Positions in Human Serum Albumin which Involve the Indole Binding Site

SEQUENCE OF 107-RESIDUE FRAGMENT*

KANWAL K. GAMBHIR,† RAPIER H. McMENAMY,§ AND FRANK WATSON

From the Department of Biochemistry, The Medical and Dental Schools, State University of New York at Buffalo, Buffalo, New York 14214

SUMMARY

The first 107 residues of Fragment C of human serum albumin have been sequenced and two positions at which affinity labels block the indole site determined. Histidine 23 is the position of blockage by bromoacetyl-L-tryptophan and lysine 67 is the position of blockage by 5-dimethylaminonaphthalene-1-sulfonyl chloride and probably pyridoxal 5'-phosphate. The presence of an indole ligand at the binding site markedly reduces incorporation of the label into the above lysyl residue, and in the case of 5-dimethylaminonaphthalene-1-sulfonyl chloride, increases incorporation into three other positions, lysine residues 13, 39, and 84. It is concluded that binding of the indole ligand on the site brings about conformational changes in the albumin structure exposing new reactive positions for 5-dimethylaminonaphthalene-1-sulfonyl chloride. There is a large accumulation of basic and hydrophobic residues and no glycine, serine, threonine, valine, aspartate, or cysteine residues in the sequence 10 to 43. Lysine 71 has been identified by amino acid analyses and sequence studies as the position acetylated by acetylsalicylic acid (HAWKINS, D. R., PINCKARD, N., CRAWFORD, C. P., AND FARR, R. S. J. Clin. Invest. (1969) 48, 536), establishing the structural relationships of two major ligand binding sites on albumin. The lone tryptophan is at position 86. Evidence indicates that within residues 1 to 86 of Fragment C and within residues of the A-Phe fragment ($M_r = 10,000$), the latter known to be adjacent to Fragment C in the whole albumin structure, exists the major binding sites of all ligands for human serum albumin.

It was found earlier that reacting human serum albumin with affinity labels bromoacetyl-L-tryptophan, dansyl$^1$ chloride, or pyridoxal 5'-phosphate, led to a blockage of the binding of acetyl-L-tryptophan (1). It was further found that the greater part of each label was incorporated into Fragment C ($M_r = 18,000$). Thus, at conditions of 1:1 molar ratio of labeling agent to albumin 65 to 95% of the label was incorporated into Fragment C, accounting for essentially an equivalent stoichiometric blocking of the binding site when bromoacetyl-L-tryptophan was used as a reagent, and 35 to 50% of the equivalent stoichiometric blocking when dansyl chloride or pyridoxal 5'-phosphate were used as the labeling reagents. If labeling was conducted in the presence of indoleproprionate, a ligand strongly bound by albumin, the indole site was largely protected. Bromoacetyl-L-tryptophan reacted with imidazole groups, dansyl chloride reacted with e-amino lysyl(s) and tyrosyl-OH group(s), and pyridoxal 5'-phosphate reacted with e-amino lysyl group(s).

The indole binding site on albumin is unusual in several aspects. First, the above described labeling agents which block the indole site all have diverse structural features. The site, thus, is apparently readily adaptable to different types of ligands. A further example of this adaptability is shown with thyroxine, a compound much different from the indole compounds in shape and structural groups, yet, as shown by Tritsch (2), binding competitively with L-tryptophan. Another unusual feature of the site is that fatty acids at low concentrations inhibit the indole site in the alkaline pH region, but have little effect on indole binding in the neutral and acid pH regions (3). Fatty acid inhibition at low concentration is not directly competitive with indole binding. In addition to providing a broad accommodation to different types of ligand, there is also evidence that occupation of the site by an inducible compound provides considerable protection to the albumin. Thus, the reaction of trypsin with albumin is much retarded when albumin is associated with L-tryptophan (4). Also, acetyl-L-tryptophan has, for many years, been known to stabilize albumin against thermal denaturation. Finally, it has recently been postu-

* This investigation was supported by National Science Foundation Research Grant GB-7224 and by a grant from the James H. Cummings Foundation. An initial report was presented at the 57th Annual Meeting of the Federation of American Societies for Experimental Biology, April, 1973. Part of the data are taken from a thesis submitted by K. K. Gambhir for his Ph D degree.
† Present address, Department of Obstetrics, Gynecology and Medicine, Howard University, Washington, D. C. 20001.
§ From whom reprints should be requested.

1 The abbreviations used are: dansyl, the prefix for 5-dimethylaminonaphthalene-1-sulfonyl group; TPCK, L-tosylamide-2-phenylethyl chloromethyl ketone; acetyl-L-tryptophan-HSA, albumin reacted with bromoacetyl-L-tryptophan; dansyl-HSA, albumin reacted with dansyl chloride; pyridoxal-P-HSA, albumin reacted with pyridoxal 5'-phosphate and reduced with sodium borohydride; PTH, phenylthioblydantoin.
lated that the indole binding site plays a role in neural regulation (5). For these reasons, it is of much interest to study further the structure and other features of the indole binding site.

In the current study, a major part of Fragment C has been sequenced, and the positions of the labeled residues determined. Information is provided on several regions of the peptide chain which are involved in the indole binding site. The natures of the side groups of the residues near the binding region, which likely supply the forces for binding, are considered.

**Experimental Procedure**

The source of human serum albumin, affinity labeling conditions, fragment isolation, and electrophoretic (Acryl Corp.) C were as described previously (1). Affinity labels bromoacetyl-L-tryptophan, daniyl chloride, and pyridoxal 5'-phosphate, were reacted with protein at 1:1 molar ratios in the presence and absence of indole-propionate. Binding studies were conducted to evaluate the extent to which the affinity labels blocked the site, and the albumin preparations were then cleaved to isolate Fragment C.

