[CONTRIBUTION FROM THE BIOCHEMICAL RESEARCH DEPARTMENT OF THE ARMOUR LABORATORIES]

Paper Chromatography of the 3-Phenyl-2-thiohydantoin Derivatives of Amino Acids with Application to End Group and Sequence Studies

BY WENDELL A. LANDMANN, MAURICE P. DRAKE AND JANIS DILLAHA

RECEIVED JANUARY 5, 1953

A paper chromatographic method for identifying *per se* the 3-phenyl-2-thiohydantoin derivatives of amino acids is presented, by which N-terminal residues and amino acid sequences of peptides and proteins may be determined conveniently. The chromogenic agent is Grote's solution which gives a variety of colors with the thiohydantoins. This color difference, together with R_f values, makes possible the identification of most of the amino acid derivatives with two unidimensional chromatograms. The reported N-terminal residues of salmine, glutathione, B-lactoglobulin, insulin and lysozyme have been confirmed by this method. By stepwise removal of phenylthiohydantoins and subsequent identification chromatographically, it has been possible to confirm the sequence of the A and B chains of insulin as far as the fifth positions. The sequence of the first five amino acids in lysozyme has also been obtained and appears to be lysyl-valyl-phenylalanyl-glycyl-seryl-.

The method for the determination of the amino acid sequence in peptides, reported by Edman[1] in 1950 has been employed quite successfully to identify the N-terminal amino acids.[1,2] In this method the protein or peptide is allowed to react with phenyl isothiocyanate and the N-terminal amino acid is then removed from the peptide chain as the corresponding 3-phenyl-2-thiohydantoin. Hydrolysis of the thiohydantoin liberates the amino acid which may be subsequently identified chromatographically. An alternate procedure of identifying the amino acid consists of quantitatively determining the amino acids present in a peptide prior to and after removal of the terminal amino acid as the thiohydantoin, and noting the residue which disappears.[3]

Direct identification of the 3-phenyl-2-thiohydantoin derivative possesses several distinct advantages over the original procedure. It eliminates the time-consuming hydrolysis of the thiohydantoin, and reduces to a minimum the manipulations inherent in the original method, particularly with respect to end group analysis. It also overcomes the discrepancies which may arise as a result of the hydrolysis of any peptide material carried along with the thiohydantoin, since amino acids and peptides produce no color with the reagent used for the detection of the derivatives.

A mixture of the 3-phenyl-2-thiohydantoins may be resolved on buffered filter paper (Whatman #1) by the use of two separate unidimensional chromatograms. Both the paper and solvents are saturated with a phthalate buffer (Clarke and Lubs, *p*H 6). The two solvent systems consist of xylene–acetic acid–*p*H 6 buffer (3:2:1) and *s*-butyl alcohol–*p*H 6 buffer (7:1). A standard mixture of reference compounds was always included on each chromatogram, after numerous runs on each compound had been carried out to establish R_f values. The R_f values as determined for each derivative are listed in Table I.

It will be noted that several of the synthetic thiohydantoins give multiple spots. These were confirmed in the case of leucine, phenylalanine and proline by the use of peptides whose N-terminal residues were known. All reference compounds were prepared according to Edman[1] and their identity checked by melting point determinations and by the identification of the amino acid obtained upon hydrolysis of the derivative.

TABLE I

R_f VALUES OF THE 3-PHENYL-2-THIOHYDANTOINS OF AMINO ACIDS

PTH[a] amino acid	Xylene:acetic:buffer			2-Butanol: buffer	Color with Grote's reagent
Ala	0.08			0.78	Blue
Asp	.03			.13	Blue
Arg	.00			.44	Blue-violet
Cys	.00			.00	Bluish
Gly	.30			.78	Red
Glu	.00			.29	Blue
His	.00			.78	Yellow
Iso	*.47*	0.90[b]		.92	Blue
Leu	*.41*	0.90		.92	Blue
Lys	.49			.92	Blue-violet
Met	.23			.92	Blue
Phe	.30	0.83	*0.90*	.92	Blue, *yellow*[b]
Pro	.45	*0.87*		.92	Blue, *purple*
Ser	.81			.92	Red
Thr	.77			.92	Blue
Tyr	.10			.92	Yellow
Try	.50			.92	Yellow
Val	.30	*0.85*		.92	Blue

[a] Phenylthiohydantoin. [b] The definitive spot and its color are in italics for each compound showing multiple spots in the xylene system.

