activity than heretofore reported. This is because all the components of such preparations have nearly the same biological activity. There was no indication of loss of total activity during a run, but it has not been feasible to assay quantitatively all the fractions in order to settle this point.

Acknowledgment.—We are greatly indebted to Dr. E. D. Campbell of the Eli Lilly Co., Sir Jack Drummond of the Boots Pure Drug Co., and Dr. K. Linderstrøm-Lang for generous supplies of insulin. All biological assays reported here were made by the companies supplying the insulin. The technical assistance of Miss Dorris McNamara is acknowledged. Microanalyses were made by Mr. D. Rigakos.

NEW YORK 21, N. Y.

[CONTRIBUTION FROM THE LABORATORIES OF THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH]

The Molecular Weight of Insulin

BY ELIZABETH J. HARFENIST AND LYMAN C. CRAIG

RECEIVED JANUARY 25, 1952

The molecular weight of insulin has been examined using the method of partial substitution by 1-fluoro-2,4-dinitrobenzene, separation of the reaction products and colorimetric analysis of the monosubstituted derivative for the dinitrophenyl group. A molecular weight of 6,500 has been found.

In a previous report[1] it was shown that it is possible to fractionate insulin preparations by countercurrent distribution and thereby obtain a preparation whose behavior approaches that of a single solute partitioning almost ideally. The next step in the characterization of insulin is therefore the determination of the molecular weight.

Molecular weights for insulin determined by several methods have already been published. The earlier work with diffusion and the ultracentrifuge[2] indicated values in the range 36,000 to 48,000, but more recently Gutfreund[3] observed that the sedimentation was less rapid at lower protein concentrations and *p*H values. As a result of osmotic pressure measurements he was able to assign a value of 12,000 for the molecular weight at infinite dilution. This value is in agreement with later studies.[4] The earlier higher values were interpreted as being due to association of monomer units in solution. Even the value of 12,000 has been questioned by Fredericq and Neurath[5] whose results indicated that in the proper environment insulin may be dissociated still further to a species with a molecular weight of about 6,000.

This state of affairs makes another completely independent approach to the question of molecular weight highly desirable. For this reason it appeared of great interest to learn whether the approach suggested by Battersby and Craig[6] for the molecular weight determination of the higher peptides could be applied to a substance of the complexity of insulin. The results of such studies up to the present time are reported in this paper.

(1) E. J. Harfenist and L. C. Craig, THIS JOURNAL, **74**, 3083 (1952).

(2) (a) B. Sjögren and T. Svedberg, *ibid.*, **53**, 2657 (1931); (b) G. L. Miller and K. J. I. Andersson, *J. Biol. Chem.*, **144**, 459 (1942); (c) H. Gutfreund and A. G. Ogston, *Biochem. J.*, **40**, 432 (1946).

(3) (a) H. Gutfreund, *ibid.*, **42**, 156 (1948); (b) H. Gutfreund, *ibid.*, **42**, 544 (1948).

(4) (a) J. L. Oncley, E. Ellenbogen, D. Gitlin and F. R. N. Gurd, *J. Phys. Chem.*, **56**, 85 (1952); (b) F. Tietze and H. Neurath, *J. Biol. Chem.*, **194**, 1 (1952). NOTE ADDED IN PROOF: While this paper was in press, two papers appeared supporting the value of 12,000 as the molecular weight of insulin: H. Gutfreund, *Biochem. J.*, **50**, 564 (1952); P. Doty, M. Gellert and B. Rabinovitch, THIS JOURNAL, **74**, 2065 (1952).

(5) E. Fredericq and H. Neurath, *ibid.*, **72**, 2684 (1950).

(6) A. R. Battersby and L. C. Craig, *ibid.*, **73**, 1887 (1951).

Experimental

The sample of insulin chosen for the initial molecular weight studies was one which was well characterized by countercurrent distribution and known to contain 90% or better of the A component. The method used in the work reported here was partial substitution by the Sanger reagent, 1-fluoro-2,4-dinitrobenzene.[7]

Five hundred milligrams of insulin (Boots Lot 9011 G) was dissolved in 25 ml. of a 1% aqueous solution of sodium bicarbonate at a temperature of 24°. To this was added a solution of 250 mg. of 1-fluoro-2,4-dinitrobenzene (FDNB) in 50 ml. of ethanol. At the same time a control solution, of the same composition except that the insulin was omitted, was made up. The progress of the reaction was followed by measuring the ultraviolet absorption at 350 mμ in a Beckman quartz spectrophotometer, using $1/_{20}$-cm. cells. The reaction was quenched by the addition of 0.25 ml. of glacial acetic acid when the optical density had reached 0.65, a value which corresponds to the substitution of 1.5 moles of FDNB per insulin molecule if the molecular weight is 12,000. The reaction time was about 15 minutes. Unreacted FDNB was removed by extraction with ether, followed by evaporation to dryness of the ether–alcohol layer. The residue was taken up in water and ether and the aqueous layer from this was added to the aqueous layer from the first extraction. The combined aqueous extracts were then evaporated to dryness in the rotary evaporator.[8]

The entire yellow residue was distributed in the system used for the original characterization of the protein, 2-butanol/1% aqueous dichloroacetic acid (DCA). A 220 tube automatic countercurrent distribution apparatus[9] was used. After 216 transfers analyses by weight and by ultraviolet absorption at 277 mμ and at 350 mμ gave the pattern shown in Fig. 1a. Bands 1, 2 and 3, all colorless, were withdrawn and the apparatus was arranged for recycling the yellow bands. At 530 transfers the pattern shown in Fig. 1b was obtained.

Material was recovered from bands 2–6 of Fig. 1a. Band 1 was found to contain only inorganic salts from the reaction. Band 2 contained unchanged insulin which was isolated in crystalline form by the usual procedure. Since intact solute from band 3 was not needed for study, material for hydrolysis was recovered by the addition of hydrochloric acid, extraction of the DCA and 2-butanol with ether, and evaporation of the aqueous layer. A portion of the residue was hydrolyzed and studied by paper chromatography, but no spots other than those obtained from an hydrolysate of insulin were observed.

Products from bands 4, 5 and 6 were all isolated in the following manner. The solutions were concentrated *in*

(7) F. Sanger, *Biochem. J.*, **39**, 507 (1945).

(8) L. C. Craig, J. D. Gregory and W. Hausmann, *Anal. Chem.*, **22**, 1462 (1950).

(9) L. C. Craig, W. Hausmann, E. H. Ahrens, Jr., and E. J. Harfenist, *ibid.*, **23**, 1236 (1951).

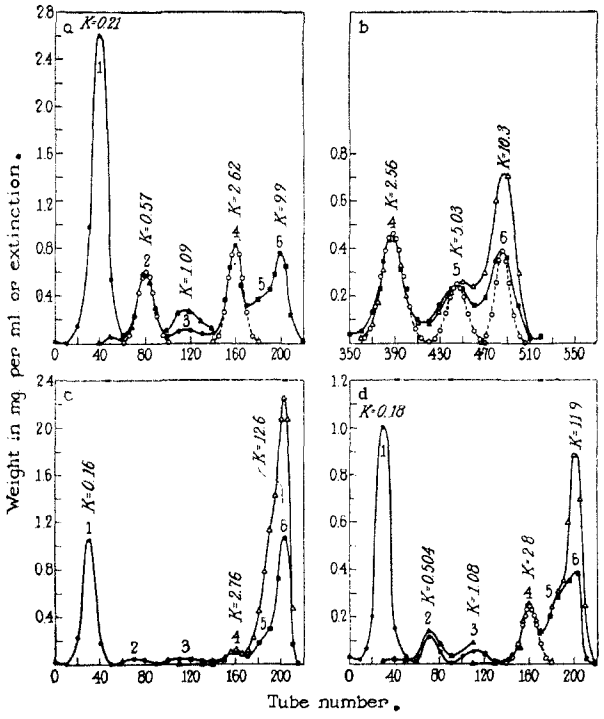

Fig. 1.—Distribution patterns of partially substituted DNP-insulin: a, reaction time of 16 minutes, 217 transfers; b, 530 transfers; c, reaction time of 38 minutes, 218 transfers; d, reaction time of 21 minutes, 218 transfers; ●, weight, lower layer; ■, weight, upper layer; O, theoretical; ▲, extinction × 1.3 at 277 mμ, lower layer; Δ, extinction × 0.5 at 350 mμ, upper layer.

