While this work was in progress an outline of a fluorimetric method for the determination of *N*-methyl-2-pyridone-5-carboxylamide in urine was reported by Rosen, Perlzweig & Handler (1948). The authors considered that their values for normal urines (5·0–13·2 mg./day) may have been somewhat low, but it is obvious that they are of the same order as those obtained by the present colorimetric method (see Table 1, and also Dean & Holman, 1949).

SUMMARY

1. Methods are described for the determination of *N*-methyl-2-pyridone-5-carboxylamide and *N*-methyl-2-pyridone-3-carboxylamide in human urine. The methods are based on the nitration of the compounds, after isolation from urine by adsorption on Lloyd's reagent followed by extraction with chloroform, and on their colorimetric determination by means of the yellow colours given by the nitro derivatives in alkaline solution.

2. The accuracy and specificity of the methods is discussed.

3. Results are shown which indicate that the 2:3-amide is of no significance in nicotinic acid metabolism, but that the 2:5-amide is an important metabolite. The 2:5-amide was excreted by two normal adults in amounts of the order of 7 mg./24 hr. After the ingestion of 500 mg. of nicotinamide, the urinary excretion of the 2:5-amide increased greatly, 47–57 % of the dose being excreted in 72 hr. as the latter compound.

4. It was possible to account for 73–80 % of an orally administered dose of 500 mg. of nicotinamide in the urine of two adult subjects as *N*-methyl-2-pyridone-5-carboxylamide, *N*-methylnicotinamide and the total acid-hydrolysable derivatives of nicotinic acid.

REFERENCES

Dean, R. F. A. & Holman, W. I. M. (1949). *Nature, Lond.*, **163**, 97.
Holman, W. I. M. & Wiegand, C. (1948). *Biochem. J.* **43**, 423.
Knox, W. E. & Grossman, W. I. (1946). *J. biol. Chem.* **166**, 391.
Knox, W. E. & Grossman, W. I. (1947). *J. biol. Chem.* **168**, 363.
Najjar, V. A. & Wood, R. W. (1940). *Proc. Soc. exp. Biol., N.Y.*, **44**, 386.
Perlzweig, W. A. & Huff, J. W. (1945). *J. biol. Chem.* **161**, 417.
Perlzweig, W. A., Levy, E. D. & Sarett, H. P. (1940). *J. biol. Chem.* **136**, 729.
Rosen, F., Perlzweig, W. A. & Handler, P. (1948). *Fed. Proc.* **7**, 181.

The Terminal Peptides of Insulin

By F. SANGER (Beit Memorial Fellow)
Biochemical Laboratory, University of Cambridge

(*Received* 1 *June* 1949)

One of the outstanding problems of protein chemistry is the elucidation of the relative positions occupied by the amino-acid residues in the protein molecule. A method for the identification of those terminal amino-acids in which the α-amino group is free has already been applied to several proteins (Sanger, 1945; Porter & Sanger, 1948). The protein is treated with 1:2:4-fluorodinitrobenzene (FDNB) and submitted to complete hydrolysis. Thereafter the *N*-2:4-dinitrophenyl (DNP) derivatives of the amino-acids are separated from each other and quantitatively estimated. By this means insulin has been shown to contain glycine and phenylalanine in the terminal position, each present to the extent of two molecules per insulin molecule of molecular weight 12,000.

The method can be adapted in determining the sequence of amino-acids which occupy positions near to the terminal amino-acid residues. For this purpose the protein is submitted to partial hydrolysis leading to the liberation of a series of *N*-2:4-dinitrophenyl (DNP) peptides. These differ from the other products of partial hydrolysis in that they are acids which can be extracted from acid solution by organic solvents and can thus be obtained relatively free from other unsubstituted peptides. The peptide mixture so produced is relatively simple and can be fractionated by chromatography on silica gel. Determination of the structure of the individual *N*-2:4-dinitrophenyl peptides then reveals the order of the amino-acid residues in proximity to the free α-amino groups. The present paper describes the results obtained by applying the method to insulin. The nature and order of the amino-acid residues contiguous to the lysyl residues have also been determined.

Since insulin contains four free α-amino groups, it was concluded that the molecule is built up of four

open polypeptide chains joined together by –S–S– bridges. These –S–S– bridges could be broken by oxidation (Sanger, 1949*a*), and on subsequent separation two main products could be obtained; an acidic fraction, *A*, which had only glycyl terminal residues, and a basic fraction, *B*, which had phenylalanyl terminal residues. For the identification of peptides described below these fractions have been used in place of the intact insulin so as to reduce the complexity of the peptide mixture dealt with.

METHODS

Materials

Insulin. The insulin used in all this work and in the earlier work (Sanger, 1945, 1949*a*) was crystalline bovine insulin having a S content of 3·2%. Fractions *A* and *B* of oxidized insulin and their DNP derivatives (referred to as DNP-*A* and DNP-*B* respectively) were prepared as previously described (Sanger, 1949*a*).

Silica gel. A number of workers who have used the silica gel solvent system procedure originally described have stated in print (Consden, Gordon, Martin & Synge, 1947; Phillips & Stephen, 1948), or in private communication, that they have experienced difficulty in that the bands have moved too rapidly down the columns. This was probably due to the properties of the silica used, although the factors enumerated below may also be contributory: (i) *Unwashed chloroform.* All $CHCl_3$ should be thoroughly washed with water to remove the ethanol added as a stabilizer. If unwashed $CHCl_3$ is used, all DNP amino-acids will travel fast on a column of silica gel. (ii) *Overloading of the column.* If too high a concentration of DNP derivative is applied to a column it will form a fast-moving band which 'tails' badly. No more than about 2 μmol. of each DNP derivative should be used on a column of diameter of 1 cm. to obtain satisfactory bands. Nevertheless, the optimum amount will depend on the particular experiment. To obtain the best separations, as are needed in the identification of a DNP amino-acid by mixed chromatograms, the smallest amount (0·1–0·2 μmol.) that can conveniently be seen on the column should be used. In quantitative estimations, a higher concentration minimizes the importance of incidental losses which are probably due to irreversible adsorption on the silica. The maximum concentration that will give a satisfactory separation should be used. In the case of the DNP peptides the limit is often set by the solubility of the derivative in the moving phase. (iii) *Water content of the silica.* The *R* values (Martin & Synge, 1941) of the derivatives vary considerably with the amount of water added to the silica, a high water content causing bands to move too rapidly. The actual capacity of the silica appears to vary from batch to batch, so that the optimal amount of water is best determined by experiments with known DNP derivatives.

All these factors apply also to the fractionation of peptides, as described below, and should be considered before discarding a gel as unsuitable for use with DNP derivatives. All gels prepared in this laboratory have been found to be satisfactory, though *R* values vary considerably from one batch to another. The gels were prepared by the method of Gordon, Martin & Synge (1943) from commercial waterglass obtained in 5 gallon drums from Joseph Crosfield Ltd., Warrington. It would seem that the source of the waterglass is more important than the actual method of preparation.

Recently Middlebrook (1949) has worked out a method for fractionating the DNP amino-acids using buffered columns (Moyle, Baldwin & Scarisbrick, 1948), which may be generally reproducible on any batch of gel. With peptides we have found that the bands tend to 'tail' rather badly on buffered columns, though they were useful for peptides that could not be dissolved easily in the solvents used on unbuffered columns, and satisfactory fractionations could often be carried out.

