Biochem. J. (1969) **115**, 587
Printed in Great Britain

Acetoacetylation of Insulin

By D. G. LINDSAY and S. SHALL
School of Biological Sciences, University of Sussex, Falmer, Sussex

(*Received* 28 *July* 1969)

Insulin was treated with diketen at pH 6·9. The reaction mixture was resolved into four components by DEAE-Sephadex chromatography. The first component was unchanged insulin. The second and third components were shown by end-group analysis to be substituted on phenylalanine B-1 and glycine A-1 respectively. The fourth component was disubstituted on both phenylalanine B-1 and glycine A-1. The ϵ-amino group of lysine B-29 was not involved in the reaction at low reagent concentrations. The purity of these derivatives was checked by their electrophoretic behaviour and by measurement of the rate of their reaction with trinitrobenzenesulphonic acid. The hormonal activity of the derivatives was determined. The effect of the modifications on the hormonal activity and the tertiary structure of insulin is discussed.

Although the first X-ray-diffraction pattern of single crystalline insulin was obtained in 1935 (Crowfoot, 1935) it has been difficult to prepare a series of isomorphous heavy-atom derivatives of the hormone, which are required to determine its tertiary structure. Diffusion of heavy-metal salts into the crystal has produced a number of derivatives but many of these have proved unsatisfactory because of the binding of the metal at many sites. The ideal solution would be covalent modification of the hormone at a specific site leading to the introduction of a heavy-metal atom at that site alone.

The obvious site of modification of insulin is on one of the three free amino groups, glycine (A-1), phenylalanine (B-1) and ϵ-lysine (B-29). However, until now no information existed about the effect such modification would have on the ability of the modified insulins to crystallize isomorphously with the parent hormone. Scoffone, Marzotto, Pajetta & Galzigna (1968) have shown that diketen is a versatile reagent for the acetoacetylation of the free amino groups of proteins. Acetoacetylation also has a potential application to X-ray-crystallographic analysis because the resulting acetoacetamides might show a preferential affinity for heavy-atom salts. Sen & Thankarajan (1968) have prepared uranyl complexes of acetoacetanilides and a number of transition-metal complexes of acetoacetamides have been described (Sen & Umapathy, 1968).

The possibility of directly forming a carbon–mercury bond at the activated methylene group in the resulting acetoacetamide also exists. Naik (1932) has prepared, in simple acetoacetamides, a number of such compounds using reasonably mild conditions.

From the observation by Fraenkel-Conrat & Fraenkel-Conrat (1950) that insulin could be acetylated without any appreciable loss of activity it was inferred that no gross alteration of the tertiary structure occurred with simple acetylating agents.

Scoffone *et al.* (1968) also demonstrated that the native proteins could be regenerated by removal of acetoacetyl groups with aqueous hydroxylamine. The reaction has therefore the additional advantage of utilizing the acetoacetyl group as a protecting group for insulin, enabling modification studies on other reactive sites in the molecule to be examined.

Previous work has shown that the three amino groups of insulin react at different rates with a variety of reagents. The order of reactivity would appear to depend on the structure of the modifying reagent. Phenyl isothiocyanate reacts preferentially with phenylalanine B-1 (Africa & Carpenter, 1968) and a similar reactivity is observed with fluorescein isothiocyanate (Tietze, Mortimore & Lomax, 1962; Bromer, Sheehan, Berns & Arquilla, 1967). However, treatment of insulin with phenyl isocyanate (Anderson, 1956*a*) and 2,4,6-trinitrotoluene (Mills, 1953) appears to cause preferentially reaction with glycine A-1.

These reactivity differences have enabled some derivatives to be isolated with varying degrees of modification. It has not been possible until now, however, to produce a series of derivatives modified

on each of the terminal amino groups and to study the effect of this modification on the biological activity and the crystallization properties of the hormone.

MATERIALS

Crystalline ox insulin (batch no. 26918), six times recrystallized, was used without further purification and was the product of British Drug Houses Ltd. (Poole, Dorset).

Urea (A.R. grade) was de-ionized by passing a 7M solution through a mixed-bed Amberlite MB 3 column and was used immediately after preparation.

Diketen was a product of Ralph Emmanuel Ltd. (London S.E. 1). The crude product was redistilled before use: b.p. 96° at 400mm.Hg.

Acetoacetyl glycine ethyl ester was prepared according to the method of Lacey (1954); m.p. 49–51° (lit. m.p. 48–50°).

2,4,6-Trinitrobenzenesulphonic acid and 'dansyl chloride' were obtained from British Drug Houses Ltd.

Cellulose acetate strips were the product of Oxoid Ltd. (London E.C. 4).

Ponceau S was acquired from Gelman Associates Ltd. (Hawksley's Ltd., Lancing, Sussex).

METHODS

Cellulose acetate electrophoresis. Cellulose acetate electrophoresis was performed exactly as described by Carpenter & Hayes (1963), with a Shandon electrophoresis tank fitted with a water-cooled plate (20cm. × 20cm.) and a power supply of 1kv. The buffer used was 0·05M-tris–HCl, 7M-urea, pH7·0, and the cellulose acetate strips (20cm. × 5cm.) were stained in a 0·2% solution of Ponceau S in 3% acetic acid.

Conductance. This was measured with an LKB Conductolyzer model 5300B fitted with a dip-cell.

pH. This was determined in a Radiometer (model 26) direct-reading pH-meter.

