THE N-TERMINAL SEQUENCE OF SERUM ALBUMINS; OBSERVATIONS ON THE THIOHYDANTOIN METHOD*

BY E. O. P. THOMPSON

(From the Laboratory for the Study of Hereditary and Metabolic Disorders, and the Departments of Biological Chemistry and Medicine, University of Utah College of Medicine, Salt Lake City, Utah)

(Received for publication, December 21, 1953)

For the determination of N-terminal sequence of amino acids in peptides and proteins the 1,2,4-fluorodinitrobenzene (FDNB) method (dinitrophenyl (DNP) method) of Sanger (1) has been successfully used. From a series of DNP peptides identified in partial acid hydrolysates of the DNP protein, the N-terminal sequence can be deduced (2-5). A limitation of this method is that, if an acid-labile peptide bond involving serine, threonine, or tryptophan is close to the N-terminal residue, few, or possibly no DNP peptides, will be obtained. Moreover, no DNP peptides higher than pentapeptides have been identified by this method.

An alternative approach involves stepwise degradative procedures (6) and recently the thiohydantoin method of Edman (7) has been applied to proteins (8-11). With this method it is theoretically possible to proceed down the chain with small quantities of material, provided that a highly selective method for the cleavage and cyclization of the phenylthiocarbamyl (PTC) derivative is available and that good yields are obtained at each stage.

The results with insulin in aqueous media (10) suggested that the cleavage did not proceed smoothly and resulted in scission of other peptide bonds, probably those involving the amino groups of hydroxyamino acids. With saturated guanidine hydrochloride solutions (8) the reaction has been shown to be reasonably quantitative, with losses of 15 to 20 per cent in experiments with tripeptides and proteins. By continuous extraction of the phenylthiohydantoin during the cleavage reaction in aqueous buffer solutions, almost quantitative recoveries have been achieved with simple peptides (12). However, the phenylthiohydantoins of arginine, lysine, histidine, and cystine are not extractable from aqueous solutions (10). The original method of Edman (7, 13) used non-aqueous media for the cleavage and cyclization step, and a modification of this type of reagent followed by

* This investigation was aided by a grant from the Rockefeller Foundation to Professor E. L. Smith and by grants from the National Institutes of Health, United States Public Health Service.
identification of the phenylthiohydantoins on paper chromatograms has been successfully used with several proteins (9).

This technique has now been applied to carboxypeptidase, the N-terminal sequence of which is asparaginylserine (14). Preliminary experiments gave the expected thiohydantoin of asparagine, identified by paper chromatography (9, 15), but the yield was poor. With human serum albumin, which was more readily available and the DNP derivative of which has been shown (16, 17) to give 1 mole of DNP-aspartic acid after acid hydrolysis, the phenylthiohydantoin of aspartic acid was obtained in poor yield. The low recovery was due to loss of the thiohydantoin after the cleavage reaction, since the residual protein, after coupling with FDNB followed by acid hydrolysis, gave a much higher yield of the DNP derivative of the 2nd amino acid residue which proved to be alanine. The sequence aspartylalanine was confirmed by isolation of DNP-aspartylalanine from partial hydrolysates of the DNP protein. Bovine serum albumin was also examined by both the Sanger and Edman techniques. The results indicated the N-terminal sequence aspartylthreonine, different from that in human serum albumin.

EXPERIMENTAL

The human serum albumin (kindly supplied by Dr. E. L. Smith) was prepared by Method 5 of Cohn et al. (18) and reprecipitated at its isoelectric point with 40 per cent ethanol until electrophoretically homogeneous.

Bovine serum albumin (Armour) and carboxypeptidase (Worthington, recrystallized three times) were commercial preparations. The 3-phenyl-2-thiohydantoin (PTH) derivatives of amino acids were prepared as described by Edman (19), but acetic acid-hydrogen chloride (13) was used for cyclization of the PTC derivatives.

