Fragments of Bovine Serum Albumin Produced by Limited Proteolysis. Isolation and Characterization of Tryptic Fragments†

Theodore Peters, Jr.,* and Richard C. Feldhoff‡

ABSTRACT: Several fragments of bovine serum albumin have been isolated following limited tryptic hydrolysis and their positions then determined in the bovine serum albumin sequence published by J. R. Brown ((1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 591). When bovine serum albumin was coupled to palmityl-aminoethylamino-agarose and digested with trypsin, two fragments were obtained: (a) peptide 115–184, containing the highly aromatic disulfide loop 3 of Brown's model, and (b) a larger fragment, residues 377–581, containing disulfide loops 7–9. This fragment constitutes the third of the three domains of the albumin molecule. From bovine serum albumin digested in solution, peptide 115–184 was again obtained, as well as (c) a 39,000-dalton fragment identified as residues 198–581, loops 4–9 of the second and third domains, but with a long, tryptophan-containing segment 204–238 missing from loop 4. The ability to isolate these fragments without cleaving disulfide bridges is partial confirmation of the proposed model of bovine serum albumin as a series of nine independent loops.

Large fragments of serum albumin are of use in the study of the location and properties of ligand-binding sites and in the determination of other structural and functional features of the molecule. Fragments obtained by proteolysis have intact disulfide bonds and unaltered amino acid side chain groups, and so are more nearly native than fragments prepared by chemical cleavage techniques.

Thus far only one large fragment of bovine albumin obtained by limited tryptic digestion has been well characterized. This is the ~40,000-dalton peptide termed T-23 by King and Spencer (1970). They showed that it included the carboxyl end of the albumin chain, and had one binding site for octanoate and one for L-tryptophan. On the other hand, a slightly larger fragment obtained by cyanogen bromide cleavage, which should include all of T-23, bound these ligands only weakly if at all.

The present work employed two types of digestion in an effort to obtain additional fragments of bovine albumin. In the first, digests of albumin in solution were prepared by the method of King and Spencer (1970), but with added precautions to remove or inactivate residual trypsin. In the second, tryptic hydrolysis of bovine albumin bound to palmityl-aminoalkylamino-agarose was employed as a means of obtaining fatty acid binding fragments. Four fragments have been purified and their locations within the parent molecule determined in accordance with the recently published sequence and disulfide bridge structure of bovine albumin (Brown, 1975). Two of the fragments, one from each digestion procedure, were found to be the same peptide. The T-23 fragment of King and Spencer (1970) was isolated and studied further in an effort to define its structure and location within the albumin chain. The fourth fragment was a fatty acid binding polypeptide of molecular weight 23,000 isolated from the digestion of albumin on fatty acyl-agarose. A preliminary description of this polypeptide has appeared (Peters et al., 1973).

Experimental Section

Materials. Commercial materials used were crystalline bovine serum albumin, Armour Pharmaceutical Co., Kankakee, Ill.; trypsin (tosylamido-2-phenylethyl chloromethyl ketone treated), soybean trypsin inhibitor, Pronase B, crystalline swine pepsin, and crystalline swine carboxypeptidases A and B, both diisopropyl fluorophosphate treated, Worthington Biochemical Corp., Freehold, N.J.; Sepharose 4B and Sephadex G-75, G-100, and G-150, Pharmacia Fine Chemicals, Piscataway, N.J.; reagents for Edman degradation and automated amino acid analysis, Pierce Chemical Co., Rockford, Ill.; DEAE-cellulose (DE-52) and CM-cellulose (CM-52), Reeve-Angel Co., Clifton, N.J. Hydrazine (95%) was further purified by azeotropic distillation with toluene. Formic acid used for cyanogen bromide reactions was purified by repeatedly freezing and thawing (McMenamy, 1971).

Half-cystinyl bovine serum albumin, with a half-cystine residue coupled to the lone thiol group of the albumin to minimize dimerization, and palmityl-aminoethylamino-agarose were prepared as previously described (Peters et al., 1973). The binding capacity of the palmityl-agarose was about 8 mg of albumin/ml of packed wet gel. Bovine serum albumin was defatted by passage through a column of anion exchange resin (AG 1-X8, 200–400 mesh, Bio-Rad Laboratories, Richmond, Calif.) at pH 3 as described by Scheider and Fuller (1970). Soybean trypsin inhibitor coupled to agarose was the gift of Dr. Hanspaul Hagenmaier of the University of Tübingen; 1 ml of this preparation inactivated in excess of 0.5 mg of trypsin.

Methods. Amino-terminal residues were identified by paper chromatography of the dinitrophenyl derivatives obtained by the procedure of Fraenkel-Conrat et al. (1955).

† From The Mary Imogene Bassett Hospital (affiliated with Columbia University), Cooperstown, New York 13326. *Received March 24, 1975.* This paper is the first in a series describing the purification and properties of fragments of bovine serum albumin. This work was supported by U.S. Public Health Service Research Grant HL-02751 and the Stephen C. Clark Research Fund of The Mary Imogene Bassett Hospital.

‡ Present address: Department of Biological Chemistry, The Milton S. Hershey Medical Center, Hershey, Pennsylvania 17033.

Solvents used in the chromatography were *tert*-amyl alcohol–1 *N* NH_4OH, 4:1, v/v, and 1.5 *M* sodium phosphate buffer (pH 6.0). In ambiguous cases dinitrophenyl amino acids were hydrolyzed (sealed tube, 2 *N* NH_4OH, 100°, 2 hr) and dried under vacuum, and the resulting free amino acids were identified by amino acid analysis.

Amino-terminal sequences were determined on 0.1-μmol specimens by a manual adaptation of the Edman procedure (Taniuchi and Anfinsen, 1966); the thiazolinone derivatives were hydrolyzed in vacuo in sealed ampoules with hydriodic acid at 128° (Smithies et al., 1971) and the resulting free amino acids quantitated on the amino acid analyzer. Recovery factors given by those authors were employed; carboxymethylcysteine was found to appear as alanine with a yield of 56%.

Carboxyl-terminal residues were determined by hydrazinolysis (Niu and Fraenkel-Conrat, 1955). Approximately 0.1 μmol of peptide was heated in vacuo in a sealed ampoule at 100° for 10 hr with 0.5 ml of hydrazine, then taken to dryness over H_2SO_4. Amino acid hydrazides were removed by shaking with 0.2 ml of benzaldehyde and 1 ml of water for 2 hr at 23°. The aqueous phase was applied to the amino acid analyzer.

