IDENTIFICATION OF C-TERMINAL RESIDUES IN PEPTIDES AND PROTEINS THROUGH FORMATION OF THIOHYDANTOINS*

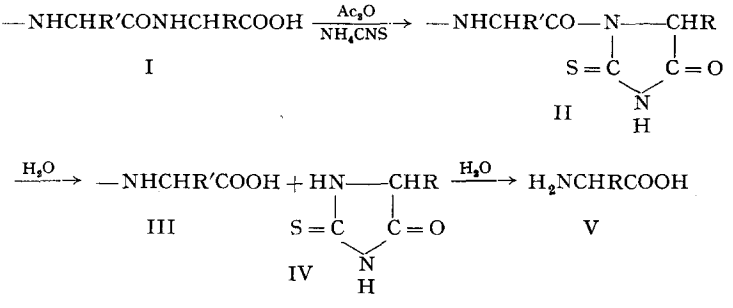
by

ROBERT A. TURNER AND GERTRUDE SCHMERZLER

Department of Biochemistry, College of Medicine, State University of New York, Brooklyn, N.Y. (U.S.A.)

INTRODUCTION

Although there are several methods available for the determination of residues bearing free amino groups in peptide chains, there are few methods applicable to the determination of C-terminal residues, and even these are of uncertain usefulness[1,2]. In 1926 SCHLACK AND KUMPF reported the first method for identifying the C-terminal residue of a peptide, which consisted in its conversion into a 1-acyl-5-alkyl-2-thiohydantoin II, scission of the latter to yield a 5-alkyl-2-thiohydantoin IV, and isolation



of IV[3]. The N-benzoyl derivatives of certain simple peptides were favoured for studying the method. By isolation of the peptide fragment III and subjection of it to their procedure, SCHLACK AND KUMPF showed that stepwise degradation was possible in some instances. The method was used by NICOLET to elucidate the structure of glutathione[4].

More recently other methods of identifying C-terminal residues have been reported: the formation of acylureas by reaction of the peptide with di-4-tolyl-carbodiimide[5]; the reduction of the C-terminal residue to an α-amino alcohol by means of lithium aluminum hydride[6,7] or sodium boron hydride[8]; the anodic oxidation of the C-terminal residue in methanol to an amino-methoxy residue that may be hydrolyzed to an aldehyde[9]. While the last of these may prove to be quite valuable, none of them appeared to be of such wide application nor to have such a favourable chemistry as the thiohydantoin method. Besides the investigations already cited[3,4], in recent years others have reported procedures using the thiohydantoin method, notably TIBBS[10], and WALEY

* This investigation was supported by a grant from the National Science Foundation.

References p. 559.

AND WATSON[11]. While our investigation was in progress, BAPTIST AND BULL published their results on the recovery of amino acids after hydrolysis of the thiohydantoins IV derived from acylamino acids and from peptides[12]. A thorough study of thiohydantoins and their formation from amino acids has been reported by SWAN[13,14,15].

In developing a method of degradation, suitable conditions for treatment of proteins are desirable, but a more important aim is to achieve a reliable method for small peptides since there is a practical limit to stepwise degradation of a large chain. At this time the only means of elucidation of an entire protein structure are the identification of the fragmentary peptides obtained through partial hydrolysis, and the deduction of the amino acid sequences of the protein from those found in the fragmentary peptides.

EXPERIMENTAL

Materials

The peptides were synthesized as described by TURNER[16] or were purchased. Ovomucoid was purchased from the Worthington Biochemical Sales Corp.

Degradative Procedure for Peptides

About 30 mg of peptide dissolved in 3 to 4 ml of a solution of 90% acetic anhydride in acetic acid was treated with 15 mg of ammonium thiocyanate and heated for thirty minutes on a steam bath. The mixture was shaken seven times with 8 ml portions of petroleum ether, which was drawn off each time with a pipette. The gummy residue was dissolved in sufficient 0.4 *N* barium hydroxide to make the pH 12.5 or higher. When the solution had stood for ninety minutes, it was acidulated with sulfuric acid to pH 6.5. Barium sulfate was removed by centrifugation. After several extractions of the solution with 10 ml portions of ethyl acetate, the united extracts were concentrated nearly to dryness. The remaining material was diluted with water to a volume of 15 ml and extracted five times with ethyl acetate as before. The united extracts were dried over magnesium sulfate and concentred to dryness in a small combustion tube with the aid of a stream of nitrogen.

