial sequence of minor component of human liver apoferritin. (6) Partial sequence of minor component of human spleen apoferritin. (7) 8 residue insertion in rat liver apoferritin.

residues 52, 56, 57, 59, 60, 63 and 64. They may have a functional role in iron core nucleation [4], although not all these residues are conserved in the human liver cDNA sequence recently determined in [19].

Many substitutions have occurred in residues 150–160, including the insertion in rat liver. This region is on the outer edge of the 4-fold channel, contrasting with the 3-fold channel where few changes have taken place.

Four small peptides (yields 4–6%) were isolated and sequenced during the course of the human liver sequence determination (fig.2). They can be aligned with the major sequence, with some changes and may correspond to a minor sequence. These peptides have also been found in the human spleen sequence. The recent studies on cDNA of human liver apoferritin have established a sequence (43% differences in amino acid sequence) which encompasses the minor sequence found in the human liver and spleen sequence determinations [19]. Work on cDNA for the major subunit in rat liver ferritin has provided evidence in favour of separate mRNAs for the two subunits and the existence of a multigene family encoding for the major or light subunit [20]. This evidence concerning the origin of heterogeneity in tissue ferritins is accumulating in favour of H and L subunits being separate gene products from primary sequence studies and cloned DNA techniques.

ACKNOWLEDGEMENTS

We wish to thank Dr J. Findlay (SERC Amino Acid Sequencing Service, University of Leeds) for the liquid phase sequence determinations on 2 large peptides analysed in this study, Dr Anne Dell (Imperial College of Science & Technology, London) for the mass spectrometry determination of the N-terminal pentapeptide and the Medical Research Council for financial support. We thank Professors R.R. Crichton and H.N. Munro for permission to reproduce their amino acid sequence data in fig.2.

REFERENCES

[1] Chukrov, F.V., Zoyagin, B.B., Gorshkov, A.I., Yermilova, L.P. and Balashova, V.V. (1973) Int. Geol. Rev. 16, 1131–1143.
[2] Banyard, S.H., Stammers, D.K. and Harrison, P.M. (1978) Nature 271, 282–284.
[3] Arosio, P., Adelman, T.G. and Drysdale, D.W. (1978) J. Biol. Chem. 253, 4451–4458.
[4] Ford, G.C., Harrison, P.M., Rice, D.W., Smith, J.M.A., Treffry, A., White, J.L. and Yariv, J. (1984) Phil. Trans. R. Soc. Lond. B., in press.
[5] Heusterspreute, M. and Crichton, R.R. (1981) FEBS Lett. 129, 322–327.
[6] Wustefeld, C. and Crichton, R.R. (1982) FEBS Lett. 150, 43–48.
[7] Leibold, E.A., Brown, A.J.P., Aziz, N. and Munro, H.N. (1983) in press.
[8] Penders, T.J., De Rooij-Dijk, H.H. and Liejnse, B. (1968) Biochim. Biophys. Acta 168, 588–590.
[9] Konigsberg, W. (1972) Methods Enzymol. 25, 185–188.
[10] Hirs, C.H.W. (1967) Methods Ezymol. 11, 199–203.
[11] Lewis, W.G., Basford, J.M. and Walton, P.L. (1978) Biochim. Biophys. Acta 522, 551–560.
[12] Atassi, M.Z. and Habeeb, A.F.S.A. (1972) Methods Enzymol. 25, 546–553.
[13] Landon, M. (1977) Methods Enzymol. 47, 145–149.
[14] Suran, A.A. and Tarver, H. (1965) Arch. Biochem. 111, 399–406.
[15] Gray, W.R. (1967) Methods Enzymol. 11, 139–151.
[16] Woods, K.R. and Wang, K.T. (1967) Biochim. Biophys. Acta 133, 369–370.
[17] Gray, W.R. (1972) Methods Enzymol. 25, 121–138.
[18] Addison, J.M., Fitton, J.E., Ford, G.C., Harrison, P.M., Lewis, W.G., Rice, D.W., Smith, J.M.A., White, J.L., Crichton, R.R., Wustefeld, C., Heusterspreute, M., Brown, A.J.P., Liebold, E.A., Aziz, N. and Munro, H.N. (1983) Proteins of Iron Storage and Transport, Elsevier/North Holland, New York, in press.
[19] Constanco, F., Santoro, C. and Cortese, R. (1983) unpublished results.
[20] Brown, A.J.P., Leibold, E.A. and Munro, H.N. (1983) Proc. Natl. Acad. Sci. USA 80, 1265–1269.