Ultraviolet absorption. Spectral measurements were made in a Cary model 15 recording spectrophotometer.

Zinc-free insulin. This was prepared by passing a solution of zinc insulin in 7M-urea through a column of Amberlite MB-3 mixed-bed resin. The eluent was extensively dialysed and freeze-dried.

Reaction of insulin with diketen. Zinc insulin (1·0g.) was dissolved to a final concentration of 0·167mM in 0·1M-HCl (45ml.) and the pH adjusted with M-NaOH to 6·9 in the titration vessel of a Radiometer (type TTT1) pH-stat. Freshly distilled diketen was added in 5μl. portions, the pH being maintained at 6·9 by the addition of alkali. When 1 equiv. of alkali was consumed, after approx. 5min., a sample was removed for ninhydrin analysis (Moore & Stein, 1954). Further portions of diketen were added until the ninhydrin colour yield had decreased by 30%. The amount of diketen added was 40μl. (final concn. 0·49mM).

The reaction mixture was made 0·2M in Na_2CO_3–$NaHCO_3$ buffer, pH9·5, and left overnight. After extensive dialysis against distilled water and freeze-drying, the yield of modified protein was 950mg.

In a second experiment, zinc insulin (500mg.; final concn. 0·085mM) was treated with diketen (100μl.; final concn. 1·2mM) at pH6·9 exactly as described above.

Chromatographic separation of acetoacetyl insulins. The acetoacetyl insulins were separated on a column of DEAE-Sephadex A-25 (2·5cm. × 40cm.) by a modification of the method described by Bromer & Chance (1967).

The column was equilibrated with buffer containing 0·01M-tris and 0·05M-NaCl in 7M de-ionized urea adjusted to pH7·20 with M-HCl. Acetoacetyl insulin (500mg.) was dissolved in this buffer (50mg./ml.) and the column was developed at a flow rate of 54·0ml./hr., 12ml. fractions being collected. After 120ml. of eluent was collected a linear gradient, obtained by running 0·01M-tris–0·15M-NaCl in 7M de-ionized urea (1l.) at pH7·20 into the stirred reservoir of the starting buffer (1l.), was applied.

The monoacetoacetylated insulins were rechromatographed on a column of DEAE-Sephadex A-25 (2·5cm. × 40cm.) was described above with the exception that the starting buffer was adjusted to pH7·30 with M-HCl. A linear gradient was applied by running 0·01M-tris–0·10M-NaCl in 7M de-ionized urea (1l.) at pH7·30 into the stirred reservoir of the starting buffer (1l.).

The protein concentration was determined from the extinction of the solution at 277nm.

End-group analysis. (*a*) Dansylation. Samples containing 5–10nmoles of the isolated derivatives were dansylated according to the method of Gray (1967). The residue after acid hydrolysis was dissolved in 15μl. of acetone–acetic acid (3:2) and the extracted dansyl amino acids were separated by flat-plate electrophoresis at pH4·38.

(*b*) Carbamoylation. The modified insulins were carbamoylated as described by Stark (1967). The carbamoylated derivatives were cyclized to hydantoins and, after separation, converted into the corresponding amino acids by hydrolysis *in vacuo* in 0·2M-NaOH for 24hr. Amino acid analyses were performed on a Beckman–Spinco model 120C automatic amino acid analyser. Insulin was put through these steps to determine the correction factors needed to make the method quantitative.

Trinitrobenzenesulphonic acid analysis. The method used was essentially that described by Freedman & Radda (1968). Insulin was dissolved (to 31·0μM) in 0·1M-phosphate buffer, pH7·4, and TNBS* in phosphate buffer was added to a final concentration of 4·6mM. A blank of the same concentration of TNBS in phosphate buffer was placed in the reference cell. A time-course analysis for 1100hr. was performed for each derivative by reading the extinction at 345nm. in 5mm. silica cells.

Determination of acetoacetyl groups. The number of carbonyl groups present in each of the modified derivatives was determined by a modification of the method of Jordon & Veatch (1964). The protein (3mg.) was dissolved in water (0·5ml.) and carbonyl-free ethanol (1·0ml.). A saturated solution of 2,4-dinitrophenylhydrazine in carbonyl-free ethanol (1.ml.) and conc. HCl (50μl.) was added. After being heated at 55° for 30 min. the solution was made alkaline with ethanolic KOH, prepared by dissolving KOH (6g.) in water (18ml.) and diluting to 100ml. with ethanol.

Extinction at 480nm. was read 10min. after the addition of KOH. A blank of insulin (derivative A) was put through the same procedure.

A calibration curve was constructed by using acetoacetylglycine ethyl ester.

* Abbreviation: TNBS, trinitrobenzenesulphonic acid.

Isoelectric precipitation of acetoacetyl derivatives. The pH at which isoelectric precipitation occurred for each of the derivatives was determined by measuring the protein concentration in 0·1 M-NaCl at various pH values after centrifugation of the resulting precipitate. The protein concentration was determined by measurement of the extinction of the supernatant at 277 nm.

Crystallization of derivatives. Crystallization studies were performed at the Laboratory of Molecular Biophysics, University of Oxford, by Dr G. G. Dodson. The crystals were grown from sodium acetate buffer, pH 5·8, as described by Schlichtkrull (1956) for the preparation of 2 Zn rhombohedral insulin. The X-ray-diffraction pattern of glycine monoacetoacetyl insulin was obtained by using Cu $K\alpha$ radiation with a Ni filter as a precession photograph (μ/21°). The time of exposure was 12 hr.

Biological assay. The various insulin derivatives were assayed against a neutral insulin solution (Nuso insulin batch A.354) as standard, by the mouse-convulsion method as described in the British Pharmacopoeia (1968; Appendix XVC, p. 1339). All injections were of solutions in acetate buffer, pH 8·0, and the assay was performed at the Wellcome Research Laboratories by Dr G. A. Stewart.

Removal of acetoacetyl groups. Acetoacetylated insulin (5 mg.) was treated with 100 μl. of the reagents listed in Table 3. After a given time the mixture was made 7 M with respect to urea and a constant portion (5 μl.) was applied to a cellulose acetate strip. After electrophoresis, the relative amounts of the products formed were estimated visually. With trifluoroacetic acid the acid was removed under vacuum and 7 M-urea (100 μl.) was added to the residue.

An insulin control was used throughout to determine whether modification by any of the reagents had occurred. In no case was this observed. However, the use of higher concentrations of the reagents caused precipitation of material that was not soluble in urea.

Diacetoacetyl insulin (80 mg.) was treated with hydrazine acetate solution (2 M, 2·5 ml.) for 12 hr. at 37°. Urea (0·9 g.) was added to make the material 5 M with respect to urea and the product was extensively dialysed against distilled water. A precipitate formed and was centrifuged off; it was found to be insoluble under all conditions. The yield of protein in the clear supernatant was less than 25 mg.

RESULTS

The effect of increasing the concentration of diketen on the reaction of insulin with ninhydrin is shown in Fig. 1. A decrease of approx. 30% of the ninhydrin value occurs after a threefold molar excess of diketen/amino group. A tenfold molar excess of diketen gives a 60% decrease in the initial ninhydrin value. After this point only a slight decrease in the ninhydrin colour value occurs. These results suggest that the reaction with diketen occurred in three distinct steps.

Chromatographic separation. Chromatography results for insulin (500 mg.) treated with a threefold molar excess of diketen can be seen in Fig. 2. Four main peaks are discernible. The position of elution of unmodified insulin is shown on the figure. Component I was mainly unmodified insulin; component IV (80 mg.) was chromatographically homogeneous, requiring a higher salt concentration for elution, and would be expected to be more highly substituted. No further protein was eluted.

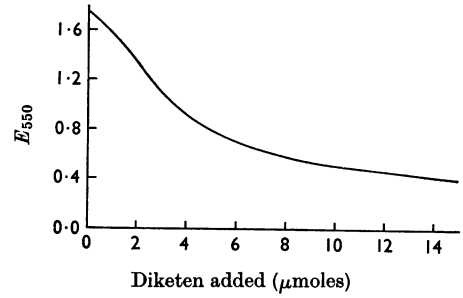

Fig. 1. Reaction of zinc insulin (0·167 mM) with diketen. The extent of reaction was determined by removal of 0·35 μmole portions of insulin for development with ninhydrin. The ordinate records the extinction of the ninhydrin colour at 550 nm.

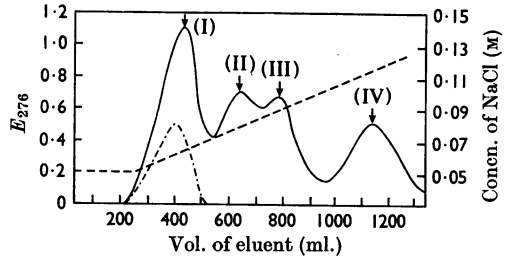

Fig. 2. Chromatographic separation of insulin (500 mg.) treated with a threefold molar excess of diketen per amino group, on a 40 cm. × 2·5 cm. column of DEAE-Sephadex A-25, pH 7·20. ——, E at 276 nm; – – –, gradient of NaCl applied; –·–·, elution pattern of zinc insulin on same column.

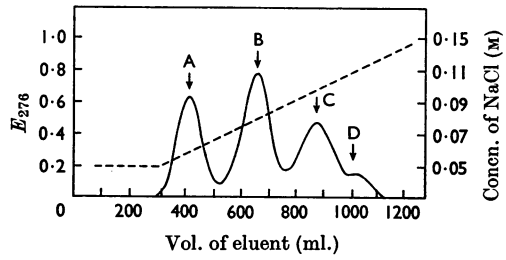

Fig. 3. Rechromatography of fractions II and III (250 mg.) on a 40 cm. × 2·5 cm. column of DEAE-Sephadex A-25, pH 7·20. ——, E at 276 nm.; – – –, gradient of NaCl applied.

Material from peaks II and III (250mg.) was combined and re-chromatographed under exactly the same conditions. The elution pattern is shown in Fig. 3. This material was divided into three chromatographically homogeneous fractions A, B and C, which together with fraction IV from the previous chromatography were further characterized. To demonstrate that fractions B and C were chromatographically homogeneous they were run again on a column of DEAE-Sephadex at pH 7·30, with a gentler salt gradient (from 0·05 to 0·10M-sodium chloride). The elution pattern (Fig. 4) showed only two components; the position of elution of insulin and component IV under the same conditions are shown on the figure.