Degradative Procedure for Proteins

Proteins were handled in a slightly different manner in some cases in order to effect as much solution of the protein as possible in the acetic acid[17]. For example, ovalbumin was added in small portions, with agitation, to the cold acetic acid solution. A fraction of the protein, apparently denaturated, formed gummy, insoluble particles that did not dissolve during the period of heating.

Hydrolysis of Thiohydantoins

a. *With Hydrobromic Acid.* The residue of thiohydantoin in the combustion tube was heated at 150° for six hours with 0.2 ml of 48% hydrobromic acid. The contents of the tube were diluted with water, filtered, and distilled to dryness *in vacuo*. Addition of water and distillation to dryness were repeated alternately three times. The residue was dissolved in 0.3 ml of water for use in paper chromatography.

b. *With Barium Hydroxide.* The residue in the combustion tube was heated at 140° for five hours with 3.0 ml of 1.2 *N* barium hydroxide. After the tube was opened, a stream of carbon dioxide was passed through its contents until they were approximately neutral. Removal of the barium carbonate was effected by filtration, and the filtrate was concentrated before use in paper chromatography.

Penultimate Residues

In order to identify the amino acid moiety next to the C-terminal one, the aqueous solution remaining after the first extractions with ethyl acetate was lyophilized, and the residue was subjected to the procedure described above.

Paper Chromatography

Following hydrolysis of the thiohydantoin IV, the terminal amino acid V was identified by means of paper chromatography. The most frequently used solvents were *n*-butanol:acetic acid:water in volume proportions 4:1:5; and phenol:water, 4:1 (atmosphere of ammonia). In order to distinguish Leu from Ileu, the irrigating solvent tertiary butyl alcohol:butanone:water:88% formic acid in volume proportions 160:160:39:1, was employed[18].

References p. 559.

Insulin

Two experiments were performed with insulin. In the first, 50 mg of insulin was subjected to the procedure described for peptides. The protein dissolved readily and completely. After hydrolysis with hydrobromic acid and paper chromatography a single spot of Ala was revealed.

In the second experiment, 50 mg of insulin and 20 mg of ammonium thiocyanate were dissolved in 4 ml of a 90% solution of acetic anhydride in acetic acid and allowed to remain at room temperature for two days. (It has been shown that formation of an acyl-thiohydantoin II takes place at room temperature[19].) After scission of the acyl-thiohydantoin in the usual way, the deacylated thiohydantoin was hydrolyzed with barium hydroxide. Paper chromatography revealed a weak spot of *aspartic acid* and a stronger spot of *alanine*.

Deacylation Studies

In the cases of certain peptides it was found that when the thiohydantoin was extracted in one series of extractions instead of being extracted in the double series described above, all of the component amino acids in the peptide appeared on the chromatogram rather than the C-terminal amino acid alone. The following experiment exemplifies this.

Phthaloyl-glycyl-phenylalanine was converted to the acyl-thiohydantoin in the usual way in three tubes. In tube *A*, after the extractions with petroleum ether, deacylation was performed with concentrated ammonia solution at room temperature for one hour. In tube *B* deacylation was performed with 1 *N* sodium hydroxide for one hour. In the cases of both *A* and *B*, after neutralization with acid, the thiohydantoin was extracted with several portions of ethyl acetate. The combined extracts were dried and evaporated. The residue was hydrolyzed with hydrobromic acid. In tube *C* deacylation was performed with 1 *N* sodium hydroxide as for the case of tube *B*, but the thiohydantoin was extracted in two series in the manner given under the procedure for peptides (*vide supra*). Paper chromatography of the acid hydrolyzates gave the following results: from tube *A*, spots of Phe and Gly, approximately equal in intensity; from tube *B*, a spot of Phe and a much weaker spot of Gly; from tube *C*, a single spot of Phe.

RESULTS AND DISCUSSION

The C-terminal residues, and some C-penultimate residues, found for several peptides and proteins through the use of the thiohydantoin method, are listed in the Table.

