Summary

Investigations on pyridine derivatives have lead us to the study of the bromination of pyridine in the gaseous phase. At a reaction temperature of about 300°C the hydrogen atoms are substituted by bromine mainly at the 3: or 3:5-positions, while at a reaction temperature of 500°C the hydrogen atoms in 2: or 2:6-positions are substituted by bromine atoms.

Many dibromopyridines and tribomopyridines have become available by these bromination methods; some of these compounds are useful starting materials for synthetical work.

By the study of the high temperature bromination of monohalogenobenzenes or of naphthalene results of theoretical importance have been obtained.

The monobromination of bromobenzene in the gaseous phase, using *a pumice-ferric-bromide catalyst*, belongs to the ortho-para substitution type in the temperature range of 200–450°C; the ratio of the isomeric dibromobenzenes is determined by the differences in energy of activation required for substitution in the ortho-meta- or para-position in the bromobenzene molecule.

When the vapour phase bromination of bromobenzene or of chlorobenzene is carried out in a tube filled with pumice or graphite the substitution reaction belongs to the *ortho-para-type below 410°C*; from 410–600°C meta-substitution predominates.

In the *non-catalytic bromination of liquid naphthalene* in the temperature range of 85–215°C α-bromonaphthalene, besides small amounts of β-bromonaphthalene are formed. The ratio in which the two isomeric monobromonaphthalenes are formed is determined by the difference in energy of activation required for substitution in the α- or β-position in the naphthalene molecule. The same conclusion is valid for the non-catalytic bromination of gaseous naphthalene in the temperature range 250–300°C.

The value of the difference in activation energies for the bromination in the α- and β-positions of the naphthalene molecule is in accordance with the result of quantum-mechanical calculations.

If the non-catalytic bromination of naphthalene vapour is carried out in the temperature range 500–650°C equal amounts of α- und β-bromonaphthalene are formed. At these high temperatures the α:β-ratio is not determined by the difