chloroform melted at 214–216°. The pure ester of the terephthalic acid is recorded[11] as melting at 225°. Neutralization equivalent of the mixture of acids was 84.3; calculated for phthalic acid 83.3. No trace of benzoic acid could be detected among the oxidation products. Apparently the mixture is largely isophthalic acid. The yield calculated on the basis of four propyl chloride molecules to one of the butylpropylbenzene mixture varied from 14 to 38%, an estimation based upon the quantity of high boiling residue left after distilling the butylbenzene from each experiment.

About 6 to 12 g. of a mixture of saturated and unsaturated compounds, boiling in the hexane and nonane range, was obtained from each of these preparations, the quantity being larger the higher the temperature of reaction. Careful fractionation in an eighteen-plate column failed to show any single product. All fractions also showed presence of unsaturated material. Upon ozonolysis of a 63 to 66° cut, decomposition of the ozonides with zinc and water, and careful fractionation of the resulting aldehydes and ketones, definite evidence in the form of the dinitrophenylhydrazone derivative was obtained for the presence of acetaldehyde. The melting point observed was 146–150°; that of an authentic sample was 147.5–148°; and that of the mixture was 148–150°. A mixed melting point with the like derivative of propionaldehyde (m. p. 190°) was 141°. Owing to the small amount of material left after attempts at purification, a satisfactory melting point for 2-butanone-2,4-dinitrophenylhydrazone could not be secured but the crystals obtained were shown to be identical with an authentic sample in having parallel extinction, pleochroism, centered acute bisectrix, and elongated obtuse bisectrix, and were not in whole agreement with observations on known samples of the like derivatives of acetone and butyraldehyde. A trace of propionaldehyde may also be present, for a few crystals prepared from one of the fractions appeared to have similar optical properties.[12]

Summary

Organosodium compounds from *n*-butyl and *n*-propyl chlorides have been prepared, the ease of reaction being progressively less, the yields lower, and the ratio of di- to monocarboxylic acid greater than observed with the amyl homolog.

n-Butyl- and *n*-propylsodium are progressively less reactive with benzene and toluene than is *n*-amylsodium.

n-Butylbenzene can be prepared conveniently by adding propyl chloride to sodium in toluene at 72°.

Metalation of butylbenzene takes place largely in the meta position.

(12) We are greatly indebted to Mr. Gibb for assistance on the optical studies.

CAMBRIDGE, MASS. RECEIVED OCTOBER 7, 1940

[CONTRIBUTION FROM THE CHEMICAL LABORATORY OF THE JOHNS HOPKINS UNIVERSITY]

The Geometrical Attack on Protein Structure

BY DOROTHY M. WRINCH

It is unnecessary at the present time to state the case for the cyclol hypothesis, since authoritative accounts have already been given by Langmuir of the way in which the theory accounts satisfactorily for many of the well-known properties of the globular proteins.[1,2] In a recent summary,[3] however, Pauling and Niemann repeat a number of statements purporting to disprove the theory already made by other writers. Attention must therefore be directed to a number of publications in which these criticisms have already been discussed, at least so far as their scientific importance appeared to warrant.[2,4–9] We then proceed to discuss, necessarily in a preliminary manner, two issues (first raised in my publications[10,11]) to which Pauling and Niemann refer, namely (a) certain short interatomic distances in the original cage structures and (b) the energy of formation of these structures. In particular it is shown that the claim by Pauling and Niemann that the cyclol hypothesis can be disposed of by means of Anson and Mirsky's heat of denaturation of trypsin must be rejected.

(a) In my first studies of possible protein structures, very exacting metrical conditions were adopted, mainly in order to demonstrate in a simple manner the possibility of handling problems of protein structure by strictly mathematical methods. These, it appears from crystallographical data, are unnecessarily onerous. Evidently a discussion of certain short interatomic distances in the original cyclol structures can only

(1) Langmuir, *Cold Spring Harbor Symposia Quant. Biol.*, **6**, 135 (1938).
(2) Langmuir, *Proc. Phys. Soc.* (London), **51**, 542 (1939).
(3) Pauling and Niemann, THIS JOURNAL, **61**, 1860 (1939).
(4) Wrinch, *Cold Spring Harbor Symposia Quant. Biol.*, **6**, 122 (1938).
(5) Langmuir, *Nature*, **143**, 280 (1939).
(6) Langmuir and Wrinch, *ibid.*, **143**, 49 (1939).
(7) Wrinch, *ibid.*, **143**, 482 (1939).
(8) Wrinch, *ibid.*, **143**, 763 (1939); **145**, 660, 1018 (1940).
(9) Langmuir and Wrinch, *Proc. Phys. Soc.* (London), **51**, 613 (1939).
(10) Wrinch, *Proc. Roy. Soc.* (London), **A160**, 59 (1937); **A161**, 505 (1937).
(11) Wrinch, *Nature*, **138**, 241 (1936).