The elution pattern of insulin (500mg.) formed by treating insulin with a 12-fold molar excess of diketen can be seen in Fig. 5. Only two peaks are produced, the first corresponding in elution volume to fractions IV; a new component, fraction E, was eluted at a much higher salt concentration.

Characterization of derivatives. (*a*) End-group analysis. Dansylation of each of the isolated derivatives gave the results indicated in Table 1.

Carbamoylation of the monosubstituted derivatives was done to determine the relative purity of these derivatives and confirmed the end groups found by dansylation.

Analysis of the number of acetoacetyl groups gave values roughly in agreement with the degree of modification found by end-group analysis.

(*b*) Cellulose acetate electrophoresis. Modification of insulin by acetoacetyl groups effectively increases the number of resulting charges on the molecule by -1 for each substitution. The results (Fig. 6) show that each of the separated derivatives behaves as a pure component and is not contaminated by other derivatives. It was not possible by this technique to distinguish between the two monosubstituted insulins. Insulin (derivative A) was assumed to have a relative net charge of -2 under these conditions.

(*c*) TNBS analysis. Data from the rate of reaction of zinc insulin with TNBS were analysed by the method described by Freedman & Radda (1968) in terms of three distinct reactions with differing rate constants. Results are shown in Fig. 7. Data from 200 to 1100min. were used to produce a semi-logarithmic plot, the experimentally found value of E_∞ being used. The values of $1{\cdot}15 \times 10^4 M^{-1}\,cm.^{-1}$ for the molar extinction of the slowest reacting group and $1{\cdot}00 \times 10^4 M^{-1}\,cm.^{-1}$ for each of the two faster-reacting groups (Goldfarb, 1966) were used to calculate the pseudo-first-order rate constant.

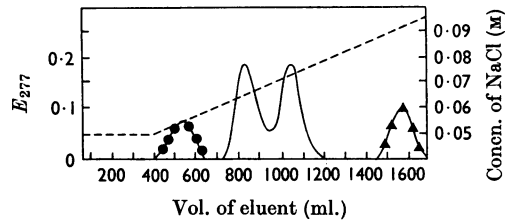

Fig. 4. Chromatographic separation of the monoacetoacetylated insulins on a 40cm. × 2·5cm. column of DEAE-Sephadex A-25, pH 7·30, with a gentle salt gradient (---) from 0·05M-NaCl to 0·10M-NaCl: ●, position of elution of insulin; △, position of elution of diacetoacetylated insulin.

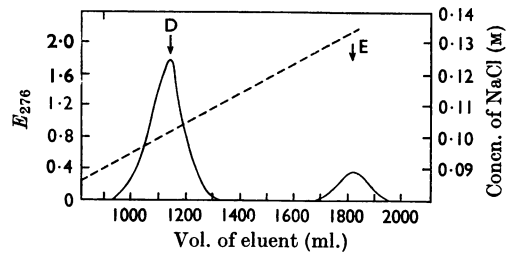

Fig. 5. Chromatographic separation of insulin (500mg.) treated with a 15-fold molar excess of diketen per amino group, on a 40cm. × 2·5cm. column of DEAE-Sephadex A-25, pH 7·20. ——, E at 276nm.; ---, gradient of NaCl applied.

Table 1. *Characterization of the acetoacetylated insulins*

I.E.P. denotes the pH at which derivatives show minimum solubility. For insulin this is equal to the isoelectric point.

Derivative	End-group analysis: Dansylation	End-group analysis: Carbamoylation	No. of acetoacetyl groups (±0·5)	I.E.P. (±0·2)
A	Gly, Phe, ε-Lys		—	5·5
B	Gly, ε-Lys	90% Gly 10% Phe	1·2	5·2
C	Phe, ε-Lys	15% Gly 85% Phe	1·5	5·2
IV	ε-Lys		2·4	4·9
E	—		3·6	4·4

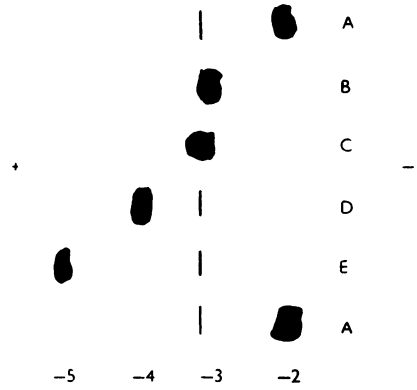

Fig. 6. Separation of the purified acetoacetyl derivatives of insulin on cellulose acetate strips. A, Unchanged insulin. B, Insulin modified on phenylalanine B-1. C, Insulin modified on glycine A-1. D, Insulin modified on phenylalanine B-1 and glycine A-1. E, Insulin modified on phenylalanine B-1, glycine A-1 and ϵ-lysine B-29.

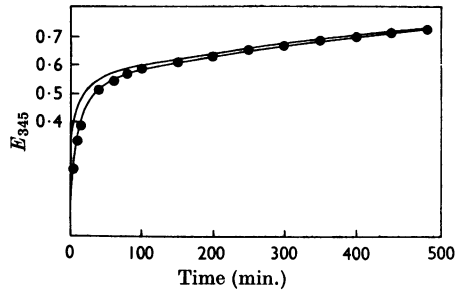

Fig. 7. Analysis of the reaction of insulin with TNBS. ——, Rate curve for unmodified insulin, calculated from the contribution to E at 345nm. at various times from each of the second-order rate constants k_1, k_2 and k_3. ●, Observed rate curve for derivative A.