Fractionation of DNP peptides

The methods of fractionation of the DNP peptides are essentially the same as those used for the fractionation of DNP amino-acids (Sanger, 1945; Porter & Sanger, 1948). For a preliminary extraction of the terminal DNP peptides from other amphoteric peptides the following procedure was used. The acidic solution of the partial hydrolysate was extracted three times with equal volumes of ethyl acetate, and the combined ethyl acetate solutions extracted three times with a 1% solution of $NaHCO_3$, each bicarbonate extract being washed twice in two separating funnels with ethyl acetate. In this way one has a simple three tube counter current procedure of the type used by Craig (1944). The three bicarbonate solutions were combined, acidified and extracted again using similar counter-current procedure with ethyl acetate. This final ethyl acetate extract should contain only acidic substances. After taking to dryness the peptide mixture is fractionated on a suitable column. To describe the various solvent systems used the method and codes of Tristram (1946) are followed (Table 1). It has been found that the most generally useful solvents for the separation of the acidic terminal DNP peptides are the $CHCl_3$-*n*-butanol mixtures. The chromatographic purity of each band was usually established by running it on columns with other solvent systems.

Table 1. *Solvent systems used on silica chromatograms*

Solvent systems	Stationary phase (ml./g. silica given in brackets)	Moving phase (saturated with stationary phase)
$CHCl_3$	Water (0·5)	$CHCl_3$
CB_1	Water (0·5)	$CHCl_3$-1% (v/v) *n*-butanol
CB_5	Water (0·5)	$CHCl_3$-5% (v/v) *n*-butanol
Etc.		
Ea-HCl	N-HCl (0·5)	Ethyl acetate
Ea-pH 7	3 M-phosphate buffer, pH 7 (0·6)	Ethyl acetate
M_{66}	N-HCl (0·5)	Ether-66% (v/v) methylethylketone
M_{66}-pH 6·5	3 M-phosphate buffer, pH 6·5 (0·6)	Ether-66% (v/v) methylethylketone
Etc.		

Acetone-*cyclo*hexane: prepared from 1 vol. water, 3 vol. acetone and 10 vol. *cyclo*hexane (Sanger, 1945).

Sometimes it is found that a soluble compound, after being passed through a column and taken to dryness, has

become apparently insoluble. This is believed to be due to the formation of salts, and the material can generally be brought into solution once again by adding a drop of conc. HCl and a solvent in which it is readily soluble such as ether or ethyl acetate. Such a mixture is then taken to dryness and rapidly dissolved in the required solvent. This technique has also proved useful for dissolving DNP peptides (such as *A* 5) which are themselves very insoluble.

It was found that the yellow peptides containing N^5-DNP-lysine could be fractionated from one another very successfully on suitable silica columns, on which they form sharp well-defined bands. Such bands, however, are often contaminated with other colourless peptides, which can be separated by adsorption chromatography on talc from acid solution. The DNP derivatives are adsorbed firmly on the talc, whereas the simple peptides and amino-acids usually pass straight through the column.

To prepare the column, talc (5 g.) was washed several times by decantation with N-HCl. The suspension was then poured into a chromatogram tube of 1 cm. diameter and allowed to settle under slight suction. The peptide mixture was then put on the column. The DNP derivatives were adsorbed in a band at the top. After washing with about 40 ml. N-HCl, they were eluted with a mixture of 4 parts ethanol and 1 part N-HCl, in which they moved rapidly down the column. This technique could be applied to the separated DNP-lysyl peptides or to the aqueous solution of the hydrolysate before fractionation.

When it was necessary to defer the working up of a DNP peptide fraction it was usually dissolved or suspended in dilute HCl and kept in the dark. No destruction of DNP derivatives preserved in this way has been observed.

Estimation of DNP peptides

The yellow DNP amino-acids and peptides were estimated colorimetrically using the Beckman spectrophotometer. Fig. 1 shows the spectral absorption curves for

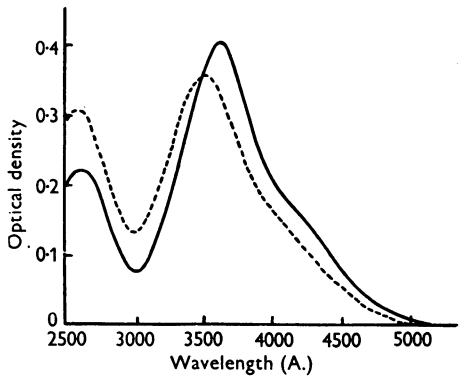

Fig. 1. Spectral absorption curves of 23·7 μM-solutions of DNP-phenylalanine (——) and DNP-phenylalanyl-valine (- - - - -) in 1% $NaHCO_3$.

DNP-phenylalanine and DNP-phenylalanyl-valine (prepared from DNP-insulin, as described below) in 1% $NaHCO_3$. It was found that the higher peptides of DNP-phenylalanine (*B* 3 and *B* 4) gave absorption curves identical with that for DNP-phenylalanyl-valine at wavelengths above 3000 A. To obtain these curves the concentration of DNP peptide was estimated from the amount of DNP-phenylalanine produced on hydrolysis. The absorption curve for DNP-glycine was almost identical with that of DNP-phenylalanine, and the curve for DNP-glycyl-glycine identical with that for DNP-phenylalanyl-valine. The DNP-glycyl peptides (*A* 4 and *A* 5) separated in the present work gave the same curve as DNP-glycyl-glycine at wavelengths >3300 A. All these DNP derivatives give identical readings at 3500 A. and the estimations have been done at this wavelength. In most experiments readings were also taken at 3900 A. and agreement between the results at the two wavelengths indicated the absence of other absorbing impurities. Table 2 shows the optical densities of 20 μM-solutions of the

Table 2. *Optical densities of* 20 μM-*solutions of DNP amino-acids and peptides*

DNP derivative	Solvent	Optical density ($\log I_0/I$) 3500 A.	3900 A.
DNP-glycine	1% $NaHCO_3$	0·309	0·210
DNP-phenylalanine	1% $NaHCO_3$	0·313	0·214
DNP-glycyl-glycine	1% $NaHCO_3$	0·316	0·173
DNP-phenylalanyl-valine	1% $NaHCO_3$	0·310	0·178
N^5-DNP-lysine	N-HCl	0·296	0·204
O-DNP-tyrosine	N-HCl	0·058	0·0

different DNP derivatives. Beer's law is obeyed in solutions of concentrations <50 μM. Standard curves for a number of other DNP amino-acids in 1% $NaHCO_3$ have also been determined for a wavelength of 3500 A. and have been found to differ only slightly from those of the above derivatives.

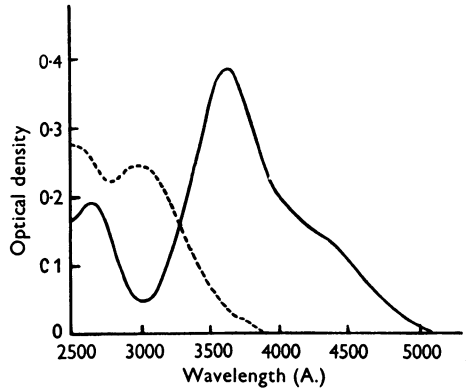

Fig. 2. Spectral absorption curves of 22 μM-solutions of N^5-DNP-lysine (——) and *O*-DNP-tyrosine (- - - - -) in N-HCl.