It has been established that a 1-acyl-5-alkyl-2-thiohydantoin II is hydrolyzed to a 5-alkyl-2-thiohydantoin IV in a few minutes in 0.1*N* alkali[11,20]. Deacylation may also be performed with hot, dilute hydrochloric acid[10,12]. The use of dilute alkali at room temperature appeared less drastic, and as it has been shown in one case that hydrolysis was practically complete in one hour in 0.01*N* alkali[11], the procedure followed here has been to allow the acyl-thiohydantoin to stand for ninety minutes in barium hydroxide solution at pH 12.5 or somewhat higher. Concentrated ammonia solution was found to be unsatisfactory; it would be so, in any event, when stepwise degradation was contemplated, for the peptide fragment III is not formed, but rather its *amide*[3].

The double series of extractions of the thiohydantoin IV frees it from substances that tend to be extracted with it. Obviously, if the peptide fragment III detached during deacylation is extracted with the thiohydantoin IV, the result is the same as if deacylation had not taken place, and the acyl-thiohydantoin II had been extracted. The cases which offered difficulty when only a single series of extractions was performed were those that yielded a peptide III having some solubility in ethyl acetate. When III was a once degraded protein chain, no difficulty was encountered. Solubility may not be the only property involved, for thiohydantoins are weak acids (p*K* 9.1)[21] and may tend to associate with peptides. In the case of phthaloylglycyl-phenylalanine, when only one series of extractions was performed, glycine as well as phenylalanine appeared on the chromatogram. This observation may be explained as due to some extraction of phthaloyl-glycine along with the thiohydantoin. However, the double series of extractions led to a cleaner chromatogram which had a single spot of phenylalanine.

References p. 559.

TABLE I

THE C-TERMINAL RESIDUES AND C-PENULTIMATE RESIDUES OF PEPTIDES AND PROTEINS AS REVEALED BY THE THIOHYDANTOIN METHOD

Substances	*Residues found**
Peptides	
Leucyl-glycine	Gly (Leu)
Alanyl-methionine	Met**
Carnosine	His
Phthaloyl-β-alanyl-asparagine	Asp
Phthaloyl-β-alanyl-leucine	Leu
Phthaloyl-β-alanyl-tyrosine	Tyr
Phthaloyl-glycyl-histidine	His
Phthaloyl-glycyl-phenylalanine	Phe
Phthaloyl-β-alanyl-serine	O
Leucyl-glycyl-glycine	Gly
Glycyl-phenylalanyl-glycine	Gly (Phe)
Leucyl-prolyl-glycine	Gly (Leu)
Glutathione	Gly, Glu
Proteins	
Bacillomycin B	Tyr
Insulin	Ala, Asp**
Ovalbumin	Ala
Ovomucoid	Phe (Leu)

* Hydrolyses of thiohydantoins were performed with hydrobromic acid unless otherwise indicated. C-penultimate residues are given in parentheses.

** Hydrolysis of the thiohydantoin was performed with barium hydroxide.

Among the dipeptides, alanyl-methionine gave a satisfactory result only when barium hydroxide was used to hydrolyze the thiohydantoin. When hydrobromic acid was used several spots appeared on the chromatogram. A similar result was obtained when methionine itself was heated with hydrobromic acid under the hydrolytic conditions. In both instances there was considerable decomposition.

After subjection to the thiohydantoin procedure phthaloyl-β-alanyl-serine gave no spot on the chromatogram. SWAN has been unable to prepare a thiohydantoin from serine[13]. BAPTIST AND BULL identified a small amount of alanine when serine was used in their procedure[12].

In the case of glutathione glycine ought to be revealed as a C-terminal residue, while glutamic acid might be revealed under certain conditions. After detachment of the acyl residue the alkaline solution would be expected to contain the thiohydantoin of glycine and N,N′-bis-(β-(2-thiohydantoin-5-yl)-propionyl)-cystine. If the latter substance were soluble in ethyl acetate, then the chromatogram would be expected to reveal spots of glutamic acid and cystine as well as glycine. Under our conditions glycine was identified, with a faint spot of glutamic acid.