be profitable if accompanied by a detailed study of the propriety and the feasibility of relaxing the metrical conditions. Further, it must be remembered that cases are known (*e. g.*, in the terpenes, the sterols and phenanthrene derivatives) in which the accumulated evidence of organic chemistry makes it necessary to place certain nonbonded atoms nearer together than the minimum distances allowed by Pauling and Niemann. These distances in any case refer to interatomic distances in different molecules. Thus the statement that methyl groups in the case of hexamethylbenzene are 4.0 to 4.1 Å. apart refers to intermolecular distances. Within a single molecule of hexamethylbenzene, methyl groups approach as near as 2.93 Å. Further, these distances have been derived from a study of simple molecules which bear no resemblance to the very complex highly organized structures which, on any plausible theory of protein structure, must characterize the globular proteins. Nothing is known as to the nearest approach of groups *within a single molecule* in such cases. (b) It is known that, under certain closely defined conditions, the heat of denaturation of trypsin amounts to less than 1 kcal./mole of amino acid residues.[12] Pauling and Niemann use this figure in an attempt to disprove the cyclol hypothesis as follows: (1) they claim that a cyclol cage structure is about 28 kcal./mole of amino acid residues less stable than the corresponding linear polypeptide containing NH·CO groups: (2) they claim that Anson and Mirsky's denatured trypsin consists of linear polypeptides. They then deduce that the trypsin molecule cannot be a cyclol cage structure, since if it were the heat of denaturation obtained by Anson and Mirsky would have been about −28 kcal./mole of amino acid residues. It is my opinion that both claims (1) and (2) were unfounded, and that the suggested deduction therefore falls to the ground.

(2) Little or nothing is known as to the structure of any denatured protein, nothing as to the structure of the denatured trypsin under discussion. It is assumed by Pauling and Niemann that denaturation on the cyclol theory means the opening of all cyclol bonds. This is not the case. (In particular the difference between trypsin and denatured trypsin in this case cannot consist in opening of all cyclol bonds, since, as Mirsky and Pauling state,[13] there is no change in molecular weight.) Denaturation, on the cyclol therory, corresponds to any breakdown of the intact structure produced by the opening of any subset of cyclol bonds.[14] The theory thus readily offers an interpretation of the well-known fact that very different denatured products can be obtained from one and the same protein (*cf.* the film studies of pepsins which have been shaken for different times, heated for different times at 65°, etc.[1]), regarding each type of denatured product as due to the opening of some definite subset of cyclol bonds and possibly the re-formation of some of them. A denatured product (particularly when, as in the case under discussion, the denaturation is reversible) may have only a small or even negligible decrease in its complement of cyclol bonds. In the case of a C_2 cyclol cage, such as Pauling and Niemann are assuming for trypsin, a net decrease of one in n cyclol bonds, where n is 12 or even 24, is sufficient to allow a gross change in structure and so may be sufficient to account for denaturation. A degradation product of trypsin may therefore possess $^{11}/_{12}$ or even $^{23}/_{24}$ of its full complement of cyclol bonds. Thus if it were the case that a cyclol cage structure is about 28 kcal./mole of amino acid residues less stable than the corresponding linear polypeptides, it still would not follow that the heat of denaturation should be of this order. With reversible denaturation the figure should in this event be $28/n$, or say 2.3 or 1.17 if n were 12 or 24, respectively.

(1) But there is, in any case, no reason to suppose that the figure 28 kcal./mole of residues represents the situation even approximately.

We consider first the attempt of Pauling and Niemann to estimate the difference between the heats of formation of a glycyl linear residue and a glycyl cyclol residue by the explicit use of bond energies. Repeating without acknowledgment the calculation (yielding the value 27.3) which I published in 1936 to ventilate the question,[15] Pauling and Niemann write for the heats of formation in the gaseous state

$$\begin{aligned}&\text{linear glycyl residue} - \text{cyclol glycyl residue}\\ &= \text{C=O} + \text{N—H} + (\text{resonance in CO—NH bond}) - (\text{C—O} + \text{O—H} + \text{C—N})\\ &= 152 + 83.3 + 21 - (110.2 + 70 + 48.6) = 27.5 \quad (1)\end{aligned}$$

or say 28, since the standard energy of NH is usually taken to be 83.7.[16] It is my opinion that

(12) Anson and Mirsky, *J. Gen. Physiol.*, **17**, 393 (1934).

