[Contribution from the Department of Biochemistry, the University of Texas, M. D. Anderson Hospital and Tumor Institute, Houston, Texas]

Studies of the N- and C-Terminal Amino Acid Sequence of Human Serum Albumin¹

By Tokuji Ikenaka

Received September 28, 1959

The N-terminal amino acid sequence of human serum albumin (HA) was found to be Asp.Ala.His² by determining the chemical structure of the dinitrophenyl (DNP) peptide obtained from the enzymic hydrolysate of DNP-HA, using bacterial proteinase for the hydrolysis. A combination of the carboxypeptidase method and the hydrazinolysis method was used to investigate the C-terminal sequence of HA, which was found to be (Leu, Val, Ala₂₋₄) Gly.Leu. In an attempt to explain the incomprehensible relative rate of release of alanin¹ from HA by carboxypeptidase (CPase), carbobenzoxy-L-alanyl-L-alanyl-glycine was synthesized and digested with CPase. The release of alanine and glycine from the peptide was similar to that of human albumin.

Introduction

Several investigators have studied the N-terminal group of HA, and all have found one mole of aspartic acid as the N-terminal group.³⁻⁵ The amino terminal sequence of HA has been studied by E.O.P. Thompson, who obtained DNP-Asp, DNP-Asp.Ala, and a DNP-tripeptide in a partial hydrolysate of DNP-HA, but he could not detect the third amino acid in a hydrolysate of the DNPtripeptide.⁵ Titani, et al., recently obtained DNPcysteic acid in a hydrolysate of DNP-oxidized equine serum albumin and suggested the presence of half cystine as another N-terminal group of that protein.6 Thompson investigated the possibility of half cystine as N-terminal group of bovine serum albumin, but no DNP-cysteic acid could be detected in oxidized-DNP bovine serum albumin.7 G. Biserte has recently reported that cystine could not be found as the N-terminal amino acid of human serum albumin.8

The C-terminal amino acid of HA has been determined to be one mole of leucine by both the carboxypeptidase⁹ and hydrazinolysis methods.¹⁰ The C-terminal sequence, -Gly.Val.Ala.Leu, was postulated on the basis of the rates of release of the amino acids by CPase.⁹

This paper deals with studies concerning the N- and C-terminal animo acid sequence of HA.

Experimental

Materials.—Lyophilized HA, contributed by the Protein Foundation Laboratory, Jamaica Plain, Massachusetts, was used in this experiment. The protein was found to be better than 95% pure as found by the moving boundary method of electrophoresis using the Perkin-Elmer Model

(2) Abbreviations for the amino acid residues are those presented by E. Brand, Ann. N. Y. Acad. Sci., 47, 187 (1946), and Ann. Rev. Biochem., 16, 224 (1947).

(3) P. Desnuelle, M. Rovery and C. Fabre, Compt. rend., 233, 987 (1951).

(4) H. Van Vunakis and E. Brand, Abstracts from the 119th meeting of the American Chemical Society, Boston, Mass., April, 1951, p. 28C.

(5) E. O. P. Thompson, J. Biol. Chem., 208, 565 (1954).

(6) K. Titani, H. Yoshikawa and K. Satake, J. Biochem. (Japan), 43, 737 (1956).

(7) E. O. P. Thompson, Biochim. et Biophys. Acta, 29, 643 (1958).
(8) G. Biserte, ibid., 34, 558 (1959).

(9) W. F. White, J. Shields and K. C. Roggins, THIS JOURNAL, 77, 1267 (1955).

(10) K. Kusama, J. Biochem. (Japan), 44, 375 (1957).

38 A Tiselius Electrophoresis Apparatus. This is in agreeinent with the result of the N-terminal group analysis that the protein has one mole of aspartic acid per mole protein as the N-terminal group, and no other N-terminal amino acid could be found. The crystallized bacterial proteinase (from *Bacillus subtilis*) was kindly supplied by Nagase and Company, Ltd., Osaka, Japan. The CPase (recrystallized three times and treated with diisopropylfluorophosphate) was from the Worthington Biochemical Corporation, Ltd., and was recrystallized a fourth time before use. DNP-HA.—HA was dinitrophenylated with fluorodinitro-

DNP-HA.—HA was dinitrophenylated with fluorodinitrobenzene in 66% alcoholic solution in the presence of sodium bicarbonate for 2 hr., by Sanger's method.¹¹ The DNPprotein was washed successively with water, alcohol and ether, then dried in a vacuum desiccator.