When glycyl-phenylalanyl-glycine was subjected to the procedure, glycine was revealed on the chromatogram; after a second degradation phenylalanine was revealed. The case of leucyl-prolyl-glycine was not as satisfactory, for the first degradation revealed glycine; the second, leucine. The method cannot succeed when proline is the C-terminal residue since proline cannot form a thiohydantoin. Hence the second degradation might have been expected to reveal no amino acid at all. The fact that leucine was revealed indicates that the leucyl-prolyl bond was hydrolyzed to a considerable degree

References p. 559.

during the first degradation. Since the reaction conditions are mild, it appears that the leucyl-prolyl bond is unusually susceptible to hydrolysis*.

In applying the thiohydantoin method to the identification of the C-terminal residues of proteins, it has been found that in ovomucoid the C-terminal residue is phenylalanine[17]. The same residue has been identified by reduction of the protein with lithium aluminum hydride[6]. When the thiohydantoin procedure was repeated on the once degraded ovomucoid, the residue revealed was leucine. Thus the terminal sequence is possibly leucyl-phenylalanine; but since a prolyl residue, for example, might stand between the leucine and phenylalanine residues and yet lead to the same result, the sequence must be demonstrated in a partial hydrolyzate before it can be considered more than tentative.

In our earlier report on the thiohydantoin method it was stated that ovalbumin yielded alanine as the C-terminal residue[17]. It had been reported, from studies of ovalbumin with carboxypeptidase, that alanine was at least one of the C-terminal residues[25]. However, in these studies di*iso*propylfluorophosphate was not used to suppress the action of proteolytic enzymes in the sample of carboxypeptidase. Since in subsequent experiments with di*iso*propylfluorophosphate no C-terminal residue was revealed, it appears that there is no C-terminal residue in native ovalbumin**. Since the denaturation of ovalbumin in warm acetic acid has been observed, the alanine revealed by the thiohydantoin method is doubtless the C-terminal residue of the denatured protein. In the first studies on ovalbumin with carboxypeptidase in the absence of d*iso*propylfluorophosphate, alanine appeared by itself during the early stages of the incubation; subsequently other amino acids appeared together rather than serially[25]. If it is postulated that the proteolytic enzymes functioned in denaturing or unfolding the ovalbumin molecule, so that unexposed C-terminal residue became accessible to the carboxypeptidase, the results of the experiments with carboxypeptidease in the absence of di*iso*propylfluorophosphate and the results from the application of the thiohydantoin method are in harmony. At the present state of our knowledge it appears that while native ovalbumin has no accessible C-terminal residue, the denatured protein has a single C-terminal residue, viz., alanine.

Application of the thiohydantoin method to a sample of crystalline zinc insulin (Lilly) revealed spots of alanine and aspartic acid although the spot of aspartic acid was quite weak. Previously, WALEY AND WATSON, using the thiohydantoin method, had found alanine as the C-terminal residue, with a trace of a substance having the same R_F value as aspartic acid[11]. In view of the work of SANGER AND THOMPSON, who showed that the *A* chain of oxidized insulin gave cysteyl-asparagine as a fragment after enzymatic hydrolysis[26]; and the work of HARRIS, who obtained asparagine by hydrolysis of insulin with carboxypeptidase[27], it was important to demonstrate the presence of as-

* N-acyl derivatives of N-heterocyclic compounds are known to be easily deacylated. Like 1-acyl-thiohydantoins, 1-acyl-imidazoles are rapidly hydrolyzed[22]. Since several peptides having proline as the N-terminal residue have been found in a partial hydrolyzate of vasopressin[23], the facile release of the acyl moiety may be general. The selective hydrolysis of proline peptides is worthy of further study[24].

* The authors thank Prof. C. FROMAGEOT and Dr. D. STEINBERG for private communications on this subject. Since the submission of the manuscript to the editor a communication by STEINBERG has been published (*J. Am. Chem. Soc.*, 75 (1953) 4875) in which the author withdraws his earlier claim that alanine is C-terminal in *native* ovalbumin, and postulates that the alanine residue becomes available only after a preliminary opening of the protein molecule which is catalyzed by a contaminating enzyme in the carboxypeptidase preparation.

References p. 559.

paragine as a C-terminal residue by means of a non-enzymatic method. The thiohydantoin method might be expected to be peculiarly appropriate to this problem, as will presently be shown.