vacuo at a low temperature until most of the 2-butanol had evaporated and a yellow solid began to precipitate. The mixture was then extracted with ether to remove the free DCA, and lyophilized. Analytical studies were made on the material thus isolated from band 4. A chlorine analysis was made to determine the percentage of DCA in the residue. Loss of drying (100°, 0.2 mm.), 8.56%; Cl, 6.37%. A solution of 11.65 mg. of the yellow residue in 25 ml. of the upper phase of the DCA system gave an extinction of 0.810 at 350 mμ. Material from band 5 was useless for molecular weight determinations because a control experiment described below indicated the presence of a yellow substance other than an insulin derivative, probably 2,4-dinitrophenol. However, some of the fraction was hydrolyzed and studied by paper chromatography. No spots other than those from insulin were found.

N-DNP-glycine was made[10] and its molecular extinction coefficient determined at 350 mμ in the upper phase of the system; ε 14,360.

Discussion

The reasoning involved in deriving molecular weights by the method used in the present work has been previously set forth.[6] With any unknown substance a number of experimental requirements must be met before such reasoning can be applied. These are the following: (1) The distribution behavior of a good sample of the substance must be well known. (2) A sample of the substance must be permitted to react partially with some reagent capable of covering or liberating functional groups of high polarity so that for each group covered or liberated a significant shift in partition ratio is produced. (3) A weight distribution pattern of all the reaction products must give clear cut, well separated bands of constant composition throughout each band. (4) Quantitative analytical data for the substituting group in each band must be obtainable.

The first of these requirements has already been met as reported previously.[1] An attempt to satisfy the second with the use of FDNB was first made by using the conditions Sanger[7] gave for completely substituting the free $-NH_2$ groups of insulin except that only sufficient FDNB was used to substitute two of the six groups (assuming a molecular weight of 12,000). It was found experimentally that nearly half of the FDNB was transformed, presumably to 2,4-dinitrophenol, under these conditions. Therefore two moles of the reagent were used per mole of insulin. However, clear cut bands were not obtained and the distribution pattern gave evidence for transformation other than simple substitution.

Different reaction conditions were therefore tried. A large excess of FDNB was used and the reaction was quenched after a short time, the time being determined by following the development of absorption at 350 mμ in the Beckman quartz spectrophotometer. The first experiment was interrupted when the extinction corresponded to the substitution of 2.56 moles. Distribution gave pattern c of Fig. 1. Very little unchanged insulin appeared, and the reaction had obviously been allowed to proceed too far.

In the next experiment the time of reaction was reduced to allow the substitution of only 2.2 moles of FDNB. Distribution of the reaction products now gave pattern d of Fig. 1, showing that the reaction time was still too long. Reduction of the reaction time still further, to allow the introduction of only 1.5 moles, and distribution of the products gave pattern a of Fig. 1. This would appear to represent a suitable stage of substitution for molecular weight studies because of the comparable amounts of the various components. The colored bands in tubes 140–220 were recycled to 530 transfers in order to give a more complete separation as shown in Fig. 1b.

Two control experiments were carried out. In one the experiment which gave pattern a was repeated except that the FDNB was omitted. Only bands 1 and 2 were observed showing that insulin is not transformed under the basic conditions of the reaction. In the second control the same experiment was repeated but the insulin was omitted. Here the salt band and a yellow band were obtained. The yellow band had little weight and travelled at a rate which would make it occur at the position of band 5 of pattern b. Its absorption was sufficient in amount to account for almost half of the total absorption of band 5, and the experiment showed that 5 was a mixture unsuitable for use in molecular weight deductions. An additional experiment showed that 2,4-dinitrophenol also travelled at the same rate as band 5.

Band 4 is the yellow band occurring nearest to the unchanged insulin, band 2. It therefore should have only one of the basic nitrogens substituted and is of most interest for molecular weight determination. Pattern b shows that the weight, extinction

(10) E. Abderhalden and P. Blumberg, *Z. physiol. Chem.*, **65**, 318 (1910).