The determinations of the peptides containing N^5-DNP-lysine were carried out in N-HCl solutions, since N^5-DNP-lysine is unstable in alkaline solutions. The absorption curves for N^5-DNP-lysine and *O*-DNP-tyrosine are shown in Fig. 2. The curves for the lysyl peptides (*L* 2, *L* 3 and *L* 4) described below were identical with that obtained for N^5-DNP-lysine at wavelengths >3300 A. In order to avoid

interference from *O*-DNP-tyrosine, estimations of N^5-DNP-lysine and its peptides were carried out using a wavelength of 3900 A.

Identification of amino-acids present in DNP peptides

The amino-acids present in the DNP peptides were identified by paper chromatography (Consden, Gordon & Martin, 1944, 1947). After purification of the band on a suitable column, a sample containing 0·1–0·5 μmol. peptide was taken to dryness in a 50 ml. round-bottomed flask; 5 drops of 5·7 N-HCl were added and the mixture refluxed for 1 min. on a microflame to dissolve the peptide and wash down the sides of the flask. The contents were sucked into a small capillary tube using a piece of valve rubber as described by Consden, Gordon & Martin (1947). The capillary was sealed off and incubated at 105° for 24 hr. After cooling, the hydrolysate was transferred to a polythene strip and the HCl removed in a desiccator. A suitable portion of the residue was then taken for identification of the amino-acids by paper chromatography. Since only a few amino-acids were present in each peptide, it was best to use one-dimensional chromatograms on which known amino-acids were run parallel as controls. In the case of fraction *A* all the amino-acids present, except leucine and isoleucine, could be separated using the system phenol-0·3% NH_3-coal gas. Leucine and isoleucine were differentiated using 1:1 *n*-butanol-benzyl alcohol-HCN. All the amino-acids from fraction *B* could be identified using the above mixtures and *n*-butanol-acetic acid (Partridge, 1948). This latter solvent has been found to be very useful for amino-acid identification and to replace collidine for most purposes, though the R_F values vary considerably with the age of the mixture, and it is always advisable to run known amino-acids parallel to the unknowns. N^5-DNP-lysine and *O*-DNP-tyrosine can also be identified using this solvent, whereas they both run with the solvent front on phenol chromatograms.

If it is desired to identify the terminal residue as well as the other amino-acids present, the above technique can be modified in the following way. About 0·5 μmol. of the DNP peptide is hydrolysed as before, if necessary, as in the case of DNP-glycyl peptides, for a shorter period. The hydrolysate is transferred to a test tube, 2 ml. ether added and the tube shaken up. The terminal DNP amino-acid is extracted into the ether and can be identified on a suitable silica chromatogram. The aqueous solution after evaporation of the ether on a water bath is used for the identification of the other amino-acids present. In the case of peptides from fraction *B* it was often necessary to determine whether the yellow compound present in the hydrolysate was DNP-phenylalanine or N^5-DNP-lysine. This was readily accomplished by decanting the ether extract of the hydrolysate into another test tube and shaking it with a few drops of $NaHCO_3$ solution. N^5-DNP-lysine remains in the acid hydrolysate in the first test tube, whereas DNP-phenylalanine passes into the $NaHCO_3$ solution in the other test tube.

Estimation of amino-acids present in DNP peptides

In order to determine the amount of each amino-acid residue present in a peptide it was necessary to use an approximate micromethod, which could differentiate between one, two or three amino-acid residues per molecule. Quantitative paper chromatography, as used in the 'spot dilution' method of Polson, Mosley & Wyckoff (1947), was found to be very suitable and sufficiently accurate for the purpose.

After estimation of the DNP peptide in 1% $NaHCO_3$ a measured fraction (containing 0·5–1·0 μmol.) was acidified and extracted three times with ethyl acetate, each extract being washed with water to remove salts. The combined ethyl acetate solutions were taken to dryness *in vacuo* and the residue boiled under reflux with 5·7 N-HCl for a sufficient time to liberate all the terminal residue (2 hr. in the case of DNP-glycyl peptides and 4 hr. for DNP-phenylalanyl peptides). After cooling, the hydrolysate was extracted with ether, each extract being washed with water. The DNP amino-acid which was present in the ether extract was purified on a $CHCl_3$ column and estimated. The combined aqueous solution and washings were taken to dryness and hydrolysed for a further 20 hr. with 5·7 N-HCl. Excess HCl was removed by repeated evaporation *in vacuo* and the residue made to a suitable volume. Samples corresponding to various amounts between 0·05 and 0·005 μmol. peptide were then applied to a one-dimensional chromatogram and similar amounts of a standard solution of amino-acids were applied alongside the unknowns. After development of the chromatogram in a suitable solvent and treatment with ninhydrin the amount of each amino-acid could be determined by comparing the colour and size of the spots. In the case of the peptides of DNP-glycine and DNP-phenylalanine, phenol-0·3% NH_3-coal gas was a suitable solvent, and for the N^5-DNP-lysine peptides *n*-butanol-acetic acid was used.

With certain peptides it was desirable to confirm the estimations by a more accurate method and it was found that, where only a few known amino-acids were present, they could be estimated with reasonable accuracy by conversion to the DNP derivatives. The following technique was used. The aqueous hydrolysate from about 1·0 μmol. of DNP peptide, after extraction of the terminal DNP amino-acid, was taken to dryness and dissolved in 1 ml. water. $NaHCO_3$ (0·1 g.) was added and a solution of 0·1 ml. FDNB in 2 ml. ethanol. After shaking for 2 hr. the solution was taken to dryness, a few drops conc. HCl added and again taken to dryness. The residue was extracted with $CHCl_3$, CB_1 or CB_3, according to the nature of the DNP amino-acids present, and the extract put on the appropriate column. Excess FDNB moved rapidly down these columns and was discarded, while the desired DNP amino-acid bands were separated and estimated.

RESULTS

Estimation of the terminal residues of fraction A and B of oxidized insulin

As a preliminary to the following work, and also to determine the average molecular weight of the peptide chains of fractions *A* and *B* of oxidized insulin it was necessary to estimate the terminal residues of these fractions as accurately as possible.

The methods employed were those described in an earlier paper on insulin (Sanger, 1945); about 20 mg. of material were used in each experiment, which was carried out in duplicate. Since the main error is

probably due to the correction factor that has to be applied for the breakdown of the DNP amino-acids during hydrolysis, the times of hydrolysis were kept short, and small amounts of peptides present were estimated separately. In the earlier work the amount of protein present in a DNP protein was estimated from the content of amide N. This required more material than was available, so the molecular weight of the peptide was calculated from the molecular weight of the DNP derivative by assuming that one DNP group (mol.wt. 167) is attached to every free amino group and to every tyrosine and histidine residue (Porter, unpublished work). In the case of insulin this method gave the same results as the amide method.

Fraction A. For the estimation of the DNP-glycine in DNP-*A* a time of hydrolysis of 2 hr. was used and the DNP-glycine was purified on a $CHCl_3$ column. A small amount of a peptide band (probably *A* 4) was adsorbed tightly at the top of the column. This was eluted with CB_{17} and estimated separately. The results are shown in Table 3. The

Table 3. *Terminal residues of fractions A and B*

(Results corrected for breakdown during hydrolysis.)