Since some difficulty was experienced in the preparation of the serine, threonine and cystine derivatives, the reference compounds are probably not the thiohydantoin derivatives expected. However, repeated preparations of these compounds consistently gave the same products, as evidenced by their chromatographic behavior on paper. These products presumably would be obtained when their parent amino acids occur as peptide end groups, and may therefore be used in their identification. This view has been supported in the case of a naturally occurring peptide having serine as its terminal residue, previously established by the DNP method. When this peptide was treated with phenyl isothiocyanate in the manner described, the cleavage product obtained was identical chromatographically with that given by serine alone.

A few troublesome combinations occur, but these present little difficulty with single compounds. For example, PTH-lysine and tryptophan have the

(1) P. Edman, *Acta Chem. Scand.*, **4**, 277 (1950).

(2) H. Fraenkel-Conrat and J. Fraenkel-Conrat, *Federation Proc.*, **11**, 214 (1952).

(3) S. W. Fox, T. L. Hurst and K. F. Itschner, THIS JOURNAL, **73**, 3573 (1951).

same R_f values, but can be distinguished by color with Grote's solution, since the tryptophan derivative is yellow. PTH-glycine may also be distinguished from the interfering spots of the hydantoins of phenylalanine and valine by its red color with Grote's solution. Of the PTH-amino acids giving multiple spots, namely, isoleucine, leucine, proline, phenylalanine and valine, PTH-phenylalanine shows a characteristic yellow spot at R_f 0.90 while proline gives a purple spot at R_f 0.87. The leucine and isoleucine compounds may each be recognized by their second spots which occur at R_f 0.41 and 0.47, respectively. Only the valine derivative in the presence of PTH-phenylalanine remains indistinguishable. Consequently, hydrolysis of the thiohydantoins need be carried out only if this combination is indicated.

No interference from the sulfur-containing amino acids or the phenylthioureides of amino acids or peptides has been observed in using Grote's reagent, since very little, if any, color is produced by these compounds. The reagent as employed here is capable of producing a readily detectable color with five micrograms of the phenylthiohydantoins, and is even more sensitive for derivatives such as PTH-glycine, which give a high intensity of color with the reagent.

In order to evaluate the method, several pure peptides and proteins were treated with phenyl isothiocyanate, hydrolyzed and the resulting thiohydantoins extracted with ether. The chromatographic technique just outlined was applied in the identification of the N-terminal residues. The results so obtained are summarized in Table II and are in excellent agreement with those of other investigators.[4]

TABLE II

N-TERMINAL RESIDUES OF PEPTIDES AND PROTEINS By Chromatography of Phenylthiohydantoins (Compared to Reported Residues)

Peptides	N-Terminal amino acids Reported	Found
Leucylglycine	Leucine	Leucine
Salmine[a]	Proline	Proline
Insulin, bovine[b]	Phenylalanine	Phenylalanine
	Glycine	Glycine
Insulin, B-chain	Phenylalanine	Phenylalanine
B-Lactoglobulin[b]	Leucine	Leucine
Lysozyme[b]	Lysine	Lysine
Glutathione	Glutamic acid	Glutamic acid

[a] Paul Lewis Laboratories. [b] Armour Laboratories.

In addition to its use in the identification of end groups, the direct chromatography of phenylthiohydantoins offers a sensitive and convenient means of establishing the purity of peptide preparations, since a single peptide chain should reveal only one end group.

Application to Sequence Studies.—By modifying the procedure for the cleavage of the phenylthiohydantoin from the peptide it becomes possible to apply the chromatographic technique to a study of the sequence of amino acids in peptides. For this work, the A and B chains of insulin were prepared according to Sanger,[5] and subjected to the treatment described later.

(4) F. Sanger and R. R. Porter, *Adv. Protein Chem.*, **5**, 98 (1949).

(5) F. Sanger, *Biochem. J.*, **44**, 126 (1949).

The first five amino acids in each chain of insulin have been identified. The sequences have agreed with the findings of Sanger in every respect, except that asparagine and glutamine have not been fully distinguishable from aspartic and glutamic acids, respectively, by the method used here.

Crystalline isoelectric lysozyme from egg white was also carried through the procedure to determine the sequence of amino acids from the amino end of the polypeptide chain. The results appear in Table III.