**Reaction of Protein with Affinity Labels**

The reaction was carried out at 5°C for 10 min. The reaction was then stopped by the addition of carrier protein. The reaction products were separated from the reagents by passing the solutions through a Sephadex G-25 (coarse) column equilibrated with 0.01% triethylpropionate. Binding studies were conducted to evaluate the extent to which the affinity labels blocked the site, and the albumin preparations were then cleaved to isolate Fragment C.

**Performic Acid Oxidation**

At -5°C, Fragment C (75 to 100 mg) was dissolved in a solution containing 4.5 ml of twice recrystallized formic acid and 0.5 ml of methanol. A performic acid solution, prepared by allowing to stand 1 hour at room temperature a mixture of 5 ml of formic acid and 0.5 ml of 30% hydrogen peroxide (Mallinckrodt reagent grade) and 0.01 ml of liquid phenol (Mallinckrodt reagent grade) were used to react with the freed -SH groups, providing S-carboxymethylated and S-B(4-pyridyl)ethylated C fragments, respectively. The peptide was separated from the reagents by passing the solutions through a Sephadex G-25 (coarse) column equilibrated with 0.01% triethylpropionate. Binding studies were conducted to evaluate the extent to which the affinity labels blocked the site, and the albumin preparations were then cleaved to isolate Fragment C.

**Maleylation—**Reduce or oxidized Fragment C was dissolved in 5 ml of water, its container placed in an ice bath and, with the pH maintained at 8.5 to 9 by addition of 2 n NaOH, 2 parts of maleic anhydride (w/w) were added in small aliquots over a 15-min period. After standing for an additional 30 min, the reaction solution was diluted three times to 56 ml with water, and concentrated each time to approximately 5 ml in the indole apparatus.

**Trypsin Reaction**

To assure more complete destruction of chymotryptic activity, trypsin (TPCK-treated, Worthington) was added to 2.5 mg of the trypsin in 0.25 ml of 0.2 N NaHCO3. After 3 hours, the pH was reduced to a value of 3.0 with 0.1 N HCl. The enzyme solution was used within 16 hours.

To a 3 ml solution of reduced or oxidized, maleylated Fragment C (75 to 100 mg), after flushing with N2 and pH adjustment to 9.0, was added 0.1 ml of TPCK-treated trypsin solution. The pH was maintained at 9.0 by addition of 0.1 N NaOH for 5 min. A second aliquot of trypsin solution was added, and the pH again maintained at 9.0 by addition of NaOH solution (the volume of NaOH solution added after the second addition of enzyme solution was usually very small). After 5 min, with good mixing, 2.5 µl of disopropylphosphorofluoridate were added and allowed to react for 20 min. The solution was placed on a Sephadex G-100 column (5 X 210 cm), and eluted with a solution containing 0.1 N NH4HCO3 and 0.01 N NH3OH.

With maleylated peptides, the trypsin reactions were carried out under essentially the same conditions, except that the reaction times were extended 2 to 5 hours.

**Chymotrypsin Reaction**

Reactions with maleylated Fragment C were carried out at pH 8.0 by two additions of chymotrypsin (Worthington) at two 20-min intervals (1.100, w/w). At the end of the reaction, 50 µl of phenylmethylsulfonylfluoride solution (2 mg freshly added to 1 ml of methanol) were added. After 20 min, the digested peptide was placed on Sephadex G-100, and eluted with a solution of 0.1 N NH4HCO3 and 0.01 N NH3OH.

**Demaleylation—**After the enzyme digestions and separation of the fragments, demaleylation was conducted at pH 2.5 to 3.0 for 40 hours at 45°C.

**α-Protease Reaction**

At a 1:50 ratio (w/w), pH 8.0, 0.01 m CaCl2, α-protease (Pierce Chemical Co.) was reacted with the peptides for 2 hours. The pH of the solution was then reduced to 3.0.

**Separation of Peptides—**Fractionation of peptides was conducted on a column (0.9 X 15 cm) of Dowex 50-X8 (Spincro resin AA15) using volatile buffers following the general methods described by Hill and Delaney (7) and by paper chromatography and electrophoresis. In the separation of dansyl peptides, the strategy was first to find a system in which the dansyl peptide could be separated from other peptides by one dimensional movement on paper. After the separation and before the paper was completely dry (dansyl zones elute very poorly if allowed to dry on the chromatograms), the zone was cut out and eluted by centrifugation. The apparatus for the latter consisted of a 12 ml centrifuge tube with a serated conical disc fitted approximately 5 cm from the bottom of the tube. The paper with the dansyl zone was placed in the tube with a small amount of 20% pyridine, allowed to stand several minutes, briefly centrifuged, and the extraction step repeated. After several extractions with the pyridine solution, two or more extractions were then carried out with 10% acetic acid.

**Reaction with N-Bromosuccinimide—**Fragment C (20 mg) was dissolved in 1 ml of H2O, the container placed in an ice bath, and 20 µl of a 3% (v/v) solution of N-Bromosuccinimide solution (in 1N NaOH) added. After 30 min, the pH was reduced to 4.5, and the solution allowed to stand 2 hours. Twice recrystallized urea (1 g) was added to the solution, the pH readjusted to 4.5, and 120 µl of N-Bromosuccinimide solution (0.2 g/20 ml of H2O) added. After 5 min, the pH of the solution was raised to 9.0, and the solution placed on a column (2 X 30 cm) of Sephadex G-25 (coarse) and eluted with 0.005 N NH4OH. The protein solution obtained from the column was freeze-dried and oxidized with performic acid. The methods of Gray (9) and Sauer et al. (10) were used for manual sequencing.

**Amino Acid Analyses—**Analyses of the normal and modified amino acids were conducted on a Beckman 120 B amino acid analyzer using a two column methodology as described before (1). The analysis was conducted at 110°C, 22 hours in 6 N HCl or 4 N methanolic acid containing 0.2% 3-(2-aminoethyl)indole (Pierce Chemical Co.). In each instance, aliquots of the hydrolyzates were diluted 1:30 with water, and the solutions placed directly on the resin columns.