(13) Mirsky and Pauling, *Proc. Nat. Acad. Sci. U. S.*, **22** 439 (1936).

(14) Wrinch, *Phil. Mag.*, **25**, 706 (1938).

(15) Wrinch, *Nature*, **138**, 241 (1936). This calculation was brought to the notice of Professor Pauling in January, 1938.

(16) Pauling, "Nature of the Chemical Bond," Cornell University Press, Ithaca. N. Y., p. 53.

no weight is to be attached to this formal evaluation, since little or nothing is known as to the values of several of the bond energies in these particular cases. Particularly I question the validity of (1) assuming that the C–CH_2 bonds have the same energy in both structures, (2) assuming the C–O bonds in cyclols have an energy as small as in the primary alcohols. These uncertainties appear to me to make the calculation valueless, except in so far as it calls attention (as was my reason for publishing my calculation in 1936) to the types of data which are needed, namely, bond distances (and derivatively bond energies) in long linear peptides—particularly those containing residues other than glycyls—and in ring compounds in which a carbon bearing an OH group is linked to two nitrogens in the same ring, and to another carbon bearing a variety of substituents.

However, as Pauling and Niemann rest the greater part of their case against the cyclol hypothesis upon this calculation and upon two others, it is necessary (A) to point out that these two other calculations are strictly equivalent to the first and (B) to consider how far present data offer any guide as to the bond energies involved in the first calculations. (A) Pauling and Niemann state that they have obtained without the use of bond energies other values (in fact 32 and 24) agreeing closely with and so further supporting the value 28 found with the use of bond energies. First, we have to point out that, assuming that the heat of combustion of crystalline diketopiperazine is 474.6, it is erroneous to conclude that the heat of formation of crystalline diketopiperazine from elements in their standard states is 128. The correct value is

$$-474.6 + 4 \times 94.45 + 3 \times 68.37 = 108.3$$

These two calculations, as they stand, are certainly remarkable in that they give values which differ from the first by +4 and −4, respectively, or using the correct value for the heat of formation of crystalline diketopiperazine, by −6 and −14, respectively. The reason for these discrepancies is easily seen. For if we write for the heats of combustion of crystalline compounds for which experimental data are available a typical equation such as

$$C(CH_2OH)_4 = 4C_2H_5OH - 3CH_4 \qquad (2)$$

say A = 4B − 3C, with the values used by Pauling and Niemann to assess the heat of formation of cyclol residues, we obtain 1302.8 = 1293.0. Again the equation

$$C_6H_{12}N_4 + 12C_2H_6 = 4N(C_2H_5)_3 + 6CH_4$$

say D + 12 E = 4F + 6C, using again the values used for evaluating the glycyl cyclol residue, we obtain the result 5446.7 = 5405.6. Plainly such equations are untrue in the case of these compounds. As little reliance can be placed upon them for the evaluation of cyclol residues. Evidently the only way of interpreting such an equation, P = Q + R, is to write it in the form

Heat of combustion of P_c + K(P_c) = Heat of combustion of Q_c + Heat of combustion of R_c + K(Q_c) + K(R_c)

where K(X_c) is the heat of sublimation of X plus the amount by which the heat of formation of X_{gas} exceeds the sum of bond energies, when the standard values (*e. g.*, N − C = 48.6) are used. When written in this form, all such equations of course yield the result 0 = 0. In other words, such an equation as (2) states simply that, if equal energies be given to the corresponding bonds, then

$$4C{-}C + 8C{-}H + 4C{-}O + 4O{-}H = 4(C{-}C + 5C{-}H + C{-}O + O{-}H) - 12C{-}H$$

After the same manner, the equations for the cyclol residue constructed by Pauling and Niemann, which may be written

linear glycyl residue − cyclol glycyl residue
= ½ diketopiperazine − (F + B − 3E)
= ½ diketopiperazine − (A + D − 3C)/4

give simply my original calculation. In view of the claim that such equations give estimates without the use of bond energies, it must be emphasized that they are simply bond energy equations. They make no use of the heats of combustion cited and add nothing to my original calculation,[15] to which in fact they are strictly equivalent. (B) The fact that a C–C bond length as small as 1.47 ± 0.03 Å. has been found in the only cyclic peptide so far investigated,[17] whereas in glycine[18] the distance is 1.52 ± 0.02 Å. shows that (1) is not the most plausible assumption. Rather it would seem natural to assume that this bond is shorter and therefore stronger in the cyclol structure. Further, it seems difficult to reconcile the assumption of Pauling and Niemann that the tertiary cyclol OH has its C–O bond as low in energy as the standard value for primary alcohols, with Pauling's statement[19] that this bond in secondary