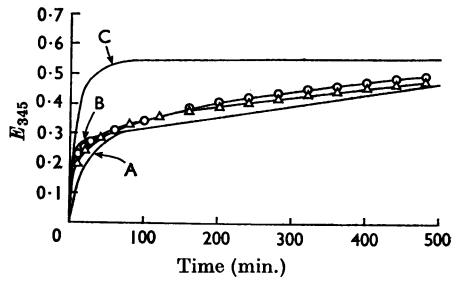

Fig. 8. Analysis of the reaction of monosubstituted insulins with TNBS. Calculated rate curves: A, for k_1+k_2; B, for k_1+k_3; C, for k_2+k_3. ○, Observed rate curve for derivative B; △, observed rate curve for derivative C.

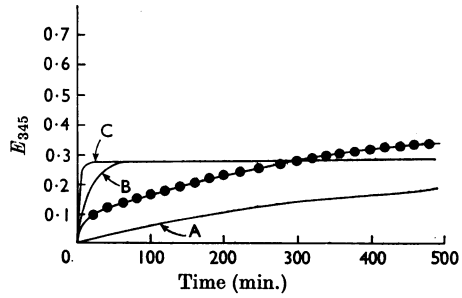

Fig. 9. Analysis of the reaction of disubstituted insulin with TNBS. Calculated rate curves: A, for k_1; B, for k_2; C, for k_3. ●, Observed rate curve for derivative D.

Subtraction of the increment in extinction due to the slowest-reacting set from the total gave data for the faster-reacting sets, which were used to produce further semi-logarithmic plots and hence calculate the corresponding rate constants.

Application of the above method gave these second-order rate constants for zinc insulin: k_1, $0{\cdot}6\,\text{M}^{-1}\text{min.}^{-1}$; k_2, $17{\cdot}0\,\text{M}^{-1}\text{min.}^{-1}$; k_3, $128\,\text{M}^{-1}\text{min.}^{-1}$. Zinc-free insulin, analysed as described above, reacted with TNBS to give second-order rate constants that, within the limits of experimental error, were in agreement with those found with zinc insulin.

As the application of this method gives a higher cumulative error the faster the rate, the values of the rate constant for the two fastest-reaching groups showed a wide scatter on repetition of the experiment. As large an error as $\pm 40\%$ could be observed in the fastest rate and the second fastest rate showed $\pm 20\%$ variation.

Calculation with the above rate constants gave theoretical rate plots that were compared with the experimental data obtained from causing each one of the products isolated by DEAE-Sephadex chromatography to react with TNBS. These results are shown in Figs. 7, 8 and 9.

Diacetoacetyl insulin, although giving an overall higher extinction at a given time than would be expected from the calculated data, gave a second-order rate constant of $0{\cdot}6\,\text{M}^{-1}\text{min.}^{-1}$.

Biological activity. Results, expressed in terms of the Fourth International Standard for insulin, are shown in Table 2 and were determined by the mouse-convulsion assay. Times to convulsion for the various derivatives were not significantly different from those for native insulin, nor was any prolongation of effect observed.

DISCUSSION

The pK of the ϵ-amino group of insulin (B-29) has been determined as 9·6–11·1 from titration data obtained by Tanford & Epstein (1954). A pH

Table 2. *Biological assay of the acetoacetylated insulins*

Material	Statistical wt. of assay	Potency (units/mg.)	Fiducial limits of error (P 0·95)
Native ox insulin (Boots batch no. 26918)	814	21·5	18·4–25·2
Unmodified insulin (fraction A)	847	16·3	14·0–19·1
Monoacetoacetyl insulin (fraction B)	723	12·6	10·7–14·9
Monoacetoacetyl insulin (fraction C)	1040	10·8	9·3–11·4
Diacetoacetyl insulin (fraction IV)	1296	9·5	8·4–10·8
Triacetoacetyl insulin (fraction E)	1614	10·4	9·3–11·7

of 6·9 was chosen as the ϵ-amino group would be completely protonated, unlike the α-amino groups, and would be less likely to enter into reaction with diketen, thus decreasing the number of possible products formed and making the separation problems easier.

The chromatography of the modified insulins showed that it was possible to distinguish between monosubstituted phenylalanine and glycine acetoacetylated insulins. A monosubstituted derivative modified on lysine B-29 would have a complete charge difference from native insulin at pH 7, unlike the monosubstituted glycine and phenylalanine derivatives, and should easily be separated from the latter derivatives by DEAE-Sephadex, appearing after the monosubstituted glycine acetoacetyl insulin. No such derivative was detected. Furthermore, chromatography of the material treated with a 15-fold molar excess of diketen showed the presence of only one disubstituted derivative. Derivatives modified on ϵ-lysine would be well separated by chromatography owing to the greater charge difference. No trace of such derivatives was found and it can be concluded that only when all of the phenylalanine and glycine groups are modified does reaction with the ϵ-amino group of insulin occur under the reaction conditions described in this paper.