	Fraction	μmol. DNP amino-acid/ 100 mg. DNP derivative	Mol. wt. of DNP derivative
A	DNP-glycine	27	—
	DNP-glycyl peptide	2	—
	Total	29	3450
B	DNP-phenylalanine	17	—
	DNP-phenylalanyl peptides	1·5	—
	Total	18·5	5100
	DNP-glycine	1	—
	N^5-DNP-lysine	19·5	4800

molecular weight obtained for DNP-*A* is 3450. Fraction *A* contains one free amino group and two tyrosine residues. It has no lysine or histidine (Sanger, 1949*a*), so that the DNP-*A* should contain three DNP groups, giving an additional molecular weight of 500. The molecular weight of the average chain in fraction *A* is thus 2950, which agrees well with the value of 2900 determined in the ultracentrifuge (Gutfreund & Ogston, 1949). This value corresponds to a molecular weight of 2750 for the individual chains in the intact insulin, the difference being due to the oxygen taken up during the treatment with performic acid.

Fraction B. DNP-*B* was hydrolysed for 4 hr. in boiling 5·7 N-HCl and the hydrolysate extracted with ether. The ether extract was used for the estimation of DNP-phenylalanine and the aqueous solution for N^5-DNP-lysine. The ethereal solution was taken to dryness and the residue put on a $CHCl_3$ column. Three bands were obtained. The main DNP-phenylalanine band was estimated directly. This was followed by a fainter band containing peptide *B* 2 and a small amount of DNP-glycine, which is usually present in preparations of fraction *B*. These two were separated on an acetone-*cyclo*hexane column. There was also a faint band remaining at the top of the $CHCl_3$ column, which consisted of higher DNP-phenylalanyl peptides and could be eluted with CB_{17}.

The aqueous solution, after extraction with ether, was hydrolysed for a further 20 hr. and the N^5-DNP-lysine purified on an M_{66} column and estimated.

The results for DNP-*B*, set out in Table 3, lead to an average molecular weight of 4950 on the assumption that the small amount of DNP-glycyl peptide has the same molecular weight as DNP-*A*. Fraction *B* contains two free amino groups, three tyrosine residues (estimated by difference between the tyrosine contents of insulin and fraction *A*) and two histidine residues (estimated directly on fraction *B* using histidine decarboxylase; Gale, 1945). DNP-*B* should therefore contain seven DNP groups and the molecular weight of fraction *B* would be 3800, and of the intact chains in insulin 3700. The molecular weight as determined in the ultracentrifuge was 7000 (Gutfreund & Ogston, 1949), so that it would appear that either two phenylalanyl peptide chains are joined together by some non-oxidizable cross-linkage or that in solutions of fraction *B* aggregates are formed containing two molecules. Since non-oxidizable cross-linkages are unknown, while aggregation reactions of insulin are well known (summarized by Sanger, 1949*b*), the latter explanation would seem the more probable

Isolation of DNP-phenylalanyl-valine from DNP-insulin

When DNP-insulin was hydrolysed for 2 hr. with boiling 5·7 N-HCl, a peptide band was present that moved at the same rate as DNP-glycine on a $CHCl_3$ column, but could be separated using an acetone-*cyclo*hexane column. In the experiment described by Sanger (1945), in which the DNP-amino-acids were isolated from 1·03 g. DNP-insulin, the DNP-glycine and the peptide were fractionated and the solution of the peptide in acetone-*cyclo*hexane allowed to stand overnight. It crystallized as needles as the acetone evaporated (yield 8·1 mg.).

A sample was hydrolysed and DNP-phenylalanine identified as the terminal residue. Paper chromatography showed that valine and no other amino-acid was present. By quantitative paper chromatography it was shown that there was one residue of valine per molecule of peptide and this was confirmed by estimation as the DNP derivative after purification on a $CHCl_3$ column. (Found: 3·3 % valine N as percentage of peptide; calc. for DNP-phenylalanyl-

valine, 3·2; for DNP-phenylalanyl-valyl-valine, 5·3 %.) The only possible structure for this peptide is therefore DNP-phenylalanyl-valine.

Identification of DNP peptides from fraction B

In the initial experiments, DNP-*B* was hydrolysed for 8 days with conc. HCl at 37°.

DNP-phenylalanyl peptides. The material extracted from the hydrolysate by ethyl acetate was first put through a CB_{17} column. Most of the colour travelled fast and only a few faint bands could be distinguished on the column. These were not investigated in detail. The fast-moving material was then fractionated on a CB_5 column. Three main bands were formed; *B*3, *B*4 and a band that moved fast; there were also a number of fainter bands moving more slowly than *B*4 and a fairly strong one that remained at the top of the column and could be eluted with CB_{17}. Bands *B*3 and *B*4 were shown to be homogeneous on Ea-pH 7 columns ($R = 0·25$ and 0·1 respectively). The material that moved fast in CB_5 consisted of DNP-phenylalanine, DNP-phenylalanyl-valine (*B*2) and a small amount of DNP-glycine. These were separated using a $CHCl_3$ column, and the second band refractionated on acetone-*cyclo*hexane. In some experiments a band was present that was adsorbed at the top of the $CHCl_3$ column and usually split into a number of faint bands on development with CB_1. It contained a rather complex mixture of amino-acids. These bands were believed to be due to some secondary products (possibly esters), formed from the DNP peptides during the fractionation. They did not occur in most experiments, and the same phenomenon was sometimes observed with DNP-*A*.

Table 4. *Peptides containing DNP-phenylalanine*

Peptide	Column used for separation	*R* value	Amino-acids present
*B*1	$CHCl_3$	0·4	None
*B*2	$CHCl_3$	0·2	Valine
	Acetone-*cyclo*hexane	0·7	
*B*3	CB_5	0·6	Valine, aspartic acid
*B*4	CB_5	0·3	Valine, aspartic acid, glutamic acid

The amino-acids identified in the main fractions are shown in Table 4. In each case the terminal residue was DNP-phenylalanine, and quantitative paper chromatography showed that the amino-acids were present in equimolecular proportions in all the peptides. The regularly increasing complexity of these peptides suggested that they were all derived from the same peptide chain containing the terminal sequence DNP-phenylalanyl-valyl-aspartyl-glutamic acid, and this was confirmed by the following partial hydrolysis experiments.

Peptide *B*3 (1·5 μmol.) was boiled under reflux with 5·7 N-HCl for 1 hr. and the degradation products studied chromatographically. The results are shown in Table 5. The only products of hydrolysis were DNP-phenylalanine and DNP-phenylalanyl-valine, which confirms the structure DNP-phenylalanyl-valyl-aspartic acid.

Table 5. *Partial hydrolysis of peptide* B 3

Peptide	Column used for separation	*R* value	Amino-acids present
*B*3·1	$CHCl_3$	0·4	None
*B*3·2	$CHCl_3$	0·2	Valine
	Acetone-*cyclo*hexane	0·7	
*B*3·3	CB_5	0·6	Valine, aspartic acid

The partial hydrolysis of peptide *B*4 was carried out for 14 days at 37° in conc. HCl. All the bands *B*1, *B*2, *B*3 and unchanged *B*4 were identified and shown to contain the same amino-acids as in Table 4. Since *B*4 gives rise to peptide *B*3, its structure can only be DNP-phenylalanyl-valyl-aspartyl-glutamic acid.