(17) Corey, THIS JOURNAL, **60**, 1598 (1938).
(18) Albrecht and Corey, *ibid.*, **61**, 1087 (1939).
(19) Pauling, *ibid.*, **54**, 3570 (1932).

and tertiary alcohols "seems to be 0.3 to 0.5 v. e. more stable than corresponds to this C–O value." Rather it would seem more reasonable to follow Pauling's 1932 lead in attributing say 0.4 v. e. or 9.2 kcal. more energy to this bond, as is in fact suggested by the observed heat of combustion of trimethylcarbinol.[20] These corrections can easily add 12 to 18 to the heat of formation of the glycyl cyclol residue, reducing the balance in favor of the linear peptide in the gaseous state to (say) 13. So far everything said has had reference only to glycyl residues and only to the gaseous state. Now proteins in general have at least one CH_2 group per residue in their R groups. If any serious estimate of the relative stability of the cyclol cage and the extended peptide is to be attempted, this fact must be taken into account. In view of the solubilities of the native proteins it may be presumed that the interarrangement in space of this large number of CH_2 groups is such as to contribute considerably more, say 2, to the heat of formation of the former than to that of the latter, and to facilitate a higher degree of hydration in the former than in the latter. This higher degree of hydration, if it amounts to two more hydrogen bonded water molecules per residue in the cyclol than in the peptide, would add a further 10 or more to the balance in favor of cyclol cages. These two terms thus may further reduce the balance in favor of linear peptides to $13-2-10=1$. Finally, the association of oppositely ionized R groups has to be taken into account. The dipole moments of the native proteins which have been measured are all low (180×10^{-18} for egg albumin, a molecule of diameter 30 Å. or more and 500×10^{-18} for hemoglobin, a molecule of about twice the weight).[21] While it must again be emphasized that the data necessary for an estimate are lacking (since nothing is known as to the arrangement of the ionizable R groups on any native protein) it appears to be illegitimate and certainly unplausible to assume that the coulomb energy terms would be unchanged when the cyclol cage structure opens into linear peptides. The fact that the stability of a native protein is impaired when the *p*H goes outside some central "stability" range suggests that these terms make some considerably greater contribution to the heat of formation of the cage cyclol than to that of the linear peptide. In egg albumin, for example, where there are say 27 basic and 27 acidic R groups, an arrangement of these groups in isolated pairs at distances apart of 4 Å. would contribute to the heat of formation of the cyclol cage an amount of the order of 2 or more per residue. This changes the balance in favor of the cyclol to 1 or more.

(20) Kharasch, *Bur. Standards J. Research*, **2**, 359 (1929).
(21) Cohn, *Chem. Rev.*, **24**, 210 (1939).

To Sum Up.—The estimate of a balance in favor of the linear peptide of 28 kcal./mole of residues rests upon many assumptions which lack plausibility. With assumptions which are less unplausible, the estimate can be changed to a balance in favor of cyclol cages of at least 1. However the lack of information makes any such estimates, ranging from 28 to −1 or less, of little usefulness. Since in any case, a reversible denaturation probably destroys only a small part of the fine structure, the heat of denaturation should tally not with the balance in favor of cyclols per residue but with some small fraction of this amount.

It must be concluded that no case has been made out for deducing that the cyclol theory is false on the basis of Anson and Mirsky's figure for the heat of denaturation of trypsin; further, that no measurements of the heat of denaturation of proteins can provide a test of the cyclol hypothesis unless the structure of the denatured products in question is known and the heat of formation of such structures and of cyclol cages can be estimated.

Summary

Arguments against the cyclol hypothesis, which have been collected by Pauling and Niemann in a recent article, are examined. It is found that they do not disprove it. In particular their statements purporting to prove that a protein with the cyclol structure would be less stable than the polypeptide chain structure by a very large amount is examined and found to be unproven.

BALTIMORE, MD. RECEIVED APRIL 30, 1940