As *O*-acetylation of tyrosine residues in proteins is known, each one of the derivatives formed by acetoacetylation of insulin was examined spectrophotometrically. Each derivative showed a molar extinction coefficient at 277nm. of $5{\cdot}3 \times 10^3 M^{-1}$ cm.$^{-1}$, which is not significantly different from that of unmodified insulin. Balls & Wood (1956) showed that modification of tyrosine by an *O*-acetyl group decreases the molar extinction coefficient at 277nm. from 1230 to 70. Scoffone *et al.* (1968) treated the reaction products, formed by reaction of ribonuclease with diketen, with 0·2M-sodium carbonate–sodium bicarbonate buffer, pH 9·5, for 11hr. at 25°. This treatment removed all the acetoacetyl groups bound to residues other than amino groups As the reaction products with insulin were similarly treated, any *O*-acetoacetyltyrosine residues formed in the initial reaction must have been removed. Under these conditions, diketen would appear to be a selective reagent for amino groups in proteins.

Evidence that diketen did not react with groups other than amino groups in insulin was also obtained by determining the number of carbonyl groups formed in the modified derivatives. Although the reaction with 2,4-dinitrophenylhydrazine was not highly reproducible, it did give results that were in accord with the other evidence presented in this paper.

Because of the design of the acetoacetylation reaction it has not been possible to determine the rate of reaction of diketen at pH 6·9 with each of the free amino groups of insulin. As ϵ-lysine does not react under these conditions, a graph of log (reactivity of lysine B-29 towards diketen) against pK must have a slope of less than unity. Chromatography also reveals that the derivative modified on the *N*-terminal phenylalanine is formed in a quantity two- to three-fold that of the derivative modified on the *N*-terminal glycine. The order of reactivity with diketen is therefore Phe > Gly $\gg$ ϵ-Lys. This order of reactivity was found also in the reaction of insulin with isothiocyanates (Africa, 1968; Bromer *et al.* 1967; Tietze *et al.* 1962).

Chromatography of unchanged zinc insulin gave no observable evidence of the presence of deamido forms of insulin. Some preparations of insulin are known to contain these forms by acid treatment of the hormone (Slobin & Carpenter, 1963). Deamido insulin should show a similar elution volume to that for fraction B. However, carbamylation of fraction B gave an analysis of 90% glycine and 10% of phenylalanine as *N*-terminal residues. This result indicated that fraction B was at least 90% pure and was contaminated by only small amounts of native insulin, deamido insulin or fraction C. However, assessment of the carbamylation reaction may be subject to error, particularly in the analysis of glycine (Stark, 1967).

TNBS analysis. The value of the second-order rate constant of the slowest-reacting group in insulin, derived from the rate of reaction of zinc insulin with TNBS, was in excellent agreement with that found by Freedman & Radda (1968). The two faster rate constants were appreciably higher,

however. Because of the errors inherent in the analysis of rate data by the method of subtraction of semi-logarithmic plots it is not possible to quote accurate data for the two faster-reacting groups, unless a large number of experimental readings are taken in the first minutes of reaction. The rate of reaction of the faster group with TNBS is so high that the reaction is complete within 5 min., whereas the amino group of intermediate rate has completely reacted after 70 min.

Analysis of the acetoacetylated derivatives of insulin with TNBS shows that in every case the overall fit was in accordance with the degree of modification suggested by the other evidence presented in this paper. However, the method was too inaccurate to distinguish between the two monosubstituted acetoacetyl derivatives or to give a reasonably close fit for the data from 0 to 100 min., that is during the reaction time of the two fast-reacting groups. It is precisely during this period that accurate data are needed if the rates are to be used to identify the reacting groups.

Although it is not possible to assign rate constants to the two α-amino groups of insulin in the reaction with TNBS, it is evident from the TNBS data that the group reacting with a second-order rate constant of $0{\cdot}6\,\text{M}^{-1}\text{min.}^{-1}$ is the ϵ-amino group of lysine B-29. The claim by Li (1956) that 2,4-dinitrobenzenesulphonic acid reacts exclusively with the lysine B-29 of insulin would seem unlikely from the above evidence.

Triacetoacetyl insulin does in fact react with TNBS although at a lower rate than the slowest-reacting amino group. As the material was extensively dialysed it is unlikely that TNBS was reacting with anything other than protein. An explanation could be that increasing modification of the amino groups of insulin affects the aggregation properties of the hormone to such an extent that groups other than amino groups can react with TNBS.

The assumption on which these calculations are based, namely that modification of one amino group in insulin does not influence the rate of reaction of other groups in the molecule, might not be justified, particularly for the glycine and phenylalanine end groups. Anderson (1956*a*) studied the rate of reaction of phenyl isocyanate with insulin and concluded that substitution of one *N*-terminal group of insulin did affect the reactivity of the other.

X-ray-diffraction studies. The only derivative to crystallize under Schlichtkrull's conditions for 2 Zn rhombohedral insulin was the monoacetoacetyl derivative modified on the glycine *N*-terminus. Slight variations in the pH of the crystallizing buffer, to account for the differences in the isoelectric point of the modified derivatives, gave only amorphous precipitates for all the other derivatives. An X-ray-diffraction pattern of the glycine modified monoacetoacetyl insulin (Fig. 10) showed it to be isomorphous with 2 Zn ox insulin. This is the first time that a chemical derivative of insulin has been crystallized in the rhombohedral form and it has important implications for the preparation of heavy-atom derivatives of the hormone by amino group modification.

Work is in progress to try to produce suitable derivatives for phase determination by diffusion of heavy-metal salts into crystals of the glycine monoacetoacetyl insulin.