Carboxyl-terminal sequences were detected by treatment of 0.1-μmol aliquots of peptide with a mixture of carboxypeptidases A and B (1:30, w/w, and 4 units/μmol, respectively) in 0.05 *M* sodium phosphate buffer (pH 7.4) at 37° for 0.2–5 hr. Appropriate blanks were included. Remaining peptide was precipitated with 5% Cl_3CCOOH. The supernatants were washed twice with ether to remove Cl_3CCOOH and then assayed on the amino acid analyzer. In some instances the buffer was 0.1 *M* ammonium bicarbonate (pH 8.0) and the free amino acids were isolated from the peptide by continuous dialysis; the ammonium bicarbonate was removed in vacuo at 40°.

Methionyl bonds in peptides were cleaved with cyanogen bromide according to the procedure of King and Spencer (1970) using 0.5 mg of CNBr/mg of peptide. Excess reagents were removed by lyophilization or, if soluble products were required, by passage through a Sephadex G-10 column.

Amino acid analyses were conducted on a single-column system employing Aminex A-6 resin (Bio-Rad Laboratories) with three buffers and an automatic programmer (Bio-Cal, LKB Instruments, Rockville, Md.). Peptides were hydrolyzed in sealed ampoules in vacuo with redistilled 5.8 *N* HCl at 110° for 24 hr. Norleucine was employed as an internal standard and was added before hydrolysis in all analyses, including the Edman, hydrazinolysis, and carboxypeptidase techniques. Tryptophan was determined on separate aliquots of the peptides by the method of Spies (1967) using Pronase B for digestion.

Electrophoresis in cellulose acetate (Millipore Corp., Bedford, Mass., or Helena Laboratories, Beaumont, Texas) was performed at pH 8.6 in sodium barbital buffer, μ = 0.075, or at pH 5 in 0.048 *M* pyridine–0.05 *M* acetic acid; the peptides were stained with Ponceau S in 10% Cl_3CCOOH. Polyacrylamide gel electrophoresis with dodecyl sulfate was conducted in 6% gels according to Weber et al. (1972). Thin layer gel exclusion chromatography followed the procedure of Andrews (1970). Peptide chromatograms were prepared as described by Benjamin and Weigle (1971) using complete digestion with trypsin of the reduced and carboxyamidomethylated peptides.

Peptide fractionation procedures were conducted at 20–25° unless otherwise indicated. Column effluents were monitored at 280 nm with a UA-5 Absorbance Monitor (Instrumentation Specialties Co., Lincoln, Nebr.). Fractions were tested for albumin by a spot-plate version of the Bromocresol Green-binding procedure of Doumas et al. (1971). Pooled fractions were concentrated under nitrogen pressure on an ultrafilter (Amicon Corp., Bedford, Mass.), using a membrane appropriate for the molecular size of the fraction. Lyophilization was avoided as it often led to insoluble products, as noted by King and Spencer (1970).

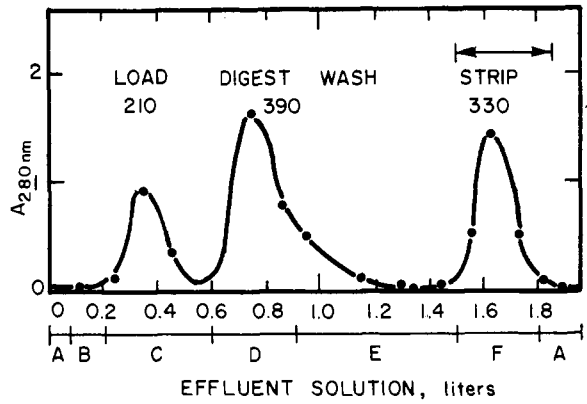

FIGURE 1: Digestion of albumin on a fatty acyl-agarose column. Figures above the peaks indicate the approximate total milligrams of protein or peptide measured by A_{280} of the effluent. The arrows indicate material taken for further fractionation. The solutions applied to the column were: (A) 0.1 *M* NaCl–0.02 *M* sodium phosphate buffer (pH 7.4); (B) same, containing 1%, w/v, defatted bovine serum albumin; (C) 0.1 *M* ammonium bicarbonate (pH 8.8); (D) same, containing 0.06% trypsin, w/v, applied in three 100-ml portions at hourly intervals; (E) 0.1 *M* ammonium acetate (pH 6.3); (F) 0.025 *N* NaOH in 0.5 ethanol, v/v.

Digestion of Albumin in Solution. In a typical digestion, 1 g of half-cystinyl bovine serum albumin was brought to pH 8.15 at 25° in 20 ml of 0.04 *M* Tris-HCl buffer. Trypsin (1 mg) was added and the solution gently stirred at 25° for 60 min, adding 1 *N* NaOH to maintain constant pH. The solution was then immediately put through a 1 × 8-cm column of soybean trypsin inhibitor-agarose in 0.1 *M* ammonium bicarbonate (pH 8.0), and eluted with 6 ml of water. This operation was completed within 20 min. To the eluate was added 2 mg of soybean trypsin inhibitor and sufficient benzamidine hydrochloride to make 0.001 *M*. The solution was adjusted to pH 3.0 with 6 *N* HCl, applied to a 5 × 95-cm column of Sephadex G-150, and eluted with 0.23 *M* ammonium formate (pH 3.0) at 42 ml/hr for the first step in fractionation.

Digestion of Albumin Bound to Palmityl-aminoethylamino-agarose. Defatted bovine serum albumin, 1 g in 0.1 *M* NaCl containing 0.02 *M* sodium phosphate buffer at pH 7.4, was loaded onto a 2.5 × 25-cm column of palmitylaminoethylamino-agarose at 22°. The unbound portion was rinsed off. The bound albumin was then digested with trypsin at 37° for several hours. The column was rinsed well to remove soluble products, mostly small peptides, and then the portion resisting digestion was eluted using NaOH–ethanol. The elution pattern is shown in Figure 1. The NaOH–ethanol eluate was chilled in ice as it was obtained and brought to pH 3.0 by the addition of formic acid. It was then diluted with an equal volume of water, concentrated to about 25 ml by ultrafiltration with a PM-10 membrane, washed twice with a tenfold excess of 0.23 *M* ammonium formate (pH 3.0), and reconcentrated to a 25-ml volume.