Aspartic and glutamic acids do not form thiohydantoins, presumably because under the reaction conditions they rapidly form cyclic anhydrides that are incapable of reacting with salts of thiocyanic acid[13,14]. A terminal residue of isoasparagine or of aspartic acid diamide could not form a thiohydantoin, but a terminal residue of asparagine can form a thiohydantoin, as demonstrated above by the case of phthaloyl-β-alanyl-asparagine. Thus, of the several possibilities, aspartic acid, isoasparagine, aspartic acid diamide, and asparagine, only the last of these will give a spot of aspartic acid in the thiohydantoin procedure. Hence the revealing of a spot of aspartic acid shows that asparagine is a C-terminal residue of insulin. While this confirms the results of others[26,27], we are unable to account for the failure to obtain approximately equal amounts of alanine and aspartic acid, which would be expected on account of what is now known of the structure of insulin.

The revealing of asparagine as one of the C-terminal residues of insulin was accomplished by hydrolysis of the thiohydantoin intermediate with barium hydroxide. In another experiment hydrobromic acid was employed for the hydrolysis, and only alanine was revealed. This result confirms that of WALEY AND WATSON, who employed hydrobromic acid[11].

In relation to investigations of bacillomycin B, a fungistatic agent obtained from *Bacillus subtilis*[28,29], it has been found that this polypeptide contains five amino acids: aspartic acid, glutamic acid, serine, threonine, and tyrosine. The thiohydantoin method has revealed tyrosine as the C-terminal residue.

CONCLUSIONS

The thiohydantoin method is applicable to the stepwise degradation of peptides and proteins on a small scale. In hydrolyzing the thiohydantoins of terminal residues hydrobromic acid is usually satisfactory; in some cases barium hydroxide is a much better reagent and causes less decomposition of the amino acid. While the method is generally successful for most of the common, neutral amino acids, it fails when the C-terminal residue is serine or proline; or when, as other investigators have found, the residue is aspartic acid[13], glutamic acid[13], lysine[12], or arginine[12]. However, asparagine is detectable.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. GEORGE W. ANDERSON of the American Cyanamid Co. for samples of tripeptides; Dr. OTTO K. BEHRENS of Eli Lilly & Co. for a sample of insulin; and Dr. JOSEPH SEIFTER of the Wyeth Institute of Applied Biochemistry for a sample of bacillomycin. It is a pleasure to acknowledge helpful discussions with Dr. JOHN M. SWAN.

SUMMARY

The thiohydantoin method has been applied to the stepwise degradation of peptide chains from the carboxyl end. In the present procedure, the thiohydantoin intermediate formed from the peptide, ammonium thiocyanate, and acetic anhydride, was deacylated with barium hydroxide; the thiohydantoin was separated from other substances by means of a double extraction and then hydrolyzed

References p. 559.

to form the C-terminal amino acid. In some cases the once degraded peptide chain was subjected again to the procedure in order to reveal the C-penultimate residue of the original peptide. Through application of the method to proteins the C-terminal residues of insulin, ovomucoid, ovalbumin, and bacillomycin B have been revealed. The limitations of the method have been discussed.

RÉSUMÉ

La méthode à la thiohydantoïne a été utilisée pour la dégradation progressive des chaînes peptidiques à partir de leur extrémité carboxylique. Dans les procédés actuels, la thiohydantoïne intermédiaire, formée par réaction du peptide avec le thiocyanate d'ammonium en présence d'anhydride acétique, est désacylée par la baryte. La thiohydantoïne est alors séparée des autres substances au moyen d'une double extraction puis est utilisée en donnant naissance à l'acide aminé C-terminal. Dans quelques cas la chaîne peptidique résiduelle a été soumise au même traitement, ce qui a permis de caractériser le résidu C-pénultième de la chaîne originale. L'application de cette méthode a mis en évidence la nature des résidus C-terminaux de l'insuline, de l'ovomucoïde, de l'ovalbumine et de la bacillomycine B. Les auteurs discutent les limites de la méthode.