From the molecular-weight determinations of Marcker (1960) on phenylcarbamoyl insulins, and the sedimentation studies of Anderson (1956*b*) on phenylthiocarbamoyl insulins, the conclusions of these workers were that the site of attachment of zinc in the hormone was on the *N*-terminal phenylalanine. Graae (1968) has suggested the participation of an α-amino group in zinc binding from titration studies on insulin in the presence of various metal salts. However, results at variance with the above have been obtained by Tanford & Epstein (1954), who examined the titration curves of zinc and zinc-free insulin, and by Brill & Venable (1968), who studied the electron-spin-resonance spectrum of cuprous insulin crystals. These workers concluded that histidine was involved in zinc binding.

However, Marcker's (1960) and Anderson's (1956*b*) results could equally well be explained by assuming that amino group modification in insulin profoundly affects the packing of the insulin molecule in the

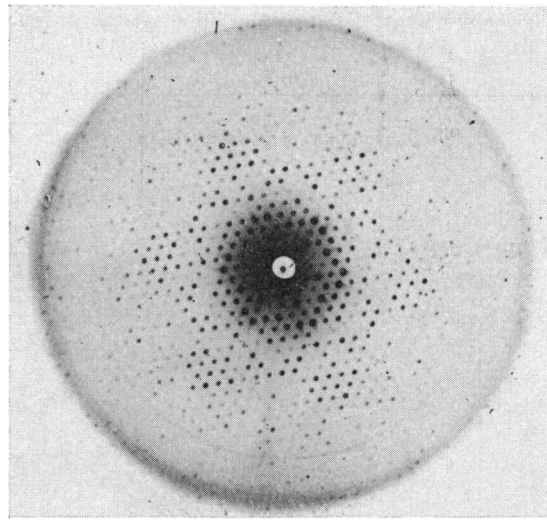

Fig. 10. Diffraction pattern of glycine modified monoacetoacetyl insulin.

rhombohedral crystal, but this need not be related to the zinc-binding site. Such modification could affect the ability of the insulin monomers to form closely knit dimers, three such dimers per two zinc atoms being required to give hexagonal packing in 2 Zn insulin. Another explanation is that the ability of the dimers to ion-pair bond with each other is affected.

The difficulty in crystallizing the monoacetoacetyl phenylalanine insulin suggests that, because of the way the insulin molecules are packed in the crystal, there is less space for additional groups on the phenylalanine *N*-terminus than on the glycine *N*-terminus. The phenylalanine group may also be more sensitive to charge differences than glycine and the presence of a free amino group on phenylalanine, as in the native hormone, may facilitate crystallization.

However, a full explanation of the aggregation phenomena of insulin and the nature of the zinc-binding site should emerge from the crystallographic analysis of the hormone.

Biological activity. The gradual decrease in activity with increasing substitution of the free amino groups in insulin has been verified by other workers (Anderson, 1956*a*; Bromer *et al.* 1967; Levy & Carpenter, 1967; Mills, 1953). Similar results are obtained for the acetoacetyl-modified insulins and appear to contradict the results of Fraenkel-Conrat. The small difference in size between the acetyl and acetoacetyl groups would not be expected to account for this difference. A specific deacylase capable of removing acetyl groups and regenerating full activity could explain this discrepancy but no evidence for this exists.

The fact that the two monosubstituted derivatives show approximately the same activity is surprising. Anderson (1956*a*) caused insulin to react with phenyl isocyanate and isolated a monophenylcarbamoyl insulin and presented evidence that it was more than 90% of the glycine-modified derivative. The activity of this derivative was close to that of the native hormone. Similar results were obtained by Mills (1953) on treatment of insulin with 2,4,5-trinitrotoluene. No appreciable decrease in activity of insulin was observed until 1·5 amino groups had been modified and the glycine end group was modified much more rapidly than the phenylalanine *N*-terminus. However, the biological assays performed were less extensive than those reported here. If the aggregation properties of insulin are important for its biological activity, then it is not surprising that a decrease of activity occurs on modification of the phenylalanine group. As the glycine-modified monosubstituted derivative is isomorphous with the native hormone, modification cannot have affected the aggregation or configuration of insulin. The fact that this derivative shows decreased activity suggests that the glycine *N*-terminus of insulin is important for hormone activity.

The biological activity of triacetoacetylated insulin is the same as the diacetoacetylated material. This suggests that the B-29 ϵ-amino group is not necessary for the convulsion-causing activity of insulin. This conclusion would, however, be more effectively demonstrated by examining derivatives of insulin modified solely on lysine B-29. The possibility of synthesizing these derivatives is being investigated.

Removal of acetoacetyl groups. The results in Table 3 show that the ready regeneration of native protein from acetoacetyl-modified proteins as described by Scoffone *et al.* (1968) was not achieved with insulin. Much higher concentrations of reagents and longer reaction times were needed to

Table 3. *Removal of the acetoacetyl protecting group by the use of various reagents and conditions*