(350 mμ) and theoretical curves all approach agreement. The peak material was recovered and carefully dried. For the molecular weight calculations some known molecular extinction coefficient must be chosen. N-DNP-amino acid and peptide derivatives are known to have approximately the same molecular extinction coefficient at 350 mμ,[11] but to show individual variations of the order of ±2% depending on the solvent used. From Sanger's work one of the end groups of insulin is known to be glycine. N-DNP-glycine was therefore prepared and was found to have a molecular extinction coefficient of 14,360 in the upper phase of the system used for the distribution. If this value is assumed for insulin it permits the calculation of a molecular weight of 6,500 after making the necessary corrections for loss in weight on drying and DCA content, the latter calculated from the chlorine analysis. At this stage of our knowledge of the reliability of the method it would be unwise to regard this value as anything but the order of magnitude. However, it could well indicate the molecular weight to within ±10% of the true value. Further work to substantiate this conclusion by partial reaction with other reagents is in progress and will be reported at a later date.

It can be seen from pattern b that band 6 has an extinction/weight ratio approximately double that of band 4. It is therefore a di-DNP derivative.

Comparison of patterns a, c and d shows that in a longer reaction time the amount of 6 increases as the amount of 4 decreases, but the extinction/weight relationships in these bands hold throughout. This behavior strongly supports the use of band 4 as a reliable basis for molecular weight calculations.

Band 3 appears to be a derivative in which the reagent is attached at some point other than an $-NH_2$ group. Such a derivative would be colorless as is this band. The –OH group of tyrosine was suspected but complete hydrolysis followed by paper chromatography failed to give an O-DNP-tyrosine spot. The high extinction (277 mμ)/weight ratio compared to insulin, band 2, is indicative of DNP substitution at some point in the molecule but the nature of this derivative is not known. The histidine residue could be the point of attachment. No interpretation can be made of band 5 since it has been shown to be a mixture.

If the interpretation given to the experiments reported here is correct, then the molecular weight deduced by Gutfreund[3b] could be based on data obtained from a dimeric form and that of earlier workers[2] from material in still higher states of association. With this background of experience the question may rightly be raised as to whether or not even the lower value suggested here is that of the monomeric form. A number of observations speak against a smaller value.

It would be difficult to interpret the present observations on the basis of a molecular weight of 3,000, since the absorption of band 4 indicates the presence of one DNP-residue in a molecule of about 6,500 molecular weight, assuming that the value of the extinction coefficient of the N-DNP group is not significantly different in a larger molecule from that in an N-DNP-amino acid. To explain a molecular weight of 3,000 it would have to be assumed either that the extinction coefficient is only half as great in the insulin derivative, or that some sort of molecular complex has been formed between insulin and FDNB or an hydrolysis product of FDNB. Both of these assumptions seem unlikely.

More progress in structural work has been made with insulin than with any other protein. Sanger's extensive studies[12] have shown that oxidation of disulfide linkages results in the formation of two peptides, A and B. No other large fragments were found. The molecular weight of A was found to be about 2,700, and that of B about 3,800.[11] At the time of his publications Sanger postulated, on the basis of a molecular weight of 12,000, two A chains and two B chains in each insulin molecule. Since no evidence for two different A chains or two different B chains has thus far been obtained,[11,13] a molecular weight of about 6,500 would be entirely compatible with his findings, but a smaller subunit could not be explained in terms of the A and B peptides.

The published amino acid analyses[14] do not support a minimum molecular weight as small as 6,500. However, results from different laboratories are not in agreement with each other, perhaps because most insulin preparations are mixtures. At least all of those studied in this Laboratory[1] have shown from 10 to 50% of protein components other than the main insulin component. However, the published values are even more strongly in disagreement with a molecular weight smaller than 6,000 since this would require less than one residue each of lysine, proline, arginine, isoleucine and threonine in a molecule of insulin, and fractions of residues of many other amino acids.

The technical assistance of Miss Dorris McNamara is acknowledged.

NEW YORK 21, N. Y.

(11) F. Sanger, *Biochem. J.*, **45**, 563 (1949).

(12) F. Sanger, *ibid.*, **44**, 126 (1949).

(13) F. Sanger and H. Tuppy, *ibid.*, **49**, 463, 481 (1951).

(14) C. Fromageot, *Cold Spring Harbor Symposia on Quant. Biol.*, **14**, 49 (1950).