Peptides containing N^5*-DNP-lysine.* After extraction with ethyl acetate the aqueous solution from the hydrolysate of DNP-*B* was taken to dryness and fractionated on an M_{66} column. Three main bands were formed, and the faster one was refractionated on M_{66}-pH 6·5 to give two bands (*L*1, *L*2). There was also a fainter fast band that appeared to consist of higher peptides of N^5-DNP-lysine. Band *L*2 was shown to be homogeneous on Ea-HCl ($R = 0·2$) and CB_{30} ($R = 0·15$) columns, and bands *L*3 and *L*4 were homogeneous on CB_{30} ($R = 0·07$ and 0·04 respectively). Each peptide was purified by chromatography on talc, and the amino-acids identified and shown to be present in equimolecular amounts. The results are shown in Table 6. The terminal residues were determined by deamination with nitrosyl chloride (Consden, Gordon & Martin, 1947), and confirmed by the DNP technique. It was found that the latter method gave a rather more clear-cut result, whereas the former was much simpler to carry out. In a preliminary report of this work

Table 6. *Peptides containing* N^5*-DNP-lysine*

Peptide	*R* values: On M_{66}	*R* values: On M_{66}-pH 6·5	Amino-acids present	Terminal residue
*L*1	0·6	0·3	N^5-DNP-lysine	—
*L*2	0·6	0·1	N^5-DNP-lysine, alanine	N^5-DNP-lysine
*L*3	0·2	—	N^5-DNP-lysine, threonine, proline	Threonine
*L*4	0·13	—	N^5-DNP-lysine, threonine, proline, alanine	Threonine

(Sanger, 1948) the peptide $L4$ was described as threonyl-(N^5-DNP)-lysyl-alanine. The presence of proline had been overlooked in the paper chromatography, the yellow spot being confused with that due to N^5-DNP-lysine. It can, however, be clearly detected if the solvent *n*-butanol-acetic acid is used.

The results in Table 6 show that peptide $L2$ is N^5-DNP-lysyl-alanine. Peptides $L3$ and $L4$ were hydrolysed for 11 days with conc. HCl at 37°. The hydrolysate of $L3$ contained only the bands $L1$ and $L3$, and a faint band that moved at the same rate as $L2$ on M_{66}-pH 6·5 and contained N^5-DNP-lysine and proline. From peptide $L4$ all the bands $L1$, $L2$, $L3$ and $L4$ were obtained and were shown to contain the expected amino-acids. Since $L4$ has a terminal threonyl residue and gives rise to N^5-DNP-lysyl-alanine ($L2$) and a peptide ($L3$), containing threonine, proline and DNP-lysine, the only possible structure that can be assigned to it is threonyl-prolyl-(N^5-DNP-)-lysyl-alanine, and $L3$ must be threonyl-prolyl-(N^5-DNP-)-lysine.

Estimation of peptides containing DNP-phenylalanine and N^5-DNP-lysine in insulin

Having established the presence of the two peptide sequences, phenylalanyl-valyl-aspartyl-glutamic acid and threonyl-prolyl-lysyl-alanine, in fraction B of the oxidized insulin, the next problem was to estimate the yields of the various peptides from insulin, and to discover what proportion of the lysyl residues and the terminal phenylalanyl residues of insulin was present in these peptide sequences. When DNP-insulin was hydrolysed for 8 days with conc. HCl no peptides could be detected other than those which were obtained from DNP-B and very small amounts of DNP-glycyl peptides which did not interfere with the estimation of peptides containing DNP-phenylalanine and N^5-DNP-lysine. Almost all the terminal glycine was present as DNP-glycine itself, owing to the great lability of the glycyl peptide bond in strong acid.

About 100 mg. DNP-insulin were hydrolysed for 8 days at 37° with conc. HCl, the products fractionated as previously described for DNP-B and estimated colorimetrically. Peptides $B5$, $B6$ and $B7$ were obtained from bands that moved slower than $B4$ on CB_5. It was shown that each gave rise to band $B4$ on partial hydrolysis. $B8$ was composed of a number of faint bands that moved slowly on CB_{17}. These were eluted from the column with 80% (v/v) acetone, hydrolysed and the DNP-phenylalanine estimated after purification on a $CHCl_3$ column. The yields of the various peptides are shown in Table 7, expressed as moles of peptide as a percentage of the total DNP-phenylalanine in the original DNP-insulin, assuming there are exactly two phenylalanyl terminal residues per insulin molecule of molecular weight 12,000. It can be seen that 92% of the total DNP-phenylalanine is accounted for as compounds that fit into the peptide sequence DNP-phenylalanyl-valyl-aspartyl-glutamic acid, and it would seem evident that in fact all the terminal phenylalanyl residues are present in this sequence. An actual estimation of the amount of the peptide $B4$ in DNP-insulin can, however, be obtained by studying the hydrolysis of the peptide itself under

Table 7. *Yields of DNP-phenylalanyl peptides from an 8-day hydrolysate of DNP-insulin*

(Results expressed as mol. of DNP peptide as percentage of the total terminal phenylalanyl residues.)

	Peptide		Yield (%)
$B1$	DNP-phenylalanine		13
$B2$	DNP-phenylalanyl-valine	16	
$B3$	DNP-phenylalanyl-valyl-aspartic acid	13	
$B4$	DNP-phenylalanyl-valyl-aspartyl-glutamic acid	30	
	Total known peptides ($B2$, $B3$, $B4$)		59
$B5$		6	
$B6$		4	
$B7$		10	
	Total higher peptides from the same sequence		20
$B8$	Unknown mixture of peptides		6
	Total		98

Table 8. *Yields of DNP peptides from an 8-day hydrolysate of peptide* B4

Peptide	Yield from $B4$ (%)	Yield from DNP-insulin (%)
$B1$	13	13
$B2$	14	16
$B3$	12	13
$B4$	55	50–56

the conditions that were used for hydrolysing the DNP-insulin. Thus, if all the DNP-phenylalanine in DNP-insulin is combined in the single tetrapeptide sequence, then the partial hydrolysis of peptide $B4$ should give the same products in the same proportions as are obtained from DNP-insulin. This assumes that the stability of the various peptide bonds is the same whether the DNP-peptide is free or combined in the protein. While this may not be completely true it is probably a safe approximation. Table 8 shows the yields of DNP-peptides from peptide $B4$ (2 μmol.) that had been hydrolysed for 8 days in conc. HCl at 37°, compared with the yields from DNP-insulin. The figure for the amount of $B4$ obtained from DNP-insulin includes higher peptides containing this sequence. It can be seen that within the limits of the methods the two sets of figures agree. The degree of accuracy of these results probably indicates that not more than about 10% of the terminal phenylalanine of insulin is present in sequences other than that detected here. Thus, not

only do these results prove that both the phenylalanine terminal residues in the insulin molecule of molecular weight 12,000 are combined in the same tetrapeptide sequence, but they do justify the assumption of this value for the molecular weight.

Table 9. *Yields of peptides containing* N^5*-DNP-lysine from an 8-day hydrolysate of DNP-insulin*

(Results expressed as mol. of peptide as percentage of the total lysine.)

	Peptide	Yield from DNP-insulin (%)	Yield from *L*4 (%)
*L*1	N^5-DNP-lysine	14	14
*L*2	N^5-DNP-lysyl-alanine	19	23
*L*3	Threonyl-prolyl-(N^5-DNP)-lysine	32	32
*L*4	Threonyl-prolyl-(N^5-DNP)-lysyl-alanine	23	21
	Total known peptides (*L*2, *L*3, *L*4)	74	
*L*5	Unknown	6	
	Total	94	90

Table 9 shows the yields of the various peptides containing N^5-DNP-lysine obtained from DNP-insulin and from a similar hydrolysate of peptide *L*4. *L*5 is the uninvestigated band that moved fast on an M_{66} column. The N^5-DNP-lysine in it was estimated after complete hydrolysis. The yields of the peptides from DNP-insulin make it very improbable that there is a second type of peptide containing lysine and agreement between the two sets of figures confirms this.