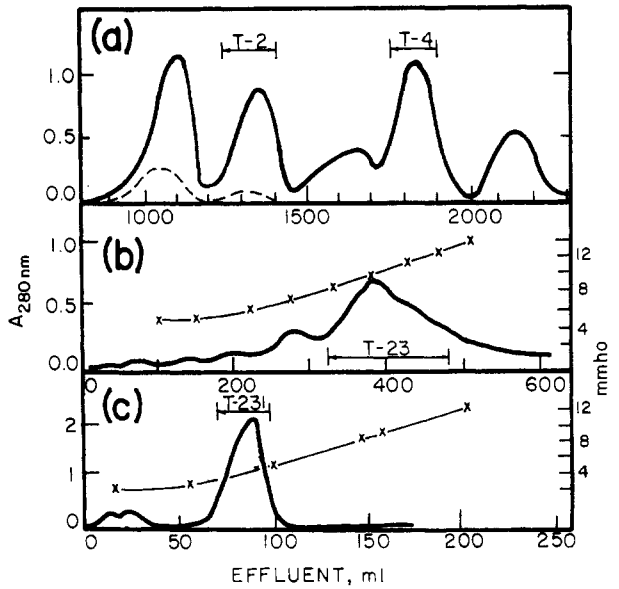

FIGURE 2: (a) Fractionation of a digest of soluble albumin on a Sephadex G-150 column pH 3.0. (—) A_{280} of the effluent; (- - -) relative color change of Bromocresol Green noted in a spot test. The arrows outline the fractions taken for further study. (b) Purification of fraction T-2 on a 2.3 × 7.5-cm column of DEAE-cellulose. The gradient was formed using two vessels each containing 500 ml of 0.1 *M* Tris-Cl buffer (pH 7.6) with 0.2 *M* NaCl in the second vessel. Flow rate was 50 ml/hr. Conductivity (22°) is shown by X—X. The numbering of peak T-23 followed the practice of King and Spencer (1970). (c) Repurification of fraction T-23 on 1 × 11-cm column of CM-cellulose. The gradient was 500 ml of 0.04 *M* sodium acetate buffer (pH 5.0) and 500 ml of 0.2 *M* sodium acetate buffer (pH 5.0) containing 0.1 *M* NaCl. Flow rate was 20 ml/hr. The lines have the same meaning as in Figure 2b.

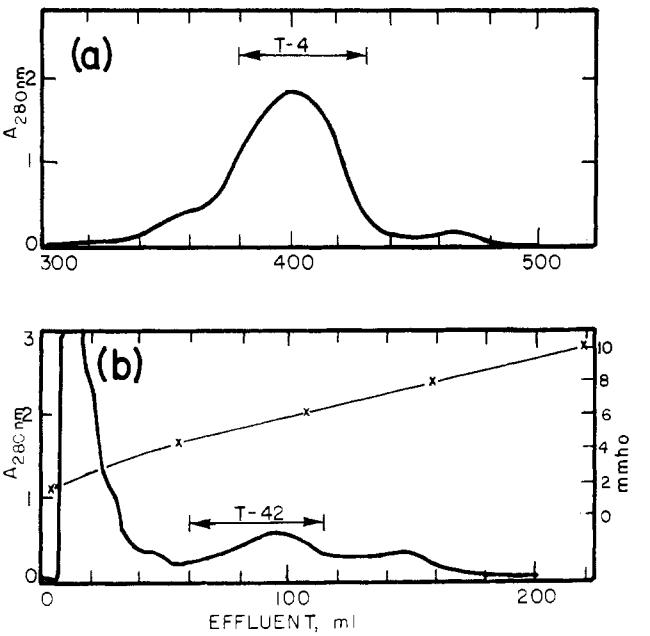

FIGURE 3: Purification of fraction T-42. (a) Rerun of fraction T-4 from Figure 2 on a G-100 column, 2.5 × 100 cm, in 0.23 *M* ammonium formate (pH 3.0), flow rate 20 ml/hr. (b) Fractionation on a CM-cellulose column, 1 × 12.5 cm. The gradient was formed with 400 ml of 0.03 *M* ammonium acetate (pH 5.8) and 400 ml of the same buffer containing 0.3 *M* NaCl. Flow rate was 20 ml/hr. Solid lines indicate A_{280}; (X—X) Conductivity at 22°; arrows indicate fractions taken for further study.

Solutions were stored frozen pending fractionation.

Results

Peptides Obtained from Albumin Digested in Solution. In preparing digests of albumin in solution several precautions were taken to minimize the possibility of continued tryptic activity during subsequent fractionation procedures at pH 7–9. Compared to the procedure used by King and Spencer (1970), the ratio of trypsin to albumin was reduced from 1:100 to 1:1000, w/w, and the temperature raised from 0 to 25°. Preliminary studies showed no qualitative differences in the patterns of the two types of digestions on cellulose acetate electrophoresis, and no differences were detected in the T-23 fractions prepared from them. In addition, remaining trypsin was removed by passage through a column of immobilized trypsin inhibitor, and trypsin inhibitors were added in the first fractionation step.

The first step of fractionation on a Sephadex G-150 column duplicated the conditions of King and Spencer (1970), and similar elution patterns were obtained (Figure 2a). Testing with Bromocresol Green helped to locate albumin or large fragments in the fractions. The numbering system of King and Spencer was retained for the peaks, in which T-2 means the second peak of the first fractionation step of a tryptic digest. Peak T-2 was rerun on the G-150 column until it appeared homogeneous.

Figure 2b and c shows the subsequent steps in the purification of fraction T-2. The use of DEAE-cellulose (Figure 2b) followed the conditions of King and Spencer (1970); the product, fraction T-23, was named as the third peak in the second purification step of fraction T-2. In our hands T-23 showed two bands on cellulose acetate electrophoresis at pH 5, and the less positively charged of these bands was purified on CM-cellulose as shown in Figure 2c to yield fraction T-231.

Peak T-4 was further fractionated by first passing through a G-100 column followed by purification on CM-cellulose (Figure 3). In accordance with the above nomenclature, this fraction is called T-42. Peaks 3 and 5 of the G-150 eluate (Figure 2a) contained little material, and were not studied further.

The homogeneity of fractions T-231 and T-42 on sodium dodecyl sulfate gel electrophoresis and cellulose acetate electrophoresis is shown in Figure 4. Both peptides are less negatively charged than is intact bovine serum albumin. Their molecular weights were estimated as 37,000–40,000 and 9000, respectively (Table II).