Derivative	Reagent	pH	Time (hr.)	Temp.	% Products formed (approx.) Insulin	Mono-substituted	Di-substituted
IV	$H_2N{\cdot}OH$ (sixfold excess)	7·0	12	25°			100
IV	0·5M-$H_2N{\cdot}OH$	7·0	12	37	10	60	30
IV	1·0M-$H_2N{\cdot}OH$	7·0	12	25	10	60	30
IV	1·0M-$H_2N{\cdot}OH$	8·5	12	37		60	40
B	0·5M-$H_2N{\cdot}OH$	8·5	12	25			100
C	0·5M-$H_2N{\cdot}OH$	8·5	12	25	40	60	
IV	$CF_3{\cdot}CO_2H$		1	25			100
IV	$CF_3{\cdot}CO_2H$		1·5	25			100
IV	0·5M-$H_2N{\cdot}NH{\cdot}CO{\cdot}CH_3$*	8·0	48	37	50	30	20
IV	1·0M-$H_2N{\cdot}NH{\cdot}CO{\cdot}CH_3$	8·0	24	37	40	60	
IV	2·0M-$H_2N{\cdot}NH{\cdot}CO{\cdot}CH_3$	8·0	12	37	100		
IV	2·0M-$H_2N{\cdot}NH{\cdot}CO{\cdot}CH_3$	6·2	36	25	70	30	

* Conditions used by Geiger & Siedel (1968) for the removal of the *N*-formyl protecting group.

remove the acetoacetyl groups. Experiments on the two monoacetoacetylated insulins suggest that it is more difficult to remove the acetoacetyl group from phenylalanine B-1 than from glycine A-1. A large-scale reaction of diacetoacetyl insulin with 2M-hydrazine acetate at pH 8·0, although giving complete removal of the acetoacetyl groups present, caused a 60% irreversible denaturation of protein. With 2M-acetylhydrazine at pH 6·2 protein denaturation did not occur, but complete regeneration of insulin did not occur even with a reaction time of 36hr.

Experiments are in progress to find a reversible protecting group for the free amino groups of insulin that can be removed under less vigorous conditions, so that derivatives modified exclusively on lysine B-29 can be prepared and the effect on the activity and the crystallization properties of insulin can be examined.

We are indebted for the assistance given in this project by Dr G. G. Dodson of the Laboratory of Molecular Biophysics, University of Oxford, who undertook the crystallization and X-ray-diffraction studies, and to Dr G. A. Stewart of the Wellcome Research Laboratories, Dartford, Kent, for the biological assays of the various derivatives. Reproduction of Fig. 10 is by kind permission of Dr Dodson. This work was supported by a grant from the Medical Research Council.

REFERENCES

Africa, B. & Carpenter, F. H. (1968). *Fed. Proc.* **27**, 766.
Anderson, W. (1956*a*). *C.R. Lab. Carlsberg*, **30**, 104.
Anderson, W. (1956*b*). *C.R. Lab. Carlsberg*, **30**, 126.
Balls, A. K. & Wood, H. N. (1956). *J. biol. Chem.* **219**, 245.
Brill, A. S. & Venable, J. H. (1968). *J. molec. Biol.* **36**, 343.
Bromer, W. W. & Chance, R. E. (1967). *Biochim. biophys. Acta*, **133**, 219.
Bromer, W. W., Sheehan, S. K., Berns, A. W. & Arquilla, E. R. (1967). *Biochemistry* **6**, 2378.
Carpenter, F. H. & Hayes, S. L. (1963). *Biochemistry*, **2**, 1272.
Crowfoot, D. (1935). *Nature, Lond.*, **135**, 591.
Fraenkel-Conrat, J. & Fraenkel-Conrat, H. (1950). *Biochim. biophys. Acta*, **5**, 89.
Freedman, R. B. & Radda, G. K. (1968). *Biochem. J.* **108**, 383.
Geiger, R. E. & Siedel, W. (1968). *Chem. Ber.* **101**, 3386.
Goldfarb, A. R. (1966). *Biochemistry*, **5**, 2570.
Graae, J. (1968). *Biochem. J.* **106**, 777.
Gray, W. R. (1967). In *Methods in Enzymology*, vol. 11, p. 139. Ed. by Colowick, S. P. & Kaplan, N. O. New York: Academic Press Inc.
Jordon, D. E. & Veatch, F. C. (1964). *Analyt. Chem.* **36**, 120.
Lacey, R. N. (1954). *J. chem. Soc.* p. 850.
Levy, D. & Carpenter, F. H. (1967). *Biochemistry*, **6**, 3559.
Li, C. H. (1956). *Nature, Lond.*, **178**, 1402.
Marcker, K. (1960). *Acta chem. scand.* **14**, 2071.
Mills, G. L. (1953). *Biochem. J.* **53**, 37.
Moore, S. & Stein, W. H. (1954). *J. biol. Chem.* **211**, 907.
Naik, R. G. (1932). *J. Indian chem. Soc.* **9**, 185.
Sen, D. N. & Thankarajan, N. (1968). *Indian J. Chem.* **6**, 278.
Sen, D. N. & Umapathy, P. (1968). *Indian J. Chem.* **6**, 516.
Schlichtkrull, J. (1956). *Acta chem. scand.* **10**, 1455.
Scoffone, E., Marzotto, A., Pajetta, P. & Galzigna, L. (1968). *Biochim. biophys. Acta*, **154**, 450.
Slobin, L. I. & Carpenter, F. H. (1963). *Biochemistry*, **2**, 22.
Stark, G. R. (1967). In *Methods in Enzymology*, vol. 11, p. 125. Ed. by Colowick, S. P. & Kaplan, N. O. New York: Academic Press Inc.
Tanford, C. & Epstein, J. (1954). *J. Amer. chem. Soc.* **76**, 2163.
Tietze, F., Mortimore, G. E. & Lomax, N. R. (1962). *Biochim. biophys. Acta*, **59**, 336.