Identification of DNP-glycyl peptides from fraction A

Hydrolysis in concentrated hydrochloric acid. Owing to the great lability to strong acid of the peptide bond involving the carboxyl group of the terminal glycyl residues, partial hydrolysates of DNP-*A* carried out in conc. HCl contained large amounts of DNP-glycine and of long peptides that could not be investigated easily, and only poor yields of small peptides that could readily be studied. Nevertheless, it was possible to determine the structure of three of these small peptides, and as most of the work was carried out on such hydrolysates, the results are reported here, although much more satisfactory yields were obtained at a later stage when the hydrolysis was carried out in dilute acid.

DNP-*A* was hydrolysed for 24 hr. with conc. HCl at 37°, and the material that was extracted with ethyl acetate first put on a CB_{17} column. Besides the main fast-moving band there were a number of fainter bands. The strongest of these, which travelled at a gradually decreasing rate down the column, gave rise to the amino-acids leucine, serine, glutamic acid and *O*-DNP-tyrosine on hydrolysis. A peptide containing these amino-acids had been reported by Woolley (1948), and it was at first assumed that the two were the same (Sanger, 1948). On quantitative estimation, however, it was found that there were about five residues of each of the above amino-acids per residue of DNP-glycine, so that the fraction was a mixture of a small amount of a DNP-glycyl peptide (probably *A* 5), which was responsible for the colour of the band on the column, and a large amount of a non-terminal peptide containing the above four amino-acids. The exact structure of this peptide has not yet been determined, but its presence does indicate that DNP-peptides, separated by the above methods, may be contaminated by non-terminal peptides.

The coloured material that moved fast on the CB_{17} column was next put on a CB_5 column. This separated a band (*A* 4) which was run out by development with CB_8. The fast-moving material from this column was fractionated on a $CHCl_3$ column. A very strong DNP-glycine band (*A* 1) was present, and this was followed by two peptide bands (*A* 2 and *A* 3) which were developed with $CB_{0\cdot5}$ and CB_1. Band *A* 2 was relatively faint and sometimes difficult to detect. Bands *A* 3 and *A* 4 were homogeneous on Ea-pH 7 columns ($R = 0\cdot4$ and $0\cdot1$ respectively) and band *A* 4 was homogeneous on M_{66}-pH 7 ($R = 0\cdot4$) and M_{66}-pH 8 ($R = 0\cdot07$).

Table 10. *Peptides of DNP-glycine*

Peptide	Column used for separation	*R* value	Amino-acids present
A 1*	$CHCl_3$	0·2	—
A 2	$CHCl_3$	0·06	Isoleucine
	$CB_{0\cdot5}$	0·2	
	CB_1	0·6	
A 3	$CHCl_3$	0·03	Isoleucine, valine
	$CB_{0\cdot5}$	0·1	
	CB_1	0·35	
A 4	CB_5	0·25	Isoleucine, valine, glutamic acid
	CB_8	0·8	

* DNP-glycine.

Table 10 shows the amino-acid composition of the three bands. In each case the terminal residue was DNP-glycine. Estimation showed that the amino-acids in *A* 3 and *A* 4 were present in equimolecular proportions to the DNP-glycine. The composition of the peptides suggests that they are all derived from the same peptide chain containing the terminal sequence DNP-glycyl-isoleucyl-valyl-glutamic acid and that this was the structure of *A* 4 was shown by partial hydrolysis. A sample was refluxed for 12 hr. with 0·1 N-HCl. The hydrolysate contained the peptide bands *A* 2 and *A* 3 and isoleucine and valine were identified in the *A* 3.

Hydrolysis in hydrochloric acid (0·1 N). In an attempt to find a more specific method for hydro-

lysing proteins, the action of boiling 0·1 N-HCl on DNP-*A* was studied. It was found that the specificity of this hydrolytic agent was very different from that of concentrated HCl, and that the smaller DNP-peptides were produced in much better yields. It seems probable that the hydrolysis of these oxidation products in dilute acid is analogous to the hydrolyses catalysed by long-chain sulphonic acids (Steinhardt & Fugitt, 1942), the peptide chains containing $-SO_3H$ groups acting as long-chain anions and catalysing a particular type of hydrolysis which is characterized by the increased lability for the amide groups and no doubt a corresponding difference in specificity towards other peptide bonds. That the $-SO_3H$ groups were catalysing the hydrolysis is indicated by the difference in the rate of hydrolysis of insulin and oxidized insulin illustrated in Fig. 3. No such difference was found when the hydrolyses were carried out in conc. HCl in which the $-SO_3H$ groups are uncharged. In these experiments the liberation of free amino groups was followed using the Van Slyke apparatus.

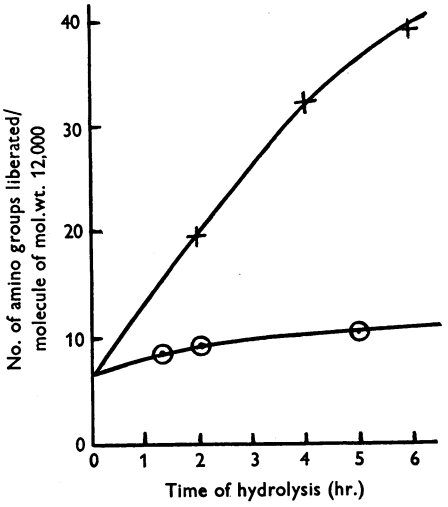

Fig. 3. Rate of hydrolysis of insulin (⊙) and oxidized insulin (+) in boiling 0·1 N-HCl.

It should be noted that the formation of diketopiperazines in dilute acid has been demonstrated (Abderhalden & Komm, 1924), so their presence was a possibility in these experiments, though the temperature employed was probably too low. The DNP derivatives could not undergo any rearrangement of this type as they contain no free amino group, and there was no evidence of any contamination of the DNP peptide bands with diketopiperazines.

In a hydrolysate of DNP-*A* that had been boiled under reflux for 12 hr. with 0·1 N-HCl there was present, besides the three previously described peptides, considerable amounts of another peptide (*A* 5) which moved with $R = 0·9$ on CB_{17} and was best purified on CB_{15} (R = approx. 0·4), though it was extremely insoluble, and only faintly coloured bands could be obtained. On an Ea-pH 6 column it gave an apparently homogeneous band with $R = 0·7$, though it 'tailed' rather badly on all buffered columns. This peptide contained DNP-glycine and the amino-acids isoleucine, valine and glutamic acid only. Quantitative paper chromatography showed that there were two residues of glutamic acid to every one residue of DNP-glycine, isoleucine and valine, and this was confirmed by estimation of the glutamic acid as the DNP derivative. No amide group was present, since no NH_3 was formed on hydrolysing for 4 hr. with boiling 2 N-HCl. Table 11 shows the results of an

Table 11. *Partial hydrolysis products of peptide* A 5

Peptide	Column used for separation	R value	Amino-acids present	Yield of peptide (%)
A 5·1	$CHCl_3$	0·1	—	29
A 5·2	CB_1	0·6	—	4
A 5·3	CB_1	0·35	Isoleucine, valine	10
A 5·4	CB_5	0·25	Isoleucine, valine, glutamic acid	19
A 5·5	CB_{15}	0·4	Isoleucine, valine, glutamic acid	27

experiment in which *A* 5 was hydrolysed for 12 hr. with boiling 0·1 N-HCl and the hydrolysis products investigated. It can be seen that it breaks down to give peptides *A* 2, *A* 3 and *A* 4, therefore the only possible structure is DNP-glycyl-isoleucyl-valyl-glutamyl-glutamic acid.