Peptides from Albumin Digested on Fatty Acyl-agarose. The rationale for the digestion of albumin bound to palmityl-agarose is that coupling of a tightly bound ligand protects a portion of the protein against proteolysis (Markus et al., 1967a). Digestion of albumin which is immobilized by coupling to fatty acyl-agarose should thus offer a means of obtaining fragments containing primary fatty acid binding sites; it has the advantage that residual trypsin is removed during the washing procedure before elution of the remaining, trypsin-resistant portion of the molecule (Figure 1).

Sephadex G-100, rather than G-150, was chosen for the initial fractionation of these digests owing to the smaller size of the peptides released, as determined by preliminary surveys using the thin-layer gel filtration technique. The elution pattern (Figure 5a) usually showed an initial peak caused by turbidity followed by intact albumin and two peptide peaks. The first peptide, which showed a slight affinity for Bromocresol Green, was named T-A, and the second peptide was named T-B. Peptide T-A was rerun on

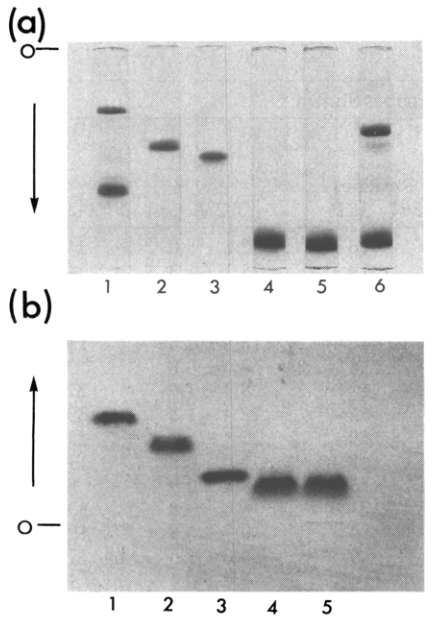

FIGURE 4: Electrophoresis of the tryptic peptides isolated from bovine albumin. (a) Electrophoresis in dodecyl sulfate acrylamide gel of the peptides, after reduction with β-mercaptoethanol. (1) Bovine serum albumin and cytochrome *c* markers, mol wt 66,000 and 12,000, respectively. (2) T-231; (3) T-A; (4) T-B; (5) T-42; (6) bovine serum albumin and the "Phe" fragment of bovine serum albumin (Peters and Hawn, 1967), mol wt 8,600, as markers; o indicates the origin; direction of migration toward the anode is shown by arrows. T-231 and T-42 were obtained from albumin digested in solution; T-A and T-B were obtained from albumin digested while bound to palmityl-aminoethylamino-agarose. (b) Electrophoresis on cellulose acetate at pH 8.6. (1) bovine serum albumin; (2–5) same meaning as in Figure 4a.

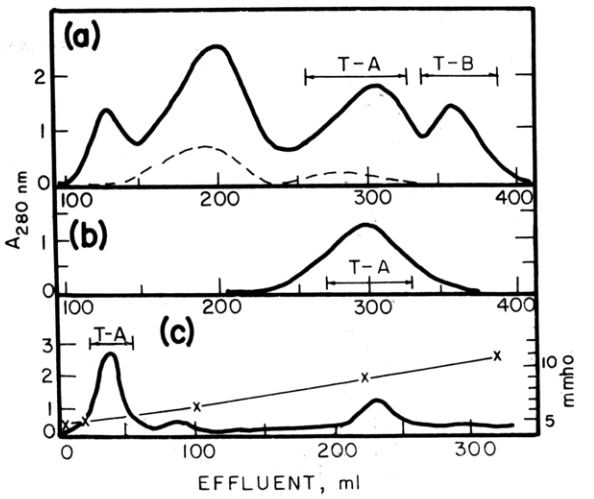

FIGURE 5: (a) Fractionation of digest of immobilized albumin. The eluates from three digests (Figure 1) were combined and applied to a 2.5 × 95-cm column of G-100 in 0.23 *M* ammonium formate (pH 3.0) and eluted at 22 ml/hr. (b) Repurification of fraction T-A on the same column. (c) Fractionation of repurified T-A on a column of DEAE-cellulose, 2.3 × 8 cm. Details are as in Figure 2b, except that the gradient was 410 ml of 0.1 *M* Tris-HCl buffer (pH 7.6) and 410 ml of this buffer containing 0.5 *M* NaCl.

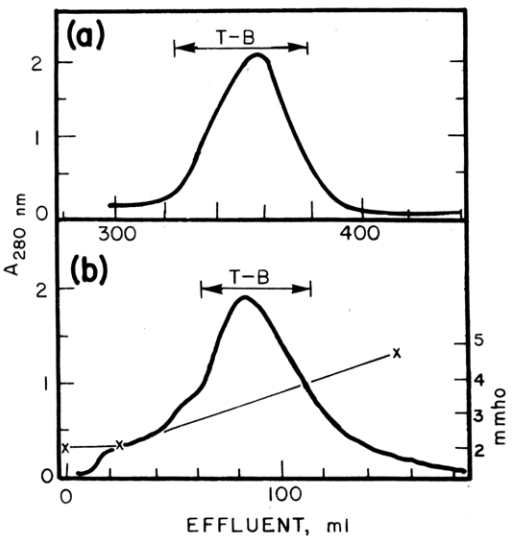

FIGURE 6: Purification of fraction T-B. (a) Rerun of fraction T-B from Figure 5 on a G-75 column, 2.5 × 95-cm, in 0.23 *M* ammonium formate, pH 3.0, flow rate 20 ml/hr. (b) Fractionation on a DEAE-cellulose column, 1.1 × 25 cm. The gradient was formed with 200 ml of 0.05 *M* Tris-HCl (pH 8.0) and 200 ml of the same buffer containing 0.2 *M* NaCl. Flow rate was 50 ml/hr.

G-100 and then purified on DEAE-cellulose (Figure 5b,c) to yield a product homogeneous on electrophoresis (Figure 4) with a molecular weight of about 23,000. Peptide T-B was repurified on G-75 and then fractionated on DEAE-cellulose, as shown in Figure 6. Upon electrophoresis on several media the T-B peak resembled peptide T-42 which had been obtained from digestions of albumin in solution (Figure 4).