**Paper Chromatography—**Large peptides were analyzed on a Beckman Sequencer (model 890C), modified with the optimal accelerated sequencing kit and N-flushing system. The Quadrol program recommended by the manufacturer was used. All solvents and reagents were from Pierce Chemical Co. For peptides less than 50 amino acid residues, 4-sulfophenylisothiocyanate was generally used to assist in keeping the peptide in the sequencer cup (8). With this reagent was carried out directly in the cup. After the peptide was dried, R1 and R3 were placed in the manual positions, Quadrol buffer was added to the 10-mm mark (low speed), the cup was stopped, opened, and 25 mg of 4-sulfophenylisothiocyanate was placed in the center of the cup. The cup was started, and when up to speed, the cover was replaced. With the effluent to waste, N2 was flushed through the cell for 1 min. The cup was then set at high speed and N2 was manually added to the 24-mm mark. The reaction was then started at Step 12 on the program and allowed to proceed to the end of the drying step after the ethyl acetate wash (Step 36 on the program). The program was then restarted from Step 1 with R1 and R3 in their automatic positions.

**Sequencing—**Large peptides were analyzed on a Beckman Sequencer (model 890C), modified with the optimal accelerated sequencing kit and N-flushing system. The Quadrol program recommended by the manufacturer was used. All solvents and reagents were from Pierce Chemical Co. For peptides less than 50 amino acid residues, 4-sulfophenylisothiocyanate was generally used to assist in keeping the peptide in the sequencer cup (8). With this reagent was carried out directly in the cup. After the peptide was dried, R1 and R3 were placed in the manual positions, Quadrol buffer was added to the 10-mm mark (low speed), the cup was stopped, opened, and 25 mg of 4-sulfophenylisothiocyanate was placed in the center of the cup. The cup was started, and when up to speed, the cover was replaced. With the effluent to waste, N2 was flushed through the cell for 1 min. The cup was then set at high speed and N2 was manually added to the 24-mm mark. The reaction was then started at Step 12 on the program and allowed to proceed to the end of the drying step after the ethyl acetate wash (Step 36 on the program). The program was then restarted from Step 1 with R1 and R3 in their automatic positions.
phenylalanine, lysine, glutamate, serine; and valine, alanine, lysine, threonine, glutamine. The standards, applied at 2.5 nmol, were alternated with unknowns in applications across the plate. Repetitive aliquots (~1 μl) of unknowns were applied at the same spot until the fluorescent intensity reached approximately that of the standards. The amount applied was recorded. Eight unknowns were usually run on each gel plate. Ascending development was carried out in Solvent V (heptane 58 ml, pro- pionic acid 17 ml, and ethylene chloride 25 ml) of Jeppsson and Sjoquist (11). The dried chromatograms were first viewed under ultraviolet light to identify as many zones as possible, then sprayed with a 1-butanol solution containing 0.5 g/100 ml of ninhydrin followed by heating of the plates in an oven at 110° for 10 min. This latter treatment assisted further identification of some of the zones by color differences (12). Overlaps were easily followed by this method. Sequencing experiments were terminated when the background became too high to recognize the un- known zones. When leucine, isoleucine, and sometimes valine and proline were indicated, their identification was confirmed by gas chromatography using the method of Pisano et al. (13) as indicated in the Sequencer Operator Manual. PTH derivatives of arginine, histidine, and cysteic acid, and sometimes lysine (when 4-sulfo- phenylisothiocyanate was used), were usually identified by amino acid analysis after hydrolysis with 5N HCl in sealed, evacuated tubes at 140° for 24 hours. In some instances, arginine was identi- fied by phenanthrenequinone (14) and histidine by p-anisidine (15) as described by Herronson et al. (6).

**RESULTS**

**Peptide Separations**—The elution profile of trypsin-digested, maleylated carboxymethylated Fragment C is shown in Fig. 1. Zones a, b, and c were found to contain single peptides with NH2-terminal residues, Phe, Tyr, and Cys. Amino acid analyses are given in Table I. Zone b appeared to be a mixture of two fragments; however, fractions taken from the leading and trailing shoulders had identical NH2-terminal groups, and upon hydrolysis, had identical amino acid contents. The reason for the evident splitting of this zone was not clear. The relative amounts in the split zones also varied; in some instances, the leading peak was much reduced. Zone d was freeze-dried, demaleylated, and chro- matographed on the peptide resin (Fig. 2). Amino acid analyses of Zones d1, d2, and d3 are reported in Table I. The other zones in this elution contained multiple peptides at low concentrations and were not further considered.

The elution profile of trypsin-digested, maleylated S-β(4-pyridyl)ethylated Fragment C is given in Fig. 3. When compared with Fig. 1, a new zone appeared between Zones c and d, Zone e. This zone was further purified by concentration and passing the second time over Sephadex, G-100, or by demaleylating and chromatographing on the peptide resin (Fig. 4). Comparison of the amino acid composition of the peptides in Zone b and c from the S-β(4-pyridyl)ethylated Fragment C digest with the peptide from b of the S-carboxymethylated Fragment C digest indicated that the latter was equal to the sum of the former two (Table II). Splitting the peptide b at its internal arginine apparently is in-hibited by S-carboxymethylene at cysteine residues. Splitting readily occurred with S-β(4-pyridyl)ethylated and performic acid-oxidized Fragment C preparations.

The fragments isolated accounted for all of the eight arginines in Fragment C. Peptide a, which contained a homoserine residue, and was, thus, the COOH-terminal peptide, contained an internal integer

![Fig. 1. Elution profile of tryptic digest of carboxymethylated Fragment C (100 mg). Column Sephadex G-100 (5 x 210 cm), eluting buffer 0.1 N NH4HCO3/0.01 N NH3OH.](image_url)
argine which did not cleave with trypsin under any conditions attempted.