ZUSAMMENFASSUNG

Das Thiohydantoin-Verfahren ist zum schrittweisen Abbau von Peptidketten vom Carboxyl-Ende her angewendet worden. In der vorliegenden Arbeit wurde das intermediäre Thiohydantoin, hergestellt durch Einwirkung von Ammoniumthiocyanat und Essigsäureanhydrid auf das Peptid, mit Hilfe von Bariumhydroxyd entazyliert. Das Thiohydantoin wurde von den übrigen Substanzen durch zweifache Extraktion getrennt und darauf hydrolysiert, um die C-entständige Aminosäure freizusetzen. In verschiedenen Fällen wurde die einmal abgebaute Peptidkette nochmals dem Verfahren unterworfen, um die nächstfolgende Aminosäure zu charakterisieren. Beim Anwenden dieser Methode auf Proteine sind die C-entständigen Aminosäurereste des Insulins, Ovomucoids, Ovalbumins und des Bacillomyzins B identifiziert worden. Die Anwendungsmöglichkeiten des Verfahren werden besprochen.

REFERENCES

1 H. G. KHORANA, *Quart. Rev.*, 6 (1952) 340.
2 P. DESNUELLE, in *Advances in Enzymology*, Vol. 14, Interscience Publishers, New York, 1953.
3 P. SCHLACK AND W. KUMPF, *Z. physiol. Chem.*, 154 (1926) 125.
4 B. H. NICOLET, *J. Biol. Chem.*, 88 (1930) 389, 395, 403.
5 H. G. KHORANA, *J. Chem. Soc.*, (1952) 2081.
6 L. PÉNASSE, M. JUTISZ, C. FROMAGEOT AND H. FRAENKEL-CONRAT, *Biochim. Biophys. Acta*, 9 (1952) 551.
7 C. FROMAGEOT, M. JUTISZ, D. MEYER AND L. PÉNASSE, *Biochim. Biophys. Acta*, 6 (1950) 283.
8 A. C. CHIBNALL AND M. REES, *Biochem. J.*, 48 (1951) xlvii.
9 R. A. BOISSONNAS, *Nature*, 171 (1953) 304; *Helv. Chim. Acta*, 35 (1953) 2226.
10 J. TIBBS, *Nature*, 168 (1951) 910.
11 S. G. WALEY AND J. WATSON, *J. Chem. Soc.*, (1951) 2394.
12 V. H. BAPTIST AND H. B. BULL, *J. Am. Chem. Soc.*, 75 (1953) 1727.
13 J. M. SWAN, *Austr. J. Sci. Res.*, A5 (1952) 711.
14 J. M. SWAN, *Austr. J. Sci. Res.*, A5 (1952) 721.
15 J. M. SWAN, *Austr. J. Sci. Res.*, A5 (1952) 728.
16 R. A. TURNER, *J. Am. Chem. Soc.*, 75 (1953) 2388.
17 R. A. TURNER AND G. SCHMERZLER, *Biochim. Biophys. Acta*, 11 (1953) 586.
18 R. J. BLOCK AND H. B. VAN DYKE, *Arch. Biochem. Biophys.*, 36 (1952) 1.
19 Personal communication from Dr. J. M. SWAN.
20 A. KJAER AND P. ERIKSEN, *Acta Chem. Scand.*, 6 (1952) 448.
21 V. DU VIGNEAUD AND D. B. MELVILLE, in *The Chemistry of Penicillin*, Princeton Univ. Press. (1949) p. 288.
22 M. BERGMANN AND L. ZERVAS, *Biochem. Z.*, 203 (1928) 284.
23 R. ARCHER, J. CHAUVET AND P. FROMAGEOT, *Biochim. Biophys. Acta*, 9 (1952) 471.
24 S. J. LEACH, *Rev. Pure Appl. Chem.* (Austr.), 3 (1953) 25.
25 D. STEINBERG, *J. Am. Chem. Soc.*, 74 (1952) 4217.
26 F. SANGER AND E. O. P. THOMPSON, *Biochem. J.*, 53 (1953) 366.
27 J. I. HARRIS, *J. Am. Chem. Soc.*, 74 (1952) 2944.
28 M. LANDY, G. H. WARREN, S. B. ROSENMAN AND L. G. COLIO, *Proc. Soc. Exptl. Biol. Med.*, 67 (1948) 539.
29 H. TINT AND W. REISS, *J. Biol. Chem.*, 190 (1951) 133.

Received October 2nd, 1953