TABLE III

AMINO ACID SEQUENCE IN LYSOZYME

Pos.	R_f values Xyl–HOAc	2-But.	Hydrolyzed PTH	Res.
1	0.49	0.92	Lys	Lys
2	.32, 0.85	.92	Val	Val
3	.90 (Yel)	.92 (Yel)	Phe	Phe
4	.30 (Red)	.78 (Red)	Gly	Gly
5	.80 (Red)	.92 (Red)	Trace of Ala	Ser

It is of interest to note that the first four residues correspond with those obtained by Schroeder,[6] and that the fifth residue is serine as was predicted.

As shown in Table III, the amino acid obtained as the hydrolysis product of the serine derivative is not serine and cannot be used to identify this amino acid. Similar difficulties are encountered when threonine and cystine derivatives are hydrolyzed. By direct chromatography of their phenyl isothiocyanate derivatives, however, it is possible to distinguish these residues from other amino acids.

Under the conditions of reaction employed here, there appears to be no cleavage of peptide bonds other than the one involved in the formation of the thiohydantoin. Thus it should be possible, with sufficient starting material and with suitable means of selectivity breaking the disulfide bonds of cystine, to continue down the peptide chain of a protein.

Experimental

Treatment of Peptides for End Group Determination.—Most peptides were allowed to react in 50% pyridine solution at 30–40° with a twenty molar excess of phenyl isothiocyanate. The *p*H of the reaction mixture was adjusted to 8–9 with triethylamine. In the case of lysozyme, the reaction was carried out at *p*H 10–11. After the reaction had proceeded for one hour, the solution was evaporated to dryness over sulfuric acid in a vacuum desiccator. The resulting residue was washed with benzene, followed by absolute ethanol and ether to remove excess reagent and traces of pyridine. In order to cleave the terminal residue, the dried phenylthiocarbamyl derivative was then hydrolyzed in 2 *N* hydrochloric acid at 70° for two hours. The phenylthiohydantoin was subsequently extracted from the acid solution with ether. The ether solution, after being dried over sodium sulfate, was concentrated to small volume and applied to the paper in portions of 5 μl. each. Quantities of peptide as small as one milligram have been treated in this manner with satisfactory results. It should be emphasized that contact with rubber stoppers or tubing is to be avoided during all stages of these operations.

Treatment of Peptides for Sequence Study.—Each peptide was treated as before with phenyl isothiocyanate in a pyridine–water solution, using triethylamine to maintain the *p*H at 8–9. After the reaction was complete, the mixture was evaporated to dryness *in vacuo* and the residue washed with benzene, alcohol and ether. The cleavage of the terminal residue was accomplished with hydrogen chloride in anhydrous dioxane,[3] the reaction being conducted in a closed flask at room temperature over a period of two hours. Subsequent evaporation, *in vacuo*, of the dioxane–hydrogen

(6) W. A. Schroeder, THIS JOURNAL, **74**, 5118 (1952).

chloride and extraction with ether of the hydantoin, left the peptide residue ready for the next cycle of treatment. A portion of the ether solution was spotted on paper for the identification of the thiohydantoin. The remainder of the solution was used for the hydrolysis in barium hydroxide solution[1] to the free amino acid for purposes of confirmation.

Chromatography.—Whatman #1 paper was found to be most suitable for the systems described. The large sheets were cut lengthwise into strips seven inches wide, which were buffered by dipping into a 0.05 *M* potassium acid phthalate–sodium hydroxide solution at *p*H 6. After the strips were dried in air, samples were applied to points three inches from one end of the paper and dried. Descending chromatography was employed.

Of the many solvent systems tried, the most satisfactory resolution was achieved with a mixture of xylene, glacial acetic acid and *p*H 6 phthalate buffer in a volume ratio of 3:2:1, respectively. The aqueous phase served as the equilibrating solvent while the organic phase was the developing solvent. After a 24-hour equilibration period, the chromatogram was allowed to develop to a length of 18 inches. At 25°, about three hours was required for development.

A second solvent system, 2-butanol–*p*H 6 phthalate buffer (7:1) was used primarily to identify the phenylthiohydantoins of arginine, aspartic acid, glutamic acid, histidine and cystine. This single phase system was used as both the equilibrating and developing solvent. After a short equilibration the chromatogram was allowed to develop for a period of four hours and attained a length of about eight inches. This length was quite satisfactory for the identification of the derivatives mentioned.