Estimation of the DNP-glycyl peptides from DNP-insulin

Since DNP-insulin is insoluble in 0·1 N-HCl it was not possible to obtain a significant yield of DNP peptides by hydrolysing it directly. However, if the insulin was first oxidized and then treated with FDNB, there was very little loss of material, and the DNP-*A* present was largely soluble and gave good yields of the peptides.

In order to obtain the maximum possible yields of peptides at the expense of the yield of DNP-glycine, the hydrolysis mixture was extracted from time to time during the course of hydrolysis, so that the larger extractable peptides (*A* 4 and *A* 5) were removed and not broken down to DNP-glycine. Also unidentified fractions were hydrolysed further to give the known peptides.

As a preliminary, the breakdown of DNP-*B* under these conditions was studied. Owing to its insolubility it was not extensively hydrolysed. The main breakdown products were peptide *B* 3 and a little *B* 2. There was also a small amount of an unidentified DNP-phenylalanyl peptide that contaminated *A* 5.

Insulin (0·2 g.) was oxidized with performic acid for 30 min. (Sanger, 1949*a*). The product was precipitated with acetone and treated with FDNB for 2 hr. in the usual manner. To remove excess FDNB about 50 ml. ether and 5 ml. water were added to the reaction mixture and, after shaking and allowing to settle, the ether layer was decanted and the aqueous layer and insoluble material were extracted twice more with ether in the same way. It was then acidified with HCl and taken to dryness *in vacuo*. Excess HCl was removed by repeated evaporation *in vacuo* with water; 50 ml. 0·1 N-HCl were added and the mixture boiled under reflux for 3 hr. It was then cooled, and extracted twice with an equal volume of ethyl acetate. After removal of dissolved ethyl acetate from the aqueous layer by evaporation *in vacuo*, it was boiled for another 3 hr. and again extracted. This process was repeated six times in all, making a total time of hydrolysis of 18 hr. The ethyl acetate solutions were taken to dryness after each extraction, and the residues stored with a little water. They were all combined and the DNP peptides purified by extraction into bicarbonate as previously described. About 3 % of the DNP-glycine was lost during this treatment. The final extract containing the terminal DNP peptides was taken to dryness, the residue dissolved in 80 % ethanol and samples taken for the analyses. Several replicate analyses were performed, but not all peptides were estimated in each case.

Peptide *A* 5 was separated on a CB_{15} column. It was usually contaminated with a small amount of a DNP-phenylalanyl peptide, which was estimated from the amount of DNP-phenylalanine produced on hydrolysis of the fraction. Peptide *A* 4 was fractionated on a CB_5 column and was free from any DNP-phenylalanyl peptides. Band *B* 3 and a faint band *B* 4 moved in front of the *A* 4 band on this column, and there were traces of other unidentified DNP-phenylalanyl peptides moving more slowly. Pure bands, *A* 2 and *A* 3, were separated from a $CHCl_3$ column which was developed with CB_1. The DNP-glycine fraction obtained from the $CHCl_3$ column contained peptide *B* 2 which was separated on an acetone-*cyclo*hexane column. The yields of the various derivatives are shown in Table 12. The figures are given as percentages of the terminal glycyl residues of the original insulin taken. The corresponding yields of the DNP-phenylalanyl peptides *B* 1, *B* 2, *B* 3 and *B* 4 were 3, 9, 20 and 3 % respectively.

On the CB_{15} column much colour was present in a number of slow-moving bands which were eluted together from the column with 80 % acetone and were designated fraction *Y*. This fraction contained about 20 % of the DNP-glycine as estimated after hydrolysis. A sample was hydrolysed for 12 hr. with boiling 0·1 N-HCl, and the products estimated as for the main hydrolysate. It gave rise to the expected peptides showing that the DNP-glycine present in it was largely in the form of higher peptides of the same sequence. The yields are shown in column 4 of Table 12.

Another 8 % of the DNP-glycine was in the insoluble residue (fraction *R*) from the original dilute acid hydrolysate. This was probably present as the DNP derivative of some incompletely oxidized insulin. In order to hydrolyse it further it was first brought into solution by incubating for 1 hr. at 37° with conc. HCl. After removal of the HCl *in vacuo*, the residue was hydrolysed for a further two periods of 12 hr. with 0·1 N-HCl, and the peptides estimated as before. The yields are shown in column 5 of Table 12.

The total yields of peptides leave no doubt that both the terminal glycyl residues are combined in the same terminal peptide sequence. It was not possible to determine the amount of peptides *A* 4 or *A* 5 directly by studying their breakdown under identical conditions of hydrolysis, as was done for peptides *B* 4 and *L* 4, since the type of hydrolysis used here could not be reproduced on a peptide. The unchanged peptide would be removed on each extraction with ethyl acetate, whereas long peptides from DNP-*A* remained in the aqueous solution and were hydrolysed further. It is also probable that the long-chain charged polypeptides have a catalytic effect on the hydrolysis, which it would be difficult to reproduce exactly on a peptide. The sum of the yields of *A* 4 and *A* 5 (55 %) does, however, give direct proof that over

Table 12. *Yields of DNP-glycyl peptides from DNP-insulin*

(Results expressed as mol. of peptide as percentage of the terminal glycyl residues in the original insulin.)

Peptide	Yield from main hydrolysate	Mean	Yield from fraction *Y* (%)	Yield from fraction *R* (%)	Total yield (%)
A 1	17·5, 16	17	3·3	1·3	22
A 2	1·4	1	2·6	0·5	9
A 3	4, 5	5			
A 4	11, 11	11	2·9	1·1	15
A 5	37, 30, 38, 35	35	3·5	1·9	40
Total known peptides					64
Total					86

half of the terminal glycyl residues are in the form of the sequence glycyl-isoleucyl-valyl-glutamic acid. Table 11 shows that neither of these peptides is completely stable to this type of hydrolysis, so that the actual amount present must be considerably greater.

DISCUSSION

It has been shown that both the terminal phenylalanyl residues are combined in the insulin molecule in the form of the single tetrapeptide sequence phenylalanyl-valyl-aspartyl-glutamic acid, and that both the lysyl residues which are in the same chains are in the form of the sequence threonyl-prolyl-lysyl-alanine. Also both the terminal glycyl residues of the other two chains are present in the pentapeptide sequence glycyl-isoleucyl-valyl-glutamyl-glutamic acid. The simplest explanation of these results is that the two phenylalanyl chains are identical, and that the two glycyl chains are identical. In the phenylalanyl chains, which contain about thirty amino-acid residues, eight of the positions relative to the terminal residues and the lysyl residues are occupied by the same amino-acids in both chains. It is, of course, possible that they might differ in the nature of one or a few residues in a position in the chain not investigated in this work. As nothing is known about the principles which govern the arrangement of amino-acids along a peptide chain, such a possibility cannot be excluded. However, if one assumes that each position in the chain can be occupied by any of the sixteen different amino-acids present in these chains, the chance that two different chains would contain the same terminal tetrapeptide is about 1 in 50,000. Certain limits are obviously imposed by the specificities of the mechanisms responsible for protein synthesis. Investigations of the free amino groups of a number of proteins by the DNP technique (summarized by Sanger, 1949*c*) have shown that the terminal position in the protein chains may be occupied by a variety of different amino-acids. There appears to be no principle that defines the nature of the residue occupying this position in different proteins and it would seem probable that this would apply to other positions in the molecule. It would thus seem a reasonable conclusion from the above results that the insulin molecule is built up of two pairs of identical chains. However, if this were indeed the case, then the simplest asymmetric unit of insulin should have a molecular weight of 6000 and all the analytical figures for the amino-acids should fit such a unit. The most recent data on the amino-acid composition of insulin have been collected by Tristram (1949). While for most amino-acids the number of residues per unit of molecular weight 6000 approximates to a whole number within the limits of the methods of assay, the values for proline and isoleucine would indicate about 1·5 and 1·3 residues respectively. It is very unlikely that these estimations are more than 10% in error, so that they do cast considerable doubt on the above conclusion, and all that can be definitely stated is that the insulin molecule is built up of two pairs of very similar chains.