Characterization of Fragments. The amino acid compositions of the four peptides T-231, T-42, T-B, and T-A are given in Table I. Their molecular weights estimated from gel filtration or dodecyl sulfate gel electrophoresis (Table II) plus their terminal sequences (Tables II and III) served to locate each peptide precisely in the amino acid sequence of bovine serum albumin derived by Brown (1975). The calculated composition of the corresponding segment from the bovine serum albumin sequence is listed in the column adjacent to that for each peptide in Table I.

The apparent identity of peptides T-42 and T-B on electrophoresis (Figure 4) was confirmed by amino acid composition (Table I), molecular weight (Table II), amino-terminal sequences (Table III), and by the carboxyl-terminal sequence, Met-Arg, demonstrated by the release only of arginine from each fragment upon cyanogen bromide cleavage (Table II). These fragments correspond to residues 115–184, or loop 3 of Brown's model, with the terminal sequences H_2N-Leu-Lys-Pro-Asp—Lys-Ile-Glu-Thr-Met-Arg-COOH. Henceforth they will be referred to by the more accurate designation $T_{115-184}$, signifying the segment of residues 115–184 obtained by tryptic cleavage. The high content of aromatic residues in this peptide causes its high absorbancy, A_{280nm} (1 g/l.) of 1.84 (Table II). The comparable absorbancy for bovine serum albumin is 0.66 (Cohn et al., 1947).

Fragment T-A is seen by composition, size, and terminal sequence to comprise the carboxyl-terminal third of the albumin molecule, residues 377–581 of Brown's model, containing loops 7, 8, and 9. The terminal sequences of this segment are H_2N-His-Leu-Val-Asp—Thr-Gln-Thr-Ala-Leu-Ala-COOH. Appearance of amino-terminal residues of proline and glutamic acid following cyanogen bromide cleavage (Table II) confirmed the presence of the -Met-Pro_{444}- and -Met-Glu_{546}- sequences. In addition, mild digestion with pepsin caused the appearance of an electrophoretic band corresponding to the known carboxyl-terminal "Phe" fragment (Peters and Hawn, 1967), which is loop 9 of the

Table I: Amino Acid Composition of Tryptic Fragments of Bovine Albumin.

Amino Acid	Residues per Total Residues Shown							
	T-42[a]	T-B[b]	Residues 115–184[c]	T-A[d]	Residues 377–581[c]	T-23[e]	T-23[f]	Residues 198–203 plus 239–581[c]
Aspartic acid	7.2 ± 0.5	7.5 ± 1.1	7	17.0 ± 1.0	16	32.6 ± 5.0	34.5	34
Threonine[g]	1.8 ± 0.2	2.3 ± 0.3	2	19.4 ± 1.1	19	25.1 ± 1.3	24.0	23
Serine[h]	0.8 ± 0.4	0.7 ± 0.2	0	11.3 ± 0.4	9	19.5 ± 1.7	18.1	17
Glutamic acid	8.8 ± 0.9	9.6 ± 0.8	9	27.4 ± 1.0	28	45.7 ± 4.1	47.1	46
Proline	5.3 ± 0.5	4.8 ± 0.6	5	12.4 ± 0.7	12	18.9 ± 2.6	19.0	18
Glycine	3.1 ± 0.3	3.0 ± 0.3	3	5.4 ± 0.2	4	7.4 ± 0.8	7.5	6
Alanine	5.9 ± 0.5	5.8 ± 0.3	6	13.0 ± 0.4	13	28.4 ± 1.6	27.7	28
½-Cystine	3.9 ± 0.6	4.9 ± 0.9	4	12.4 ± 1.0	12	22.4 ± 1.8	23.1	24
Valine	1.4 ± 0.5	1.4 ± 0.3	1	16.5 ± 0.9	19	25.1 ± 2.7	22.4	25
Methionine	0.9 ± 0.3	0.7 ± 0.3	1	1.6 ± 0.1	2	2.0 ± 0.3	1.6	2
Isoleucine	2.2 ± 0.4	1.6 ± 0.3	2	4.2 ± 0.4	5	8.8 ± 0.9	8.0	10
Leucine	6.8 ± 0.5	6.5 ± 0.3	7	19.9 ± 0.6	21	36.0 ± 2.8	33.7	36
Tyrosine	6.5 ± 0.6	5.6 ± 0.3	6	4.1 ± 0.8	4	9.7 ± 0.9	10.0	11
Phenylalanine	4.0 ± 0.7	3.8 ± 0.7	4	8.9 ± 0.5	9	14.0 ± 1.6	12.0	13
Histidine	0.7 ± 0.2	0.8 ± 0.4	1	3.3 ± 0.6	4	7.3 ± 0.4	9.2	9
Lysine	8.3 ± 0.6	8.1 ± 0.8	8	20.2 ± 0.6	20	34.2 ± 3.3	36.3	35
Arginine	2.8 ± 0.3	2.2 ± 0.9	3	7.8 ± 0.6	8	12.1 ± 0.9	13.5	12
Tryptophan	1	1	1	0.0 ± 0.1	0	0.0 ± 0.1	1.0	0
(Amide-NH_3)			(3)		(12)			(15)
Total residues	70.4	70.3	70	204.8	205	349.2	348.7	349
Molecular weight			8313		23138			39442

[a] Mean ± SD of ten analyses of two preparations. [b] Mean ± SD of seven analyses of two preparations. [c] Calculated from the bovine serum albumin sequence reported by Brown (1975). Residues 400–402, which are blank in the bovine serum albumin sequence, were assumed to be Phe-Gln-Asn by homology with the sequence of human albumin (Behrens et al., 1975). [d] Mean ± SD of nine analyses on three preparations. [e] Mean ± SD of ten analyses of two preparations. [f] Data of King and Spencer (1970) calculated on the basis of 349 total residues. [g] Corrected for 4% loss in 24-hr hydrolysis. [h] Corrected for 10% loss in 24-hr hydrolysis.

Table II: Molecular Weights, Specific Absorbances, and Terminal Residues of Tryptic Fragments of Bovine Albumin.