After the chromatograms had developed, the solvents were allowed to evaporate from the paper in a current of air until no trace of acetic acid or butanol remained. Grote's solution[7] was diluted with an equal volume of saturated sodium bicarbonate solution and applied to the chromatogram in the form of a spray. The phenylthiohydantoins appeared as red, blue or yellow spots after the paper was held over a steam-bath for several minutes. Since considerable fading occurred as the paper became dry, the location of each spot was marked while the paper was still damp. This procedure was facilitated by placing the chromatogram on a milk-glass plate, against which all spots were readily discernible.

Acknowledgment.—We are indebted to Dr. C. D. Bossinger for the preparation of some of the reference compounds utilized in this work.

(7) I. W. Grote, *J. Biol. Chem.*, **93**, 25 (1931). 0.5 g. of sodium nitroprusside, 0.5 g. of hydroxylamine hydrochloride and 1.0 g. of sodium bicarbonate are dissolved in 10 ml. of water. Two drops of bromine are added, the excess bromine removed by aeration, and the solution filtered and diluted to 25 ml. This stock solution is further diluted as specified for use.

CHICAGO, ILLINOIS

[CONTRIBUTION FROM THE RESEARCH LABORATORIES OF HOFFMANN-LA ROCHE, INC.]

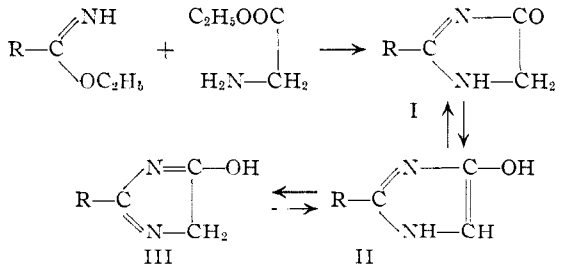
Derivatives of 4(5H)-Imidazolone

BY H. LEHR, S. KARLAN AND M. W. GOLDBERG

RECEIVED MARCH 12, 1953

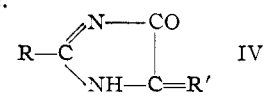
A series of new 4(5H)-imidazolone derivatives is described, obtained by the reaction of glycine ester with various imidic acid esters in the presence of a ketone. Some of the new compounds have shown hypnotic activity when tested in mice.

Only two imidazolones of the general formula I, with the oxygen atom in position 4, have so far been described, the 2-methyl-4(5H)-imidazolone (I, R = CH_3)[1] and the 2-benzyl-4(5H)-imidazolone (I, R = $C_6H_5CH_2$).[2] They were obtained by condensing glycine ester at room temperature with the ethyl esters of acetimidic or phenylacetimidic acid, respectively.[3]



These imidazolones are weak bases of limited stability and form stable hydrochlorides. Their reactions clearly indicate the existence of tautomerism. For example, the 2-benzyl-4(5H)-imidazolone gives a dibenzoyl derivative, probably derived from structure II, and a benzylidene derivative, which could be formed from I or III.[2]

The 4(5H)-imidazolones were of interest to us in connection with studies on new centrally active substances, and their synthesis was, therefore, reinvestigated. We were indeed able to obtain the two compounds described by Finger,[1,2] after modifying slightly his preparation method (cooling of the reaction mixture), but the yields were rather low, and it soon became apparent that the synthesis of new compounds of this series would require an improved method. In an attempt to find one, the condensation of glycine ester with various imidic acid esters was also carried out in the presence of solvents, such as benzene, dioxane and acetone. A new series of crystalline reaction products resulted when acetone was used. However, the new products were not the expected 4(5H)-imidazolones, but rather their 5-isopropylidene derivatives, formed from the imidazolones by a secondary condensation with acetone. Similar products were also obtained with many other aliphatic and hydroaromatic ketones, and also with ethyl acetoacetate, ethyl levulinate, acetophenone and 1-methyl-4-piperidone. Their structure is represented by the general formula IV, in which R can be an alkyl, aralkyl or aryl group (depending on the imidic acid ester used), and R′ is the radical introduced by the secondary condensation with a ketone.



The structure of the new compounds was confirmed by synthesizing one of them, 2-benzyl-5-

(1) H. Finger, *J. prakt. Chem.*, [2] **76**, 93 (1907).

(2) H. Finger and W. Zeh, *ibid.*, [2] **82**, 50 (1910).

(3) Condensation at elevated temperatures leads to other products of not yet determined structure (Finger's isoglyoxalidones).