Enzymic Digestion of DNP-HA.—The pH of a suspension of 400 mg. of DNP-HA in 20 ml. of 0.05 *M* dibasic sodium phosphate was adjusted to pH 8.0 by adding 1 *N* sodium hydroxide. Twenty mg. of crystallized bacterial proteinase was then added. The digestion was carried out at 37° for 24 hr., during this period DNP-HA was completely dissolved. The reaction mixture was then acidified to *p*H 2.0 by addition of 2 *N* hydrochloric acid. Upon acidification, a yellow precipitate was formed. The suspension was extracted three times with 15 ml. portions of ethyl acetate and then the precipitate was separated by centrifugation from the aqueous part. In order to determine a distribution of N-terminal DNP-aspartic acid and/or N-terminal DNPpeptide in each fraction (the ethyl acetate, the precipitate and the aqueous part), one tenth of each fraction was hydrolyzed with 6 *N* hydrochloric acid in the ether layer was determined by the usual method. The ethyl acetate extract contained 75% of the N-terminal group of the original protein, the water layer 6% and the precipitate 19%.

Separation and Purification of N-Terminal Peptides.— The ethyl acetate layer was used for the purification of Nterminal peptides. This layer was extracted three times with 20 ml. portions of 1% sodium bicarbonate. Each bicarbonate solution was washed twice in two other separatory funnels with 20 ml. of ethyl acetate as a simple three-tube countercurrent procedure. The three bicarbonate solutions were then combined, acidified with hydrochloric acid and extracted three times with 20 ml. portions of ethyl acetate. The purified ethyl acetate layer contained 70% of the N-terminal group in the original ethyl acetate layer (53% of the original DNP-protein).

The purified ethyl acetate layer was then dried *in vacuo*, and the residue was chromatographed on five papers, using the upper layer of a mixture of *n*-butanol, tertiary amyl alcohol, pyridine, ammonium hydroxide (28%), acetic acid and water (300:100:10:25:500) as the first solvent and 1.5*M* phosphate buffer of *p*H 6.0 as the second solvent. The chromatogram is shown in Fig. 1. Each spot was cut from the paper and placed in 1% sodium bicarbonate at 55° for 15 minutes. Each solution was then acidified and the N-terminal DNP-peptides were extracted with a mixture of *n*-butanol and ethyl acetate (1:1).

The Constituent Amino Acids of DNP-peptides.—Each DNP-peptide obtained as described above was hydrolyzed with 6N hydrochloric acid at $105-110^{\circ}$ for 7-9 hr. in a sealed tube. The hydrolysate was diluted with water and ex-

⁽¹⁾ This investigation was supported in part by a grant from the Robert A. Welch Foundation, Houston. Texas, Grant No. G-051 and the National Cancer Institute, National Institutes of Health, U. S. Public Health Service, grant NCI-1785. Presented at the 136th meeting of the American Chemical Society, Atlantic City, New Jersey, September 14. 1959.

⁽¹¹⁾ F. Sanger, Biochem. J., 39, 507 (1945).



Fig. 1.—Chromatogram of DNP-peptides in the ethyl acetate extract of the enzymic hydrolysate of DNP-HA with bacterial proteinase. The upper layer of a mixture of *n*-butanol, *t*-amyl alcohol, pyridine, ammonium hydroxide, acetic acid and water (300:100:100:10:25:500) was used as first solvent and 1.5 M phosphate buffer of ρ H 6.0 was used as second solvent. The positions of several DNP-amino acids and dinitrophenol are indicated by dotted circles.

tracted with ether, in which DNP-aspartic acid was extracted. The aqueous layer, which contained free amino acids or water-soluble DNP-amino acids, was dried and the residue was chromatographed by the usual amino acid chromatography or after conversion to DNP-amino acids by DNP-amino acid chromatography, in which the first and second solvents were the upper layer of a mixture of tertiary amyl alcohol, ethanol, chloroform and 0.8 N ammonium hydroxide (6:3:2:9) and 1.5M phosphate buffer of pH 6.0, respectively. The constituent amino acids of each peptide are given in Table I. The constituent amino acids of peptides 5, 7 and 9 could not be determined, because the amounts of these DNP-peptides were inadequate.

TABLE I

Peptide Constituent Amino Acids of DNP-Peptides

- 1 ε-DNP-Lys, Phe, Leu, Gly, Pro, Ser, Glu, Tyr, Val
- 2 e-DNP-Lys, Phe, Leu, Gly, Pro, Ala, Glu, Tyr, Val
- 3 ε-DNP-Lys
- 4 ε-DNP-Lys, Phe, Tyr, Asp
- 6 DNP-Asp, Ala, DNP(im)-His
- 8 ε-DNP-Lys, Asp
- 10 ε-DNP-Lys, Val

N-terminal DNP-peptide.—As shown in Table I, only peptide 6 contained DNP-aspartic acid. It is of interest that DNP-aspartic acid, the N-terminal amino acid of HA, could not be recognized on the paper chromatogram.

After hydrolysis of peptide 6 with hydrochloric acid, DNPaspartic acid was extracted with ether. The aqueous layer was dried and the residue was chromatographed, using the upper layer of a mixture of *n*-butanol, acetic acid and water (4:1) as the first solvent and phenol and water (4:1) as the second solvent. After the chromatogram was sprayed with ninhydrin solution and heated, two spots, alanine and an unknown spot, were found. The R_t value of the unknown spot was not in agreement with the usual amino acid R_t value but was the same as that of mono DNP(im)-His, which was prepared by dinitrophenylation of N-acetyl histidine followed by hydrolysis. A part of the aqueous layer of the hydrolysate of peptide 6 was also dinitrophenylated and extracted with ethyl acetate. DNP-alanine and di-DNP-histidine were found in the ethyl acetate layer by DNP-amino acid chromatography.

Table 11

C-TERMINAL AMINO ACIDS OF HUMAN ALBUMIN AFTER VARIOUS PERIODS OF DIGESTION

minal amino acid	Original		0.5	6	24	48 (hr.)	
Leu	0.54	(0,9)	0	0	0	0	
Gly	.05		0.45(0.9)	0.35(0.7)	0.35(0.7)	0.30(0.6)	
Ala	.06		.03 (0.05)	.06(0.1)		0,05(0.1)	
Asp.							
Glu			.03	.03	0.05		
Ser				.04	0.04	0.04	

Digestion of HA with CPase.—The pH of a solution of 500 mg. of HA in 24 ml. of water was adjusted to 8.0 by adding 0.2 M sodium carbonate (about 1 ml.). One ml. of CPase solution which contained 10 mg. of the enzyme was added to the HA solution. The reaction was carried out at room temperature (25°). After 30 minutes and after 2, 6, 24 and 48 hr., 5 ml. portions of the reaction mixture were pipetted into a flask and the pH was adjusted to 4.0 by addition of 0.1 N hydrochloric acid and stirred with about 1 g. of Dowex 50-X8(H form, 100–150 mesh) to adsorb amino acids released from HA by the CPase. The supernatant was decanted and the resin was washed two times by stirring for 5 minutes with 5 ml. portions of water. The washing solutions were combined with the supernatant and lyophilized. This then was HA digested with CPase. The resin was washed still two times with p addition and the resin. The ammonium hydroxide to elute amino acids from the resin. The ammonium hydroxide solution was then dried in a vacuum desiccator over sulfuric acid. The kind and amount of amino acids in the residue were determined by Stein and Moore's column chromatography or by the DNP-amino acid method. Figure 2 shows a rate of release of amino acid from HA by the action of CPase under the conditions mentioned above.



Fig. 2.—Rate of release of amino acids from HA by CPase; enzyme and substrate (1:50) at 25°, pH 8.0.

Hydrazinolysis of the Digested HA.—In order to confirm the next amino acid from the C-terminal leucine, the HA isolated after digestion with CPase was hydrazinolyzed by Akabori's method.¹² Table II shows the C-terminal amino acids of the HA obtained after digestion with CPase for several different periods of time. Data given in the table show mole of amino acid per mole of protein, assuming 70,000 as the molecular weight of HA. Data given in parentheses are after correction for loss of the amino acids during hydrazinolysis, extraction procedures and paper chromatography.

By the hydrazinolysis of the original HA, no glycine was found as the C-terminal amino acid. After release of about one mole of the C-terminal leucine from the original HA by CPase for 30 minutes, 0.9 mole of glycine was found as

⁽¹²⁾ S. Akabori, K. Ohno, T. Ikenaka, Y. Okada, H. Hanafusa, I. Haruna, A. Tsugita, K. Sugae and T. Matsushima, *Bull. Chem. Soc. Japan*, **29**, 507 (1956).



Fig. 3.—Rate of release of amino acids from carbobenzoxy-L-alanyl-I -alanyl-L-alanyl-glycine by CPase (2 ing. of the enzyme to 5 mg. of the substrate).

Fig. 4.—Rate of release of amino acids from carbobenzoxy-L-alanyl-L-alanyl-glycine by CPase (0.2 mg. of the enzyme to 5 mg. of the substrate).

the new C-terminal group of the digested HA, and alanine was only 0.05 mole per mole of protein. After 24 hr. of digestion, 0.2 mole of glycine was released from the original HA; and after 48 hr., 0.3 mole was released. The new Cterninal glycine of the digested protein was 0.7 and 0.6 mole, respectively. The sum of the released glycine and the new C-terminal glycine was always about one mole. Synthesis of Carbobenzoxy-L-alanyl-L-alanyl-

Synthesis of Carbobenzoxy-L-alanyl-L-alanyl-L-alanylglycine.—Carbobenzoxy-L-alanyl-L-alanyl-L-alanine hydrazide (Compound I) was synthesized by the procedure of Erlanger, Brand and Sachs.^{13,14} It was then converted to its azide by the action of nitrous acid and allowed to react with glycine ethyl ester to give carbobenzoxy-L-alanyl-Lalanyl-L-alanyl-glycine ethyl ester (Compound II), which had a melting point of 224–225°. The amino acid composition of this peptide ester was determined, after hydrolysis with hydrochloric acid, by the DNP method. The molar ratio of glycine to alanine was estimated to be 1.07: 3.00.

One hundred mg. of Compound II was dissolved in 2 ml. of dioxane and 2 ml. of water and then saponified with 0.27 ml. of 1 N sodium hydroxide for 1.5 hr. in a refrigerator. Upon addition of a slight excess of 1 N hydrochloric acid, the solution was returned to the refrigerator and allowed to stand overnight. The carbobenzoxy-L-alanyl-L-alanyl-Lalanyl-glycine (Compound III) crystallized by this treatment was filtered and washed with cold water. The crystals recrystallized from a mixture of dioxane-ether had a melting point of $241-242^{\circ}$.

Anal. Calcd. for $C_{19}H_{26}O_1N_4$: C, 54.02; H, 6.02; N, 13.26. Found: C, 53.98; H, 6.11; N, 13.30.

The C-terminal amino acid of the peptide (Compound III) was determined to be 0.9 mole of glycine per mole of peptide by the hydrazinolysis method.¹² Digestion of Carbobenzoxy-L-alanyl-L-alanyl-L-alanyl-

Digestion of Carbobenzoxy-L-alanyl-L-alanyl-L-alanylglycine (Compound III) with CPase.—To 5 ml. of a 0.1%solution of Compound III (adjusted to pH 7.5) was added 0.5 ml. of CPase solution containing 2 mg. of the enzyme. The digestion was carried out at room temperature (25°). The rate of release of the annino acids from the peptide is given in Fig. 3, which shows that the ratio of glycine to alanine released from the peptide was 1:1.9 at any time during the digestion.

Figure 4 shows the rate of release of the amino acids from the same peptide when the amount of enzyme used for the digestion was reduced to one tenth (5 mg, of the substrate and 0.2 mg, of the enzyme in 5.5 ml, of the reaction mixture).

Discussion

From a partial acid hydrolysate of DNP-HA Thompson obtained DNP-Asp, DNP-Asp.Ala, and a DNP-tripeptide for which the third amino acid in the sequence could not be determined.⁵ He suggested, however, from the chromatographic

(13) B. F. Erlanger and E. Brand, THIS JOURNAL, 73, 3508 (1951).
(14) E. Brand, B. F. Frlanger and H. Sachs, *ibid.*, 74, 1849 (1952).

behavior of the DNP-tripeptide, that the third amino acid might have some aromatic group. By the experiments described herein, the third amino acid was found to be mono-DNP(im)-His. This result is in agreement with Thompson's suggestion. It is of interest that only DNP-Asp.Ala.DNP(im)-His was found in the ethyl acetate extract of the enzymic digestion product of DNP-HA. As has been indicated in Fig. 1 and in Table I, neither DNP-Asp nor DNP-Asp.-Ala was found. This might be explained as follows: The enzyme cannot combine with the first or second peptide bond from the N-terminal group due to steric hindrance by such large radicals as DNP-groups, one of which was combined with the N-terminal amino group and another with the imidazole radical of the third amino acid. The third peptide bond was more easily attacked by the enzyme, so that only DNP-Asp.Ala.DNP(im)-His was obtained by such enzymic hydrolysis. Chymotrypsin, trypsin and papain were also tried; however, the DNP-protein could not be digested with such enzymes.

In experiments concerning the rate of release of amino acids from HA by CPase, results were similar to those of White, *et al.*, and seemed at first to support his suggestion that the C-terminal amino acid sequence of HA might be-(Gly,Val) Ala.Leu.⁹

Alanine was expected to be in the C-terminal position of the digested protein, which was separated after liberation of the C-terminal leucine from the original HA by CPase. Contrary to such expectation, however, glycine was found to be the C-terminal amino acid of the protein, as shown in Table II and Fig. 2. These results show that the the C-terminal sequence of HA might be-Gly.Leu.

In order to explain the release of alanine from HA by CPase as though it were the amino acid next to the C-terminal group of the protein, three possibilities were considered:

(1) That an endopeptidase, such as trypsin or chymotrypsin, might have been present as a contaminant in the CPase and caused the splitting of some peptide bond. This could expose a new Cterminal alanine to attack by CPase. To investigate the probability of such action, the N-terminal amino acid of the HA digested for 2 hr. was determined by the DNP method; one mole of aspartic acid was the only N-terminal amino acid found for one mole of protein. This possibility should, therefore, be given up.

(2) That HA might have two C-terminal groups, one being leucine and another being alanine amide. By the action of CPase, or some other enzyme (an amidase?) present as a contaminant in the CPase, the C-terminal amide could have been split off initially (along with leucine) and alanine would have been released next by the CPase as the new Cterminal group. In such a case, ammonia would have been liberated from the HA; however, no ammonia was detected in the progress of the enzymic action.

(3) That the greater specificity of CPase for peptide bonds containing alanine might account for the greater release of this amino acid if more units

Vol. 82

of it were available in the sequence preceding a -Gly.Leu C-terminal ending for HÅ, such as -Ala.Ala.Ala.Gly.Leu. From data concerning the specificity of CPase for a number of peptides studied by Neurath, et al.,¹⁵ it was suspected that a peptide bond containing alanine as the C-terminal group might be hydrolyzed about ten or fifteen times faster than such a bond containing glycine as the C-terminal group. To substantiate this possibility, carbobenzoxy-L-alanyl-L-alanyl-L-alanyl-glycine was synthesized and digested by CPase. As shown in Fig. 3 and 4, more alanine than glycine was released during any period of time, just as though the C-terminal amino acid of the peptide were alanine, and the ratio of glycine to alanine released from the peptide was about 1:2. These results show that more alanine than glycine can be released from HA if more alanine units follow glycine in the C-terminal sequence. They also show that the bond between the carbobenzoxy group and alanine was not hydrolyzed with CPase.

The digested HA products, which were isolated after digestion of the original HA with CPase for 0.5 hr. in one case and 2 hr. in another, were then further digested with CPase. During any period of time in this digestion, 3.5 moles of alanine, 1 mole of valine and 1.5 moles of leucine were released for 1 mole of glycine. From these results, the C-terminal amino acid sequence of HA is suggested to be -(Leu, Val, Ala₃₋₄) Gly.Leu; however, the exact sequence of the alanines, valine and leucine is not suggested from the present experimental data. As demonstrated by the experiments described herein, if the amino acids of a protein are not arranged according to decreasing rates of enzymic hydrolysis, the amino acids may not be released in order of the C-terminal amino acid sequence. HA is an example of such a protein. Therefore, it should be emphasized that the carboxypeptidase method must be used for the determination of the amino acid sequence, along with another method.

The author also confirmed independently of the report of Biserte⁸ that DNP-cysteic acid could not be detected in a hydrolysate of DNP-oxidized HA or oxidized DNP-HA. Therefore, it is supposed that HA has one peptide chain, with aspartic acid as the N-terminal amino acid and leucine as the C-terminal acid as:

Asp.Ala.His-----(Leu, Val, Ala₃₋₄) Gly.Leu.

This was supported by investigation of the physicochemical properties of oxidized and reduced carboxymethylated HA.^{16,17}

Acknowledgments.—The author is indebted to Dr. Bruno Jirgensons and Dr. Darrell N. Ward for their helpful discussion concerning this study and also to Dr. A. Clark Griffin, Head of the Department of Biochemistry, for his interest in this work.

(15) H. Neurath and G. W. Schwert, Chem. Revs., 46, 69 (1950).

(16) M. J. Hunter and F. C. McDuffie, THIS JOURNAL, **81**, 1400 (1959).
(17) B. Jirgensons and T. Ikenaka, *Makromol. Chem.*, **31**, 112 (1959).

·····

[CONTRIBUTION FROM THE DIVISION OF PURE CHEMISTRY, NATIONAL RESEARCH COUNCIL OF CANADA]

Cytodeuteroporphyrin^{1,2}

By G. S. Marks,³ D. K. Dougall,³ E. Bullock³ and S. F. MacDonald Received October 22, 1959

The syntheses of the four desmethyl derivatives of deuteroporphyrin 9 have been completed by the syntheses of the 1-, 5and 8-desmethyl derivatives. The cytodeuteroporphyrin from cytohemin (hemin-a) was identical with the last of these.

Cytohemin (hemin-a) is the prosthetic group of the sauerstoffübertragende Ferment (cytochrome-a₃) and of cytochrome-a. Crystalline cytohemin⁴ had been degraded to cytodeuteroporphyrin⁵ by fusing with resorcinol to split off the labile sidechains and then removing the iron. Cytodeuteroporphyrin contained two carboxyl groups which had survived resorcinol fusion suggesting propionic acid rather than acetic acid or ring carboxyl groups. Its oxidation had led to methylmaleimide and hematinic acid in the same yields as from deuteroporphyrin 9. This and the analysis of cytodeuteroporphyrin had suggested that it might be one of the fifteen deuteroporphyrins⁶; however, it differed from deuteroporphyrin 9 and from the syn-

(1) This was the subject of a preliminary communication, THIS JOURNAL, **81**, 250 (1959).

(2) Issued as N.R.C. No. 5643.

(3) National Research Council of Canada Postdoctorate Fellow.

(4) O. Warburg and H. S. Gewitz, Z. physiol. Chem., 288, 1 (1951).
(5) Idem, ibid., 292, 174 (1953).

(6) For the nomenclature of the deuteroporphyrin isomers see H. Fischer and H. Orth, "Chemie des Pyrrols," Vol. II/I, Leipzig, 1937, p. 409.

thetic deuteroporphyrins 1, 2, 3, 5 and 6. Cytodeuteroporphyrin had then been shown to form a tribromo derivative rather than the dibromo derivative expected of a deuteroporphyrin. It evidently had three free positions and one possible structure had been excluded by comparison with synthetic 3-desmethyl-deuteroporphyrin 9(II).⁷

As to the substituents in cytodeuteroporphyrin were not completely defined analytically and particularly as those of cytohemin were evidently not all of normal chain length,⁴ the possible structures for cytodeuteroporphyrin could not be definitely limited. According to another approach,⁸ two free positions were to be expected on opposite rings in cytodeuteroporphyrin because the spectrum ("oxo-rhodo" type) of the porphyrin corresponding to cytohemin had suggested labile side chains ("rhodifying" groups) in these

(8) R. Lemberg, Nature, 172, 619 (1953).

 ⁽⁷⁾ O. Warburg, H. S. Gewitz and W. Völker, Z. Naturforsch., 10b, 541 (1955). We have continued this work at the suggestion of Professor Warburg, who kindly provided the specimen of cytodeutero-porphyrin methyl ester from crystalline cytohemin ex horse heart.
 (2) D. Lorder Warburg, Value, 200 (1970)