	T-42	T-B	T-A	T-23
Molecular weight				
Dodecyl sulfate gel electrophoresis	9,000	9,000	23,500	37,000
Thin-layer Sephadex	10,000	10,000	23,300	40,000
A_{280nm} (1 g/l.)	1.84		0.34	0.48
Amino-terminal residues[a]	Val,Leu/Ile	Val,Leu/Ile	His His,Pro,Glu[b]	Leu,Thr,Val,Tyr
Carboxyl-terminal residues	0.56 Arg[c]	0.56 Arg[c]	0.36 Ala,0.05 Leu[d]	0.79 Ala,0.10 Leu[d]
Carboxyl-terminal sequence, carboxypeptidase	–(Ile,Leu,Glu,Tyr, Lys)-Met-Arg		–(Val,Gln,Thr, Leu)-Ala	–(Thr,Gln,Ser, Leu)-Ala

[a] Determined by use of fluorodinitrobenzene. [b] Amino-terminal residues found after treatment with CNBr as described in the Experimental Section. [c] Released as free amino acid (soluble in 10% Cl_3CCOOH, w/v), following reaction with CNBr, mol/mol. [d] By hydrazinolysis, mol/mol.

bovine serum albumin molecule. Likewise the peptide fingerprint of T-A (Figure 7) includes all of the major peptides seen in the fingerprint of the "Phe" fragment. Thus peptide T-A is more accurately called $T_{377-581}$.

The T-23 fragment was shown by King and Spencer (1970) to contain the carboxyl terminus of the albumin molecule, a conclusion supported by our results with carboxypeptidase treatment and hydrazinolysis (Table II), and by noting that its peptide chromatogram contained all of the peptides of $T_{377-581}$ as well as those of the carboxyl-terminal "Phe" fragment.

Localization of the amino terminus of peptide T-23 raised some problems. Its apparent molecular weight of 37,000–40,000 (Table II) requires that it extend beyond the link between loops 4 and 5 of the bovine serum albumin model, yet none of our preparations was found to contain tryptophan, an observation which is inconsistent with the inclusion of loop 4 in the fragment. Our preparations of T-23 were apparently contaminated with the closely related fraction designated T-22 by King and Spencer (1970), since two bands were seen on cellulose acetate electrophoresis at pH 5 and multiple amino-terminal groups were present. We purified a subfraction of this T-23, named T-231, to homogeneity as judged by dodecyl sulfate gel and cellulose acetate electrophoresis (Figure 4), but even this peptide upon Edman degradation showed a significant amount of amino-terminal alanine, serine, or cystine (all of which appear as alanine after hydrolysis with hydriodic acid) in addition to the valine reported by King and Spencer.

The nature of the amino end of this peptide T-231 was clarified by the realization that it contains two amino-terminal residues, $cystine_{198}$ and $valine_{239}$, and is lacking the tryptophan-containing segment 204–238 of loop 4 of the bovine serum albumin model. This structure is consistent with

Table III: Results of Edman Degradation of Tryptic Fragments of Bovine Albumin.

Peptide	Cycle	Residues/mole[a]												
		Asp	Glu	Pro	Gly	Ala[b]	Abu[b]	Val	Ile	Leu	Tyr	His	Lys	Arg
T-42	1									0.11				
	2												0.40	
	3			0.46										
	4	0.02												
T-B	1									0.26				
	2												0.15	
	3			0.28										
	4	0.03												
T-A	1		0.16		0.10	0.31	0.18	0.23		0.13		*0.88*	0.41	0.22
	2		0.10	0.28	0.13	0.13		0.15		*0.37*				
	3	0.10	0.12			0.15	0.12	*0.52*						
	4	*0.10*												
T-231	1					0.40		0.33						
	2					0.69						0.18		
	3					0.23							0.24	
	4		0.25			0.14			0.12					
	5		0.17			0.26								
Reduced T-231[c]	1					0.10		*0.48*						
	2					0.10								
	3					0.12							*0.16*	
	4		0.30											
	5		0.22			*0.84*								
	6		0.18			*0.38*								

[a] Based on molecular weights given in Table I. Results less than 0.1 were not shown, except for cycle 4 of T-42 and T-B, when no residue reached that value. [b] Serine, cystine, and alanine appear as alanine: threonine appears as α-amino-n-butyric acid (Abu) (Smithies et al., 1971). Carboxymethylcysteine was also found to appear as alanine. [c] The large peptide retained by an ultrafilter (10,000-dalton cutoff) after reduction and carboxamidomethylation of peptide T-231.

the findings of six steps of the Edman procedure (Table III), although cleavage at the trypsin-susceptible bonds Lys–Ser_{285} or Lys–Cys_{555} could still not be excluded. Other evidence that the second terminal group, in addition to $valine_{239}$, is $cystine_{198}$ rather than $serine_{285}$ or $cystine_{555}$ was obtained by reduction and carboxamidomethylation of T-231 followed by ultrafiltration through a PM-10 membrane (cutoff of 10,000 daltons). The retained peptide showed the amino-terminal sequence Val-( )-Lys-Glu-Cys-Cys- upon Edman degradation, consistent with the Val-His-Lys-Glu-Cys-Cys- sequence in the bovine serum albumin model (Table III). It is presumed that the histidine in the second position was modified by the reduction and carboxamidomethylation procedure or was poorly extracted in the Edman procedure. The small peptide in the ultrafiltrate had the composition shown in Table IV, and had the apparent carboxyl-terminal sequence (Table IV), -(Ala,Ser)-Ile-Gln-Lys. This corresponds to the hexapeptide 198–203 of the bovine serum albumin sequence, Cys-Ala-Ser-Ile-Gln-Lys, and so it appears that this segment is attached to the larger piece of fragment T-231 through the single disulfide bond Cys_{198}–Cys_{243}. This structure is represented diagrammatically in Figure 8.

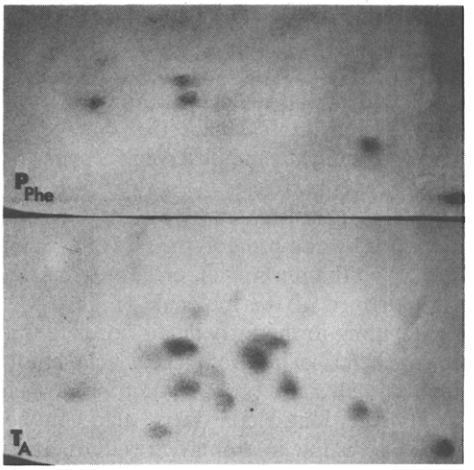

FIGURE 7: Tryptic fingerprint of T-A compared to that of the carboxyl-terminal "Phe" fragment (Peters and Hawn, 1967). Separation in the horizontal direction utilized electrophoresis at pH 3.6, and in the vertical direction chromatography in butanol–acetic acid–water, 4:1:5, v/v.