In contrast to the results of this work are the results obtained by Woolley (1949). Using a method similar in principle to that used here, he has reported the isolation of a number of peptides from a tryptic hydrolysate of DNP-insulin, which he believed to be DNP-glycyl peptides. He was unable, however, to demonstrate the presence of significant amounts of DNP-glycine on hydrolysis of the peptides concerned, although it could be detected in a hydrolysate of DNP-insulin. This and the results reported in the present paper would seem to indicate that the peptides isolated by Woolley were not DNP-glycyl peptides. He used ethyl acetate and butanol extracts which would be expected to contain non-terminal peptides, though it is difficult to say what the 'chromophore' group in these peptides could have been.

The methods of investigation described in this paper should be generally applicable to all proteins and peptides that contain a free amino group. Besides giving information about the structure of the proteins, the method can probably be applied as a sensitive test for the identity and homogeneity of proteins. A simple experiment in which DNP-peptide bands are separated and the amino-acids present in them identified can readily be carried out, and it is extremely unlikely that two different proteins would give the same result, probably less likely than that they would appear homogeneous by any of the known physicochemical methods of determining protein purity.

SUMMARY

1. The dinitrophenyl method for the identification of the terminal residues of proteins has been extended by the use of partial hydrolysis to the identification and estimation of terminal peptides.

2. The method was applied to insulin; and it was shown that both the terminal phenylalanyl residues are present in the form of the tetrapeptide sequence phenylalanyl-valyl-aspartyl-glutamic acid, both the lysyl residues which are in the same chains are present in the sequence threonyl-prolyl-lysyl-alanine and both the terminal glycyl residues are present in the pentapeptide sequence glycyl-isoleucyl-valyl-glutamyl-glutamic acid.

3. It is concluded that the insulin molecule is built up of two pairs of very similar, if not identical, polypeptide chains.

I wish to express my thanks to Prof. A. C. Chibnall, F.R.S., for his advice and encouragement.

REFERENCES

Abderhalden, E. & Komm, E. (1924). *Hoppe-Seyl. Z.* **134**, 121.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1944). *Biochem. J.* **38**, 224.
Consden, R., Gordon, A. H. & Martin, A. J. P. (1947). *Biochem. J.* **41**, 590.
Consden, R., Gordon, A. H., Martin, A. J. P. & Synge, R. L. M. (1947). *Biochem. J.* **41**, 596.
Craig, L. C. (1944). *J. biol. Chem.* **155**, 519.
Gale, E. F. (1945). *Biochem. J.* **39**, 46.
Gordon, A. H., Martin, A. J. P. & Synge, R. L. M. (1943). *Biochem. J.* **37**, 79.
Gutfreund, H. & Ogston, A. G. (1949). *Biochem. J.* **44**, 163.
Martin, A. J. P. & Synge, R. L. M. (1941). *Biochem. J*, **35**, 1358.
Middlebrook, W. R. (1949) *Nature, Lond.*, **164**, 501.
Moyle, V., Baldwin, E. & Scarisbrick, R. (1948). *Biochem. J.* **43**, 308.
Partridge, S. M. (1948). *Biochem. J.* **42**, 238.
Phillips, D. M. P. & Stephen, J. M. L. (1948). *Nature, Lond.*, **162**, 152.
Polson, A., Mosley, V. M. & Wyckoff, R. W. G. (1947). *Science*, **105**, 603.
Porter, R. R. & Sanger, F. (1948). *Biochem. J.* **42**, 287.
Sanger, F. (1945). *Biochem. J.* **39**, 507.
Sanger, F. (1948). *Nature, Lond.*, **162**, 491.
Sanger, F. (1949*a*). *Biochem. J.* **44**, 126.
Sanger, F. (1949*b*). *Ann. Rep. chem. Soc.* **45**, 283.
Sanger, F. (1949*c*). *Biochem. Soc. Symp.* **3**, 21.
Steinhardt, J. & Fugitt, C. H. (1942). *Bur. Stand. J. Res., Wash.*, **29**, 315.
Tristram, G. R. (1946). *Biochem. J.* **40**, 721.
Tristram, G. R. (1949). *Advanc. protein. Chem.* **5** (in the Press).
Woolley, D. W. (1948). *Fed. Proc.* **7**, 200.
Woolley, D. W. (1949). *J. biol. Chem.* **179**, 593.

Regulation of Urinary Steroid Excretion

2. SPONTANEOUS CHANGES IN THE PATTERN OF DAILY EXCRETION IN MENTAL PATIENTS

By M. REISS, R. E. HEMPHILL, J. J. GORDON and E. R. COOK

Biochemical and Endocrinological Research Department, Bristol Mental Hospital, Fishponds, Bristol

(*Received* 17 *June* 1949)

The daily excretion of 17-ketosteroids of a large series of mental patients has been estimated for several months—in some cases for more than a year. It was observed that in some cases the rate of daily excretion was fairly constant, but in others there were wide variations, far outside the possible error of determination. Some light has been thrown on the physiological mechanism responsible by a detailed long-term study of steroid excretion in one patient with cyclical manic and depressive phases, and by the study of diurnal variations in 17-ketosteroid output in chronic schizophrenic patients before and after treatment.

EXPERIMENTAL

Estimation of steroid excretion. In all cases where the total neutral 17-ketosteroid fraction only was required, the simultaneous hydrolysis and extraction procedure of Callow, Callow, Emmens & Stroud (1939) was employed. These estimations were carried out at intervals over long periods, on 24–72 hr. urine specimens of mental patients and normal subjects. In the case of the cyclic patient, however, the more detailed procedure for fractionation and estimation of the steroids, as previously described by us (Reiss, Hemphill, Gordon & Cook, 1949), was followed.

Diurnal variations were estimated by making use of the micromethod of Drekter, Pearson, Bartczak & McGavack (1947) which requires only 10 ml. urine; such estimations were carried out in duplicate, the margin of error being rather greater than in the macromethod. It was thus possible to estimate 17-ketosteroid output on urine collections taken every 3 hr. throughout the day (7 a.m.–10 p.m.), ending with a 9 hr. night collection (10 p.m.–7 a.m.).

RESULTS

Variations in 24 hr. excretion. Fig. 1 illustrates the excretion rates of total 17-ketosteroids in mg./24 hr. for normal subjects, over periods of 6–14 months, and this diagram may be compared with the corresponding excretion rates for chronic schizophrenic patients, which are recorded in Fig. 2. The excretion rates for normal subjects are relatively constant, varying at the most $\pm$ 1·5 mg./24 hr. Some of the patients also show excretion rates which approach this constant pattern (e.g. nos. 7, 9, 10 and 12), but others show wide variations, which may amount to $\pm$ 6·0 mg./24 hr. These variations could not necessarily be correlated with the clinical picture. Further investigations designed to ascertain whether relationships exist between psychotic condition and constancy or inconstancy of ketosteroid excretion are at present being undertaken.