The zone of the void volume from the elution of chymotrypsin-digested, maleylated, carboxymethylated Fragment C (column conditions the same as described for Fig. 1) was demaleylated and chromatographed on the peptide resin (column and gradient conditions the same as described for Fig. 4). The amino acid composition and eluting fractions for some of the peptides isolated are given in Table III. The amino acid composition and eluting fractions for two peptides isolated from an α-protease digestion of peptide b, using the same peptide resin conditions are reported in Table IV.

**Sequences**—The results of sequencing experiments are given in Fig. 5. Intact Fragment C (500 nmol) was subjected to 42 steps of automated Edman degradation. Several of the mixed residues were later obtained from chymotryptic or α-protease-cleaved peptides (Tables III and IV). Thus, residue 19 was determined to be isoleucine from the composition of peptide 30; residue 30 was determined to be glutamate from the composition of peptide 23; and residue 36 was determined to be lysine by isolation of peptide 31. Automatic sequencing of peptide b, (300 nmol) provided residues 38 to 63, and an overlap with the sequence obtained with the intact Fragment C. No PTH product was obtained for residue 26.

54. From consideration of the amino acid composition in peptide b, serine was assigned to this position. There was a large drop in the yield at aspartate 63. α-Protease digestion of the b, fragment produced a peptide with the composition aspartate, glycine, serine, alanine, glutamate, lysine, and arginine; and another with the amino acid composition glutamate, lysine, and arginine (Ta-
Neither of these peptides would sequence by the Edman procedure, presumably because blocking of the NH₂-terminal residues had occurred in their isolation. Carboxypeptidase B reacted with b₉ producing arginine. Carboxypeptidase A subsequently released glutamine. A mixture of carboxypeptidases A and B released arginine, glutamine, lysine, and alanine with a slow release of serine. On this basis, the order of assigned residues 64 to 69 seems reasonably correct. No peptide was isolated to establish evidence of overlap between residues 69 and 70. However, all other overlaps of arginine-cleaved peptides were accounted for, leading one to conclude that this assignment is correct.

Table IV

Amino acid composition (nmol) of some α-protease cleaved peptides from peptide b₉

|    | b₉a | b₉d |
|----|-----|-----|
| Lys | 81  | 9   |
| Arg | 60  | 10  |
| Asp | 66  | 1   |
| Ser | 76  | 1   |
| Glu | 60  | 1   |
| Gly | 56  | 1   |
| Ala | 62  | 1   |

Eluting fractions 40-43 125-133

Location of Affinity Labels in Fragment C—Combined tryptic and chymotryptic digestion of acetyl-L-tryptophan-albumin (labeled at 1:1 molar ratio of bromoacetyl-L-tryptophan to albumin) produced peptide 1 (Fig. 6) with the composition on hydrolysis of arginine, 3-carboxymethylhistidine, proline, and tyrosine. When compared with the total ³H label in albumin, 80% or more could be accounted for in this peptide zone. No carboxymethylhistidine (or N'-carboxymethyllysine for that matter) was found on hydrolysis of any of the other zones. Upon sequencing, the following residues were obtained: Arg—Pro-Tyr. The acetyl-tryptophan-labeled histidine was assigned to the second residue.

Fig. 5. Sequence summary of first 107 residues in Fragment C. Peptides isolated by chymotrypsin or α-protease digestions are indicated by a. Arrows (→) indicate direction from which sequence was conducted.
dized, maleylated, dansyl-Fragment C (albumin dansylated at a
in the latter permit the label to be present only at residue 23.
found to be located in Zone d. The composition of the peptides
label, after maleylation and tryptic digestion of Fragment C, was
ment C (albumin reacted with pyridoxal 5'-phosphate at a molar
molar ratio of 1:1) and a similar treatment of pyridosal-P-Frag-
this residue was found to be histidine 23. This position was further confirmed when the
composition of residues 63-69; h, identified as lysine by isolation
fication via PTH derivative; m, by difference from amino acid analysis of peptide //; n, these
residues were identified from two different peptides.
On comparison with the sequence in Fig. 4, this residue was found
to be histidine 23. This position was further confirmed when the
label, after maleylation and tryptic digestion of Fragment C, was
found to be located in Zone d. The composition of the peptides
in the latter permit the label to be present only at residue 23.

Chromatography of the tryptic digest of performic acid-oxi-
dized, maleylated, dansyl-Fragment C (albumin dansylated at a molar ratio of 1:1) and a similar treatment of pyridoxal-P-Frag-
ment C (albumin reacted with pyridoxal 5'-phosphate at a molar
ratio of 1:1) produced the results in Fig. 7. These labels were
clearly shown to be concentrated in peptide b. The dansyl content
was estimated as 0.25 mol/mol of peptide b. Amino acid analyses demonstrated that the labels were present only in
N'-lysyl derivatives.

Reduced S-β(4-pyridyl)ethylated Fragment C from albumin
after labeling with dansyl chloride in the presence of indolepropi-
one was also maleylated and reacted with trypsin. The profile