Discussion

Three fragments, isolated from bovine serum albumin following limited digestion with trypsin, have been described in terms of the bovine serum albumin sequence published by Brown (1975). Two distinct conditions for digestion were employed, the albumin substrate being either in solution or immobilized on palmityl-agarose. One fragment, $T_{115-184}$, was obtained from both types of digests. Each digest also yielded a larger fragment, arising from the carboxyl-terminal portion of the molecule. The one from the digestion of dissolved albumin was the larger and more complex of the two. The three fragments are all native in that no disulfide bonds were broken, and study of the fragments should permit evaluation of the importance of various regions of the 9-loop albumin molecule with respect to conformation, antigenicity, and ligand-binding properties of the intact protein. Subsequent reports in this series will describe the isolation of other fragments, using pepsin, and measurements of α-helix and β-form content, immunological activity, and binding of palmitate, bilirubin, Bromocresol Green and 8-anilino-1-naphthosulfonate.

The smallest tryptic fragment, $T_{115-184}$, is loop 3 of the 9-loop albumin model (Brown, 1974,1975) (Figure 8). It contains the first tryptophan of bovine serum albumin, the

Table IV: Amino Acid Composition and Carboxyl-terminal Sequence of the Small Peptide Released from T-231 upon Reduction and Carboxamidomethylation

Amino Acid	Composition, Relative[a]	Carboxyl-Terminal[b]
Cm-Cys	1 ± 0.1	
Asp	0 ± 0.3	
Thr	0 ± 0.1	
Ser	1 ± 0.1	0.1
Gln		0.7
Glu	1 ± 0.1	
Gly	0 ± 0.3	
Ala	1 ± 0.2	0.1
Val	0 ± 0.3	
Ile	1 ± 0.1	0.5
Leu	0 ± 0.4	
Phe	0 ± 0.1	
Lys	1 ± 0.1	1.0
Arg	0 ± 0.1	

[a] HCl hydrolysate; results are nearest integer, corrected for 0.02 mol/mol contamination with the large peptide of reduced T-231. [b] Aliquots, 0.05 μmol, of peptide were treated with 90 μg of carboxypeptidase A and 60 μg of carboxypeptidase B for 0 and 1.2 hr at 37°, and the mixture was analyzed directly, using lithium buffers (Vega and Nunn, 1969) to permit isolation of glutamine; values for the 0-hr specimen were subtracted from those for the 1.2-hr specimen.

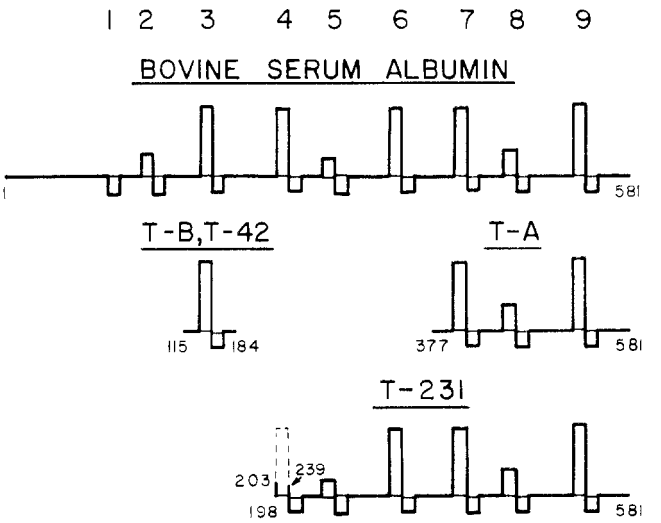

FIGURE 8: Schematic diagram of the multi-loop structure of bovine serum albumin proposed by Brown (1975), top, and, below, the proposed location of the tryptic peptides described herein. The peptide chain is shown as a heavy line and disulfide bonds as light horizontal lines. The disulfide-bridged loops are numbered 1–9. Peptides T-B and T-42 are identical, and constitute the third disulfide loop. Peptide T-A contains loops 7–9. Peptide T-231 contains loops 4–9, but is missing the tryptophan-containing segment 204–238, shown as a broken line.

one not represented in human albumin (Behrens et al., 1975). (The tryptophan common to both human and bovine albumins is situated in loop 4.) Among the 70 residues of this fragment are six tyrosines in addition to the tryptophan, causing the specific absorbance at 280 nm (Table II) to be nearly triple that of intact bovine serum albumin. This aromatic content and the resultant hydrophobic nature may cause this region of the molecule to have a conformation more resistant to tryptic cleavage than adjacent segments, and thus may account for the observed recovery of this peptide from both types of digests.

The larger fragment isolated by the combination of affinity chromatography and proteolysis is $T_{377-581}$, consisting of loops 7–9 of the albumin model, or approximately the carboxyl-terminal third of the molecule (Figure 8). The protection that binding to the fatty acid affords to this segment implies the presence of a strong fatty acid binding site in this region, as well as a probable conformational change upon binding the long-chain fatty acid. In contrast to $T_{115-184}$, $T_{377-581}$ is low in aromatic residues, and its specific absorbance at 280 nm is only half that of bovine serum albumin.

The third fragment reported is a subfraction isolated from our preparation of a fraction which corresponds to King and Spencer's (1970) fragment T-23. Termed T-231, it apparently consists of loops 4–9 of the albumin molecule, but lacks the major tryptophan-containing segment 204–238 of the longer arm of double loop 4 (Figure 8). This complex peptide is probably only one of several similar structures resulting from the tryptic digestion of soluble albumin. The missing segment of loop 4 contains nine trypsin-susceptible bonds, various pairs of which could be split to produce fragments of molecular weight about 39,000 which contain no tryptophan. One subfraction of T-23, for instance, showed on Edman degradation a sequence compatible with cleavage at such a site, -Arg-Leu-$_{217}$.

No reports other than that of King and Spencer (1970) are available on fragments isolated from bovine albumin following tryptic digestion, but several groups have studied the action of trypsin on human albumin. Štokrová and Keil (1963) noted the occurrence of three intermediate forms slightly smaller than intact albumin. Markus et al. (1967b) found fragments of molecular weight 51,000, 18,500, and 13,500, and noted that there was no significant difference in the size distribution of intermediates from human and bovine albumins. Kaminski and Tanner (1959) and Ivanyi (1963) studied the immunological reactivity of several tryptic peptides. Lapresle and Goldstein (1969) isolated a fragment, F1, of 6600 molecular weight containing one of the antigenic sites of the human albumin molecule which cross-reacts with a site in bovine serum albumin. Correlation of any of these fragments with the fragments of the present work is not apparent.