Preparation of 3-Phenyl-2-thiocarbamyl Derivatives of Proteins—The albumins (1 gm.) were dissolved in 10 ml. of water, the pH was adjusted to 8.5 with triethylamine, and 2.5 ml. of dioxane were added. A solution of phenyl isothiocyanate (1 ml.) in 7.5 ml. of dioxane was slowly added with stirring, the reaction mixture being maintained at 40° and pH 8.5. Stirring was continued for 1 to 2 hours and the reaction mixture was then extracted six times with 50 ml. portions of benzene. The residual suspension was evaporated to dryness in vacuo and washed with acetone and ether. Yield 1.14 gm.

The PTC-carboxypeptidase was prepared similarly with smaller amounts of carboxypeptidase in suspension.

Cleavage of PTC Proteins, Extraction, and Identification of 3-Phenyl-2-thiohydantoins—The PTC human albumin was cleaved with the reagents shown in Table I, and PTC bovine albumin and PTC-carboxypeptidase
were treated as described below. After completion of the reaction period the anhydrous reagents were removed in vacuo. The PTH derivatives were extracted from the dry residue by shaking three times with 50 ml. portions of ethyl acetate, since this was found to be a better solvent than ether. Extraction was also tried on an aqueous suspension of the residue at pH 3, but in both cases recovery of the PTH amino acids was low.

With the citrate buffer, continuous extraction with benzene during cleavage at 65–70° for 20 hours was used (12). The insoluble PTC derivative tended to "creep" up the sides of the flask and continuous contact with the aqueous solution could not be maintained. This may account for the small amount of cleavage (Table I).

The PTH derivatives were identified by paper chromatography (9, 15). For the separation of PTH-aspartic acid and PTH-asparagine, Solvent C of Sjöquist (15) gave best results. The 2-butanol buffer system of Landmann et al. (9) was also useful. The spots were detected with an iodine-azide reagent similar to that of Sjöquist (15), but pretreatment of the paper (Whatman No. 1) with starch was not used. The spots appeared white on a brown background.

With PTC-carboxypeptidase (25 mg.) cleavage and cyclization were effected with 5 ml. of acetic acid-hydrogen chloride (5 minutes, 30°) and anhydrous formic acid (2 hours, 25°). Aliquots equivalent to 5 mg. gave

---

**Table I**

<table>
<thead>
<tr>
<th>Cleavage reagent</th>
<th>Conditions</th>
<th>DNP amino acids identified</th>
<th>Yield* (moles per mole albumin, mol. wt. 69,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetic acid-hydrogen chloride (13)†</td>
<td>5 min., 39°</td>
<td>Alanine</td>
<td>0.5</td>
</tr>
<tr>
<td>Dioxane-hydrogen chloride (21)†</td>
<td>2 hrs., 25°</td>
<td>Threonine</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serine</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aspartic acid</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alanine</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Threonine</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serine</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aspartic acid</td>
<td>0.1</td>
</tr>
<tr>
<td>Formic acid‡</td>
<td>5 min., 39°</td>
<td>Alanine</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>2 hrs., 25°</td>
<td>Aspartic acid</td>
<td>0.6</td>
</tr>
<tr>
<td>Citrate buffer, pH 4.5 (0.05 M) (12)†</td>
<td>20 hrs., 65-70°</td>
<td>Alanine</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aspartic acid</td>
<td>0.1</td>
</tr>
</tbody>
</table>

* Corrected for destruction with factors of Porter and Sanger (20).
† Bibliographic citation.
‡ Suggested by Dr. F. Sanger.
spots of PTH-asparagine and no PTH-aspartic acid on paper chromatograms.

Both PTC human serum albumin and PTC bovine serum albumin (50 mg.) gave only PTH-aspartic acid under the above conditions with aliquots equivalent to 10 mg. However, the recoveries were poor and from the spot intensities a maximal yield of 5 per cent was estimated. Similar low yields were obtained after barium hydroxide regeneration of the amino acids (7). Dimethylformamide is a good solvent for PTH derivatives and this extracted somewhat larger amounts, though its high boiling point was a disadvantage.

Reaction of Residual Protein with FDNB—The residue, after the cleavage and cyclization step, was collected and coupled with FDNB in the usual way (22). After the mixture had been shaken for 3 hours, it was extracted three times with ether to remove unchanged FDNB and the residue was hydrolyzed with 5 ml. of 6 N hydrochloric acid (redistilled three times in glass and used in all the work reported herein) in a sealed tube at 105°. The hydrolysate was investigated by methods previously described (5).