### TABLE V

| Residue identification summary |
|--------------------------------|
| 1  Cys<sup>a</sup>  E 28  Ala<sup>b</sup>  TLC 55  Leu<sup>b</sup>  TLC<sup>d</sup>  82<sup>n</sup>  Ala<sup>a</sup>  TLC |
| 2  Thr<sup>b</sup>  TLC 29  Pro<sup>b</sup>  TLC 56  Leu<sup>b</sup>  TLC<sup>d</sup>  83<sup>n</sup>  Phe<sup>a</sup>  TLC |
| 3  Ala<sup>b</sup>  TLC 30  Glu<sup>j</sup>  AA 57  Pro<sup>b</sup>  TLC<sup>d</sup>  84  Lys<sup>a</sup>  TLC |
| 4  Phe<sup>b</sup>  TLC 31  Thr<sup>b</sup>  TLC<sup>d</sup>  58  Lys<sup>b</sup>  HI 85  Ala<sup>a</sup>  TLC |
| 5  His<sup>b</sup>  Color<sup>c</sup>  32  Leu<sup>b</sup>  TLC<sup>d</sup>  59  Leu<sup>b</sup>  TLC<sup>d</sup>  86  Trp<sup>a</sup>  TLC |
| 6  Asp<sup>b</sup>  TLC 33  Phe<sup>b</sup>  TLC 60  Asp<sup>b</sup>  TLC<sup>d</sup>  87<sup>n</sup>  Ala<sup>b</sup>  TLC |
| 7  Asn<sup>b</sup>  TLC 34  Phe<sup>b</sup>  TLC 61  Glu<sup>b</sup>  TLC 88<sup>n</sup>  Val<sup>b</sup>  TLC |
| 8  Gin<sup>b</sup>  TLC 35  Ala<sup>b</sup>  TLC 62  Leu<sup>b</sup>  TLC<sup>d</sup>  89<sup>n</sup>  Ala<sup>b</sup>  TLC |
| 9  Glu<sup>b</sup>  TLC 36  Lys<sup>b</sup>  AA 63  Asp<sup>b</sup>  TLC 90<sup>n</sup>  Arg<sup>b</sup>  Color<sup>f</sup> |
| 10  Thr<sup>b</sup>  TLC 37  Arg<sup>b</sup>  HI 64  Gly<sup>j</sup>  AA 91<sup>n</sup>  Leu<sup>b</sup>  TLC<sup>d</sup> |
| 11  Phe<sup>b</sup>  TLC 38  Tyr<sup>b</sup>  TLC 65  Ser<sup>k</sup>  AA 92<sup>n</sup>  Ser<sup>b</sup>  TLC |
| 12  Leu<sup>b</sup>  TLC<sup>d</sup>  39  Lys<sup>b</sup>  HI 66  Ala<sup>k</sup>  AA 93<sup>n</sup>  Glu<sup>b</sup>  TLC |
| 13  Lys<sup>b</sup>  TLC 40  Ala<sup>b</sup>  TLC 67  Lys<sup>k</sup>  AA 94<sup>n</sup>  Arg<sup>b</sup>  Color<sup>f</sup> |
| 14  Lys<sup>b</sup>  TLC 41  Ala<sup>b</sup>  TLC 68  Glu<sup>k</sup>  AA 95  Phe<sup>b</sup>  TLC |
| 15  Tyr<sup>b</sup>  TLC 42  Phe<sup>b</sup>  TLC 69  Arg<sup>k</sup>  AA 96  Pro<sup>b</sup>  TLC<sup>d</sup> |
| 16  Leu<sup>b</sup>  TLC<sup>d</sup>  43  Thr<sup>b</sup>  TLC 70  Leu<sup>L</sup>  TLC<sup>d</sup>  97  Lys<sup>b</sup>  TLC |
| 17  Tyr<sup>b</sup>  TLC 44  Glu<sup>b</sup>  TLC 71  Lys<sup>L</sup>  TLC 98  Ala<sup>b</sup>  TLC |
| 18  Glu<sup>b</sup>  TLC 45  Cys<sup>b</sup>  HI 72  Cys<sup>L</sup>  HI 99  Glu<sup>b</sup>  TLC |
| 19  Ile<sup>e</sup>  AA 46  Cys<sup>b</sup>  HI 75  Ala<sup>L</sup>  TLC 100  Phe<sup>b</sup>  TLC |
| 20  Ala<sup>b</sup>  TLC 47  Glu<sup>b</sup>  TLC 74  Ser<sup>m</sup>  AA 101  Ala<sup>b</sup>  TLC |
| 21  Arg<sup>b</sup>  Color<sup>f</sup>  48  Ala<sup>b</sup>  TLC 75  Leu<sup>L</sup>  TLC<sup>d</sup>  102  Glu<sup>b</sup>  TLC |
| 22  Arg<sup>b</sup>  Color<sup>f</sup>  49  Ala<sup>b</sup>  TLC 76  Glu<sup>L</sup>  TLC 103  Val<sup>b</sup>  TLC<sup>d</sup> |
| 23  His<sup>b</sup>  Color<sup>c</sup>  50  Asp<sup>b</sup>  TLC 77  Lys<sup>L</sup>  TLC 104  Lys<sup>b</sup>  TLC |
| 24  Pro<sup>b</sup>  TLC<sup>d</sup>  51  Lys<sup>b</sup>  TLC 78  Phe<sup>L</sup>  TLC 105  Ala<sup>b</sup>  TLC |
| 25  Tyr<sup>b</sup>  TLC 52  Ala<sup>b</sup>  TLC 79<sup>n</sup>  Gly<sup>L</sup>  TLC 106  Leu<sup>b</sup>  TLC<sup>d</sup> |
| 26  Phe<sup>b</sup>  TLC 53  Ala<sup>b</sup>  TLC 80<sup>n</sup>  Glu<sup>L</sup>  TLC 107  Val<sup>b</sup>  TLC<sup>d</sup> |
| 27  Tyr<sup>b</sup>  TLC 54  Ser<sup>k</sup>  AA 81<sup>n</sup>  Arg<sup>b</sup>  HI 110  Leu<sup>b</sup>  TLC<sup>c</sup> |

Abbreviations: E, flat bed electrophoresis; TLC, thin layer chromatography; AA, amino acid analysis of a peptide residue; HI, PTH derivative hydrolyzed and identification made by amino acid analysis.

a, manual-dansyl derivative; b, automatic sequenator; c, p-anisidine test; d, confirmed by gas chromatography; e, identified as either leucine or isoleucine by PTH derivative, confirmed as isoleucine by amino acid analysis of peptide containing residues 18-25; f, phenanthrenequinone test; g, identified as either glutamate or lysine by PTH derivative, confirmed as glutamate by amino acid analysis of peptide containing residues 28-33; h, identified as lysine by isolation of peptide with residues 35-38; i, obtained by difference from amino acid analysis of peptide b; j, obtained by difference from amino acid analysis of peptide bp and isolation of peptide with composition of residues 63-69; k, carboxypeptidase A and B; l, manual sequencing, identification via PTH derivative; m, by difference from amino acid analysis of peptide a; n, these residues were identified from two different peptides.
from acetyltryptophan-HSA, labeled at a 1:1 molar ratio of bromoacetyl-L-tryptophan to albumin. Gradient 250 ml of 0.2 pyridine-acetate, pH 3.1, 250 ml of 2.0 X 15 cm) of tryptic-chymotryptic digest of Fragment C obtained formic acid-oxidized C fragments obtained from pyridoxal-P-HSA and dansyl-HSA (conditions of labeling agent to HSA 1:1 molar ratios). Column and eluting conditions were the same as Fig. 3. Transmittance was the same for all digests. Solid bars tryptic digest of S-p(4-pyridyl)ethylated Fragment C obtained from albumin labeled in the presence of indolepropionate. Obtained was similar to that in Fig. 3. In this instance, the b, zone was clearly much less labeled than when labeling was conducted in the absence of indolepropionate (for comparative purposes these data are also indicated in Fig. 7). Furthermore, the label increased considerably in Zones c and d when indolepropionate was present, but was unchanged in Zone a.

The specific residues to which the dansyl labels were attached were identified when the labeling was conducted both in the presence and absence of indolepropionate. Trypsin digestion of peptide b, labeled with dansyl chloride 1:1 in the absence of indolepropionate) followed by descending chromatography in 1-butanol/glacial acetic acid/water, 4/1/5, v/v for 44 hours produced one strong fluorescent zone (~15 cm from origin) which on amino acid analysis was found to contain residues 50 to 69. This identified the major labeling position as lysine 67, identified as described above, and another which could not be identified by elution from the paper chromatogram. α-Protease digestion of peptide b, followed by resin column chromatography provided three zones containing dansyl residues. On amino acid analysis, they corresponded to residues 39 to 40, 63 to 68, and 63 to 67. Both lysine 39 and 67 were, therefore, dansylated when labeling was conducted in the presence of indolepropionate. Furthermore, lysine 39 was found to be dansylated only in the presence of indolepropionate. In view of the difficulties in elution of dansyl zones from the chromatogram, no quantitation was attempted. However, one should note that the total dansyl label in peptide b, was approximately twice as high when labeled in the absence of indolepropionate, compared to when indolepropionate was present. Labeling at residue 67, therefore, must be much reduced by the presence of indolepropionate occupying the indole binding site.

Tryptsin digestion of peptide c followed by chromatography and electrophoresis (pH 6.0) was analyzed for peptides with dansyl labels. Two dansyl zones, one very strong, and one of moderate fluorescent strength, appeared on a chromatogram developed 24 hours in the 1-butanol/glacial acetic acid/water solvent. The strongly fluorescent zone contained residues 82 to 90, the other weaker zone had an amino acid composition which did not correspond to any known peptide or amino acid sequence found in Fragment C (in nmol, lysine 6.6, proline 13.4, alanine 22, and phenylanine 12). The source of this peptide was not further investigated at this time.

When albumin was labeled with pyridoxal 5'-phosphate in the presence of indolepropionate, the amount of label incorporated into Fragment b, was approximately one-half that incorporated in the absence of indolepropionate. The behavior was similar to dansyl chloride in this respect. From analogy with the dansyl-labeling studies, the fact that pyridoxal 5'-phosphate was earlier found to react with a lysyl residue in albumin which inhibited acetyl-L-tryptophan binding, it is most likely that pyridoxal 5'-phosphate also reacts at lysine 67 in the same manner as dansyl chloride.

**FIG. 6. Chromatography profile for Dowex 50-X8 column (0.9 × 15 cm) of tryptic-chymotryptic digest of Fragment C obtained from acetyltryptophan-HSA, labeled at a 1:1 molar ratio of bromoacetyl-L-tryptophan to albumin. Gradient 250 ml of 0.2 pyridine-acetate, pH 3.1, 250 ml of 2.0 pyridine-acetate, pH 5.0.**

**FIG. 7. Elution profiles of tryptic digest of maleylated, performic acid-oxidized C fragments obtained from pyridoxal-P-HSA and dansyl-HSA (conditions of labeling agent to HSA 1:1 molar ratios). Column and eluting conditions were the same as Fig. 3. Transmittance was the same for all digests. Solid bars indicate the per cent of pyridoxal 5'-phosphate label per fragment; speckled bars indicate the per cent dansyl label per fragment; open bars show the per cent of dansyl label present per fragment in a tryptic digest of S-p(4-pyridyl)ethylated Fragment C obtained from albumin labeled in the presence of indolepropionate.**

obtained was similar to that in Fig. 3. In this instance, the b, zone was clearly much less labeled than when labeling was conducted in the absence of indolepropionate (for comparative purposes these data are also indicated in Fig. 7). Furthermore, the label increased considerably in Zones c and d when indolepropionate was present, but was unchanged in Zone a.

The specific residues to which the dansyl labels were attached were identified when the labeling was conducted both in the presence and absence of indolepropionate. Trypsin digestion of peptide b, labeled with dansyl chloride 1:1 in the absence of indolepropionate) followed by descending chromatography in 1-butanol/glacial acetic acid/water, 4/1/5, v/v for 44 hours produced one strong fluorescent zone (~15 cm from origin) which on amino acid analysis was found to contain residues 50 to 69. This identified the major labeling position as lysine 67, identified as described above, and another which could not be identified by elution from the paper chromatogram. α-Protease digestion of peptide b, followed by resin column chromatography provided three zones containing dansyl residues. On amino acid analysis, they corresponded to residues 39 to 40, 63 to 68, and 63 to 67. Both lysine 39 and 67 were, therefore, dansylated when labeling was conducted in the presence of indolepropionate. Furthermore, lysine 39 was found to be dansylated only in the presence of indolepropionate. In view of the difficulties in elution of dansyl zones from the chromatogram, no quantitation was attempted. However, one should note that the total dansyl label in peptide b, was approximately twice as high when labeled in the absence of indolepropionate, compared to when indolepropionate was present. Labeling at residue 67, therefore, must be much reduced by the presence of indolepropionate occupying the indole binding site.

Tryptsin digestion of peptide c followed by chromatography and electrophoresis (pH 6.0) provided separation of a large peptide-containing residues 1 to 14 which strongly fluoresced. Since trypsin would not be expected to cleave at N'-dansyl-lysyl residues, the labeling position in this peptide was assigned to lysine 13. When labeling was conducted in the presence of indolepropionate, labeling again was found in residue 13 which, on the basis of increased dansylation indicated in peptide c under these dansylation conditions (Fig. 7), led to the conclusion that labeling at lysine 13 was increased approximately 2-fold by the presence of indolepropionate at the binding site.

Zone d obtained from Fragment C isolated from albumin dansylated in the presence of indolepropionate was also analyzed for peptides with dansyl labels. Two dansyl zones, one very strong, and one of moderate fluorescent strength, appeared on a chromatogram developed 24 hours in the 1-butanol/glacial acetic acid/water solvent. The strongly fluorescent zone contained residues 82 to 90, the other weaker zone had an amino acid composition which did not correspond to any known peptide or amino acid sequence found in Fragment C (in nmol, lysine 6.6, proline 13.4, alanine 22, and phenylanine 12). The source of this peptide was not further investigated at this time.

When albumin was labeled with pyridoxal 5'-phosphate in the presence of indolepropionate, the amount of label incorporated into Fragment b, was approximately one-half that incorporated in the absence of indolepropionate. The behavior was similar to dansyl chloride in this respect. From analogy with the dansyl-labeling studies, the fact that pyridoxal 5'-phosphate was earlier found to react with a lysyl residue in albumin which inhibited acetyl-L-tryptophan binding, it is most likely that pyridoxal 5'-phosphate also reacts at lysine 67 in the same manner as dansyl chloride.

**DISCUSSION**

The indole site is inhibited by labeling at two positions in Fragment C, histidine 23 and lysine 67. The peptide chain undoubtedly folds back on itself in this region in development of the binding site. It had earlier been found that a tyrosyl OH group in another major albumin fragment (Fragment A, M, = ~34,000) was also involved in the indole binding site. More recent evidence indicates that this label is located in A-Phe, a subfragment of Fragment A with 92 residues. Thus, these three positions on albumin, although much removed from each other in the peptide sequence, have been identified as within or in close proximity of the indole binding site. One perhaps cannot rule out an effect induced by labeling at some distance from the site.
Since, however, blockage at the site had earlier been shown to be noncompetitive (only $n$, the amount of free site, was affected—not the binding constant itself), a distance effect such as may occur in allosteric regulation seems unlikely.

The position of residue 23 in Fragment C is adjacent to 2 arginine residues, in support of other evidence that arginine is involved in major anionic binding sites on serum albumin (17–20). There is a clustering of types of amino acids in regions of Fragment C. Thus, 5 out of 6 tyrosines, 5 out of 8 phenylalanines, 4 leucines, and 1 isoleucine fall within residues 10 to 43. Serine, glycine, threonine, aspartate, valine, and cysteine are missing in this region. There are 2 proline residues near residue 23, suggesting the absence of helical structure in this part of the site. In the sequence 10 to 43, there are 8 basic residues and only 2 acidic residues. In this segment, amino acids predominate with hydrophobic and basic side chains, which, if proper conformation exists, should provide good binding sites for anions with hydrophobic components. The latter is a property albumin strongly possesses.

Skatole was found to bind to albumin only in the presence of anions which also bind (21). Presumably, the high positively charged density of the site must be neutralized by an anion such as chloride, or thiocholate before it can accept the hydrophobic compound, skatole. The 2 positively charged arginines adjacent to the site, which would strongly favor anionic binding, are consistent with this concept. Interestingly, acetyl-L-tryptophan which has its own anionic charge, binds strongest to albumin in the absence of chloride or thiocholate. One is led to speculate that there may be carboxyl groups on the A-Phe fragment (the latter we know is adjacent to the indole site in Fragment C) which bind to the positively charged arginines in the 10 to 43 residue region of Fragment C, furnishing the forces for forming the site between the A-Phe and C fragments. In the affinity labeling experiments, no reactive lysine groups were evident in the A-Phe fragment, rather, they were overwhelmingly present in the C fragment. The C fragment, on the other hand, contained a number of tyrosines which, however, were generally unreactive with dansyl chloride. The indole site is destroyed in the N-F transition (22), concurrent with the titration of a large number of carboxyl groups. Removal of the charges on Fragment A-Phe would, thus remove the electrostatic attractions holding it to Fragment C.

In the vicinity of residue 67, no such dominance of hydrophobic or basic residues exists. There are 6 acidic and 6 basic residues, 1 phenylalanine, 6 leucines, and no tyrosine, isoleucine, or valine between residues 43 and 81. This area has a number of amino acids with hydrophilic side chains. In addition to the acidic and basic amino acids, there are 3 serines, 2 glycines, 3 half-cystines, and 6 alanines. Unless the folding of the chain fortuitously positions certain residues close to each other, this region would not appear capable of providing hydrophobic forces for binding ligands. This region probably makes up a third side of the binding site, providing mostly positions for accommodating the anionic or hydrophilic segments of the ligands. Tryptophan is located at position 67. Others have noted that it is near major binding sites on albumin. There is no evidence of involvement of residues 87 to 159 of Fragment C in the indole site. Under all conditions studied, this part of Fragment C was labeled only weakly with dansyl chloride and pyridoxal phosphate and not at all with bromoacetyl-L-tryptophan.

Lysine 71 is most likely the residue at which the acetyl group derived from acetylsalicylic acid is attached. Hawkins et al. (23, 24) isolated a fragment with the acetyl group attached, containing 2 leucines, 2 lysines, serine, glutamate, and alanine. In the absence of the acetyl group, this fragment was cleaved into two subfragments, with amino acid compositions leucine, lysine; and serine, glutamate, alanine, leucine, and lysine. The leucine-lysine peptide is apparently the only one in the trypptic digest of albumin. Except for cysteine, which, it appears, was not assayed for by the authors, these residues correspond with our own residues 70 to 76.

The presence of an indole ligand on the indole site makes available other ε-amino groups in the general site region for reaction with dansyl chloride. This is seen at lysines 13, 39, and 84 (Fig. 7), where a large increase in labeling occurred when indolepropionate was present. This latter is also consistent with the previous evidence that the binding of indolepropionate-enhanced association of l-tryptophan at secondary sites (25). Presumably indole ligands binding at the primary site induce a rearrangement of the albumin structure exposing hydrophobic and positively charged groups for the accommodation of further ligands. A very small entropy change was found on the association of skatole, a predominantly hydrophilic molecule, with albumin (26); to explain such a small value, one must allow for an alteration in the structure of albumin, which would counter the entropy change caused by the removal of skatole from the aqueous solution. This thermodynamically anticipated structural change would also be consistent with the exposure of additional hydrophilic areas, presumably to become secondary binding sites for further ligands.

Finally, evidence suggests that the region consisting of residues 1 to 86 of the C fragment and certain residues of the A-Phe fragment is the region for all major binding sites of ligands on albumin. The indole binding site is in this region, and it is seen to accommodate a variety of ligands. In unreported experiments, Mr. Douglas Karrel found that the primary binding site of bilirubin was most likely in the A-Phe fragment, probably adjacent to the C fragment. Salicylic acid binding, as described above, is taken to involve lysine 71 of the C fragment. Farr’s group (27) showed that the binding of 3-acetamido-2,4,6-triiodobenzoate (acetrizoate) was enhanced when the ε-amino group at residue 71 was acetylated. This latter observation implied that this region of the peptide involves the association site of acetrizoate. Chignell and Starkweather (28) found that the binding of phenylbutazone was similarly enhanced on the acetylation of albumin with acetylsalicylic acid. Furthermore, Pinckard et al. (24) found that a number of drugs, dyes, and other compounds, including bilirubin, inhibited acetrizoate binding. Tryptophan did not inhibit acetrizoate binding, nor did bilirubin when added to albumin at a 1:1 molar ratio inhibit tryptophan binding. In some unreported experiments, the primary site for fatty acid binding is indicated to be in the A-Phe fragment. The clustering of major binding sites in this limited area on albumin where interactions between sites are evident may be important in physiological regulation, such as, for example, plasma fatty acid effects on tryptophan binding leading to changes in serotonin levels in the brain (5).

REFERENCES

1. Gambrill, K. K., and McMenamy, R. H. (1973) J. Biol. Chem. 248, 1956-1960
2. Trutsch, G. L., and Trutsch, N. E. (1963) J. Biol. Chem. 238, 138-142
3. McMenamy, R. H. (1965) J. Biol. Chem. 240, 4235-4243
4. Marcus, G., McClintock, D. K., and Castellani, B. A. (1967) J. Biol. Chem. 242, 4402-4408
5. Fernstrom, J. D., and Wurtman, R. J. (1974) Sci. Am. 230, 64-91
6. Hermodson, M. A., Erickson, L. H., Titani, K., and Neubr, H. (1972) Biochemistry 11, 4492-4502
7. Hill, R. L., and Delaney, R. (1967) Methods Enzymol. 11, 339-351
8. Braunitzer, G., Schrank, B., and Rufus, A. (1970) Hoppe-Seyler's Z. Physiol. Chemie 351, 1589-1590
9. Gray, W. R. (1967) Methods Enzymol. 11, 469-478
10. Sauer, R. T., Niall, H. D., Hogan, M. L., Keutmann, H. T., O'Riordan, J. L. H., and Potts, J. T., Jr. (1974) Biochemistry 13, 1984-1999
11. Jeppsson, J. O., and Sjoquist, J. (1967) Anal. Biochem. 18, 264-269
12. Inagami, T., and Murakami, K. (1972) Anal. Biochem. 47, 501-504
13. Pisano, J. J., and Bronzert, T. J. (1969) J. Biol. Chem. 244, 5597-5607
14. Yamada, S., and Itano, H. A. (1966) Biochim. Biophys. Acta 130, 538-540
15. Sanger, F., and Tuppy, H. (1951) Biochem. J. 49, 465-490
16. McMenamy, R. H., Dintzis, H. M., and Watson, F. (1971) J. Biol. Chem. 246, 4744-4750
17. Pande, C. S., and McMenamy, R. H. (1970) Arch. Biochem. Biophys. 136, 260-267
18. McMenamy, R. H., Madeja, M. I., and Watson, F. (1968) J. Biol. Chem. 243, 2228-2236
19. Grossberg, A. L., and Pressman, D. (1968) Biochemistry 7, 272-279
20. Jonas, A., and Wedek, G. (1971) Biochemistry 10, 4492-4496
21. McMenamy, R. H., Madeja, M. I., and Watson, F. (1968) J. Biol. Chem. 243, 2625-2630
22. McMenamy, R. H. (1964) J. Biol. Chem. 239, 2835-2841
23. Hawkins, D. R., Pinckard, N., Crawford, C. P., and Farr, R. S. (1969) J. Clin. Invest. 48, 536-542
24. Pinckard, R. N., Hawkins, D., and Farr, R. A. (1974) Ann. N. Y. Acad. Sci. 226, 341-354
25. McMenamy, R. H., and Seder, R. H. (1963) J. Biol. Chem. 238, 3241-3248
26. Krasner, J., and McMenamy, R. H. (1965) J. Biol. Chem. 241, 4186-4190
27. Farr, R. S., Radman, G. P., and Lasser, E. C. (1961) J. Clin. Invest. 40, 1037
28. Chignell, C. F., and Starkweather, D. K. (1971) Mol. Pharmacol. 7, 226-237