The capability to locate precisely the three tryptic bovine serum albumin fragments described in this paper within the multi-loop structure of bovine serum albumin of Brown (1974,1975) helps to confirm the structure he proposed and, to a more limited degree, his proposed sequence (Brown, 1975). Peptide$_{115-184}$ is the second of the large double loops to be isolated. The first was the carboxyl-terminal "Phe" fragment, loop 9, or residues 504–581 (Peters and Hawn, 1967), which Brown (1974,1975) used as a model in formulating the double-loop structure.

Acknowledgment

Part of this work was performed while one of the authors (T.P.) was a Guest Worker at the Laboratory for Chemical Biology, National Institute for Arthritis, Metabolism and Digestive Disease, National Institutes of Health. He expresses his gratitude to Drs. C. B. Anfinsen and H. Taniuchi for their hospitality and for many helpful discussions during this period. The authors are also grateful to Dr. James R. Brown for personal communications concerning the bovine serum albumin sequence.

References

Andrews, P. (1970), *Methods Biochem. Anal. 18*, 1.
Behrens, P. Q., Spiekerman, A. M., and Brown, J. R.

(1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 591.
Benjamin, D. C., and Weigle, W. O. (1971), *Immunochemistry 8,* 1087.
Brown, J. R. (1974), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 33,* 1389.
Brown, J. R. (1975), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 34,* 591.
Cohn, E. J., Hughes, W. L., Jr., and Weare, J. H. (1947), *J. Am. Chem. Soc. 69,* 1753.
Doumas, B. T., Watson, W. A., and Biggs, H. G. (1971), *Clin. Chim. Acta 31,* 87.
Fraenkel-Conrat, H., Harris, J. I., and Levy, A. L. (1955), *Methods Bioch. Anal. 2,* 359.
Ivanyi, J. (1963), *Folia Biol. 9,* 222.
Kaminski, M., and Tanner, C. E. (1959), *Biochim. Biophys. Acta 33,* 10.
King, T. P., and Spencer, M. (1970), *J. Biol. Chem. 245,* 6134.
Lapresle, C., and Goldstein, I. J. (1969), *J. Immunol. 102,* 733.
Markus, G., McClintock, D. K., and Castellani, B. A. (1967a), *J. Biol. Chem. 242,* 4402.
Markus, G., McClintock, D. K., and Castellani, B. A. (1967b), *J. Biol. Chem. 242,* 4395.
McMenamy, R. H. (1971), *J. Biol. Chem. 246,* 4744.
Niu, C., and Fraenkel-Conrat, H. (1955), *J. Am. Chem. Soc. 77,* 5882.
Peters, T., Jr., and Hawn, C. (1967), *J. Biol. Chem. 242,* 1566.
Peters, T., Jr., Taniuchi, H., and Anfinsen, C. B. (1973), *J. Biol. Chem. 248,* 2447.
Scheider, W., and Fuller, J. K. (1970), *Biochim. Biophys. Acta 221,* 376.
Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971), *Biochemistry 10,* 4912.
Spies, J. R. (1967), *Anal. Chem. 39,* 1412.
Štokrová, S., and Keil, B. (1963), *Collect. Czech. Chem. Commun. 28,* 2846.
Taniuchi, H., and Anfinsen, C. B. (1966), *J. Biol. Chem. 241,* 4366.
Vega, A., and Nunn, P. B. (1969), *Anal. Biochem. 32,* 446.
Weber, K., Pringle, P. R., and Osborn, M. (1972), *Methods Enzymol. 26C,* 3.

Ionic Strength Induced Structure in Histone H4 and Its Fragments†

Peter N. Lewis,‡ E. Morton Bradbury, and Colyn Crane-Robinson*

ABSTRACT: The salt-induced folding and self-association of histone H4 and its fragments (1–23), (25–67), (69–84), and (69–102) have been studied at the same molar concentration (1 m*M*) by nuclear magnetic resonance (NMR), circular dichroism (CD), and ir spectroscopy. By each of these techniques intact histone H4 exhibited a fast structural change, involving the formation of α_R helix and aggregation, and also a slow change involving β-structure formation. Fragment (25–67) was found to behave in a manner similar to the intact molecule for the fast change, showing both helix formation and aggregation but exhibited no time-dependent effects. All the other fragments were found to be essentially noninteracting. It is concluded that (25–67) contains the region critical for the folding and self-association of histone H4. On the basis of these results a model is proposed for the self-association of histone H4 in which helix is located between residues 49 and 73, while the β structure lies between 74 and the C-terminus.

One of the most striking features of histone H4[1] (F2A1, IV, or GRK) is the very high degree of conservation of its amino acid sequence from species to species in eukaryotes (DeLange and Smith, 1971). Although this near invariance undoubtedly results from an evolutionary need to preserve specific and fundamental interactions in which this histone participates in vivo, physical and biological studies have yet to discover the detailed nature of the uniqueness of these interactions.

The amino acid sequence of histone H4 as determined by DeLange et al. (1969) and Ogawa et al. (1969) displays an asymmetric distribution of basic and apolar residues. Consequently, it has been widely proposed that the most basic portion of histone H4 binds to the DNA in chromatin while the least basic portion (the apolar segment) is free to participate in protein–protein interactions.

Boublik et al. (1970) have shown by nuclear magnetic resonance (NMR) measurements that when the ionic strength of aqueous solutions of histone H4 is increased an aggregation phenomenon occurs involving only a portion of the molecule that corresponds to the apolar segment. Boublik et al. (1970) from optical rotatory dispersion (ORD)

† From the Department of Physics, Biophysics Laboratory, Portsmouth Polytechnic, Portsmouth, England. *Received August 1, 1974.* This research was supported in part by the Science Research Council of Great Britain and in part by a National Research Council of Canada Postdoctoral Fellowship to P.N.L.

‡ Present address: Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada.

[1] The new histone nomenclature used here was accepted by the participants at the CIBA Foundation Symposium on the Structure and Function of Chromatin, April 3–5, 1974. This new nomenclature which has been proposed to the appropriate international nomenclature committee is as follows for each histone where the previous names are given in parentheses: H1 (F1, I, KAP), H2A (F2A2, I1b1, ALK), H2B (F2B, I1b2, KSA), H3 (F3, III, ARK), H4 (F2A1, IV, GRK), and H5 (F2C, V, KAS).