For human albumin the DNP amino acids identified and the yields obtained with various cyclization reagents are shown in Table I. The accuracy of these measurements was probably low, as no special determinations of correction factors for destruction of the DNP amino acids were made.

Investigation of DNP Serum Albumins—The coupling with FDNB followed the standard procedure (22). Amide estimations showed that 100 mg. of air-dried DNP human serum albumin corresponded to 79 mg. of moisture-free albumin. The amide ammonia content of human serum albumin was 0.94 gm. per 100 gm. This value is somewhat lower than that of Brand et al. (23) who found the amide content of human and bovine serum albumins to be identical, 1.05 gm. to 100 gm. Stein and Moore (24), however, found a maximal amide content of 0.95 gm. per 100 gm. for bovine albumin. The number of FDNB reactive groups in human and bovine albumins is very similar (25) and the content of protein in DNP bovine albumin was assumed to be identical with that for human albumin.

The methods of hydrolysis, extraction, chromatography on buffered Celite 545 columns (26), and estimation of DNP derivatives from complete and partial hydrolysates were the same as those previously described (5), ethyl acetate being used for the extraction (2) of partial hydrolysates. The yields of DNP derivatives are expressed in moles per mole of serum albumin, mol. wt. 69,000.

Complete hydrolysates of human serum albumin (16 hours) gave 0.72 mole of DNP-aspartic acid (uncorrected). No other DNP amino acids
were detected. With 12 N hydrochloric acid, after 16 hours of hydrolysis at 105°, no DNP-proline was detected (20) and only 0.26 mole of DNP-aspartic acid (uncorrected) was obtained. This is a much higher destruction than in 6 N hydrochloric acid and a much lower recovery than that found by Porter and Sanger (20).

Complete hydrolysates of DNP bovine serum albumin (18 hours) gave 0.64 mole of DNP-aspartic acid (uncorrected). Correction factors were not determined, since previous workers (16, 17) have established stoichiometric values of 1 mole for a molecular weight 69,000. If the correction factors of Porter and Sanger are applied, the recoveries are 0.98 and 0.92 moles for human and bovine albumins, respectively.

Partial hydrolysates of DNP human albumin gave several DNP-aspartyl derivatives. After 8 days of hydrolysis (12 N HCl, 37°) DNP-aspartic acid (R (the rate of the movement of the band compared to that of the solvent) = 0.3; 0.2 mole) and DNP-aspartylalanine (R = 0.02; 0.1 mole) were isolated on an ether column at pH 4. The major DNP-aspartyl derivative (approximately 0.4 mole), however, moved much more rapidly than either of these two derivatives (R = 0.13, chloroform, pH 4; R = 1.0, chloroform-ether 1:1, pH 4; R = 0.3, ethyl acetate, pH 6.5). This band tended to split incompletely on some columns, but both fractions behaved similarly on other columns and on hydrolysis. Complete hydrolysis of this band and the very slow band gave DNP-aspartic acid and alanine in equimolecular amounts, but no other major component was detected in the fast band. The rate of the fast band suggested the presence of another amino acid residue bearing a second DNP group, but this has not been detected in hydrolysates. The absorption spectrum of this DNP derivative changed rapidly in 1 per cent sodium bicarbonate but was normal in N hydrochloric acid with a maximum at 350 mμ (2). DNP-asparagine has a lower R value on chloroform and ether columns than DNP-aspartic acid, and it is improbable that either the slow or the fast band was a DNP-asparaginyl derivative, since they were obtained after 18 and 23 days of hydrolysis (12 N HCl, 37°), conditions which deamidate proteins (27).

DNP bovine serum albumin gave only DNP-aspartic acid (0.59 mole) on partial acid hydrolysis (12 N HCl, 8 days, 37°) and no DNP-aspartyl peptides were detected. After 1 day DNP-aspartic acid was again the only DNP-aspartyl derivative obtained (0.42 mole) and no DNP-asparagine could be detected. DNP-asparagine itself was recovered in 41 per cent yield under similar conditions. These results suggested that the N-terminal residue was an aspartyl rather than an asparaginyl residue and that the N-terminal sequence was different from that in human serum albumin.

After cleavage of the PTC derivative, as previously described, only the
phenylthiohydantoin of aspartic acid was obtained, though in poor yield. By coupling the residue from this cleavage reaction with FDNB and identifying the ether-soluble derivatives obtained on acid hydrolysis only DNP-threonine was detected (0.25 mole). This result, together with the absence of DNP-aspartyl peptides in partial hydrolysates of DNP bovine serum albumin, suggests the N-terminal sequence aspartylthreonine.

DISCUSSION

The quantitative yield in the cleavage of PTC peptides under anhydrous conditions has only been determined indirectly by previous workers (7, 13) on the basis of the yield of the residual amino acid or peptide. With PTC proteins Landmann et al. (9) gave no estimate of yields, and other investigators (11, 21) have measured the amino acid content before and after removal of the free amino groups.

The poor yields of the phenylthiohydantoins of aspartic acid and asparagine obtained in the experiments described suggest severe limitations in the application of this method to proteins. Moreover, with several of the reagents used, some DNP-threonine and DNP-serine were also present, indicating cleavage of peptide bonds involving the amino groups of these amino acids. This is not surprising in view of the acyl migration which these amino acids readily undergo (28), a factor which contributes to the lability of such peptide bonds. Repeated treatments with these reagents would increase the number of amino groups of these amino acids and lead to equivocal results.

That the low yield of phenylthiohydantoins is due to destruction or extraction difficulties is supported by the much greater yield of amino groups of the 2nd residue exposed in the cleavage product.

It is of some interest that Waley and Watson (29) have reported difficulty (17 per cent yield) in extracting a thiohydantoin of alanine (5-methyl-2-thiohydantoin) added to solutions of acetylated insulin.

For the more limited objective of determining amino acid sequences in peptides obtained from partial acid hydrolysates of proteins, these considerations may not apply, since bonds involving the amino groups of serine and threonine are probably absent, and the modification which employs aqueous buffer solutions for the cleavage with continuous extraction of the PTH amino acids has been demonstrated (12) to give excellent recoveries.

The free amino groups of both human and bovine serum albumin are located on aspartyl rather than asparaginyl residues, since no evidence of PTH-asparagine was obtained under the conditions which gave PTH-asparagine from PTC-carboxypeptidase. For bovine albumin this conclusion is supported by the absence of DNP-asparagine after a short period of
partial hydrolysis under conditions in which 41 per cent of DNP-asparagine was recovered. It has been shown (16, 17) that the N-terminal residues of the serum albumins from different species (human, bovine, equine, and porcine) are the same, after acid hydrolysis of the DNP proteins, but the different N-terminal sequence in human and bovine serum albumins is not surprising in view of their differences in amino acid composition (25). It is interesting that this species difference involves the replacement of an alanine residue by threonine, since the species difference in bovine and porcine insulins has previously been found (30) to involve a similar replacement and has also revealed differences in the content of amino acids of related structure (31, 32).

I wish to thank Dr. P. Edman and Dr. W. A. Landmann for forwarding copies of recent papers from their laboratories in advance of publication and Dr. Emil L. Smith for his interest in this work.

SUMMARY

The thiohydantoin method for the detection of free amino groups and stepwise degradation of proteins has been applied to human and bovine serum albumins and to carboxypeptidase. The free amino groups detected were in agreement with previous findings, but the quantitative recoveries were poor. This is due to difficulties in recovering the 3-phenyl-2-thiohydantoin derivatives rather than to difficulties of cleavage.

By the 1,2,4-fluorodinitrobenzene technique, the thiohydantoin technique, and a combination of the two, the species difference in human and bovine serum albumins has been confirmed. Whereas human serum albumin has the N-terminal sequence aspartylalanine, in bovine serum albumin it is aspartylthreonine.

BIBLIOGRAPHY

THE N-TERMINAL SEQUENCE OF SERUM ALBUMINS; OBSERVATIONS ON THE THIOHYDANTOIN METHOD
E. O. P. Thompson


Access the most updated version of this article at http://www.jbc.org/content/208/2/565.citation

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/208/2/565.citation.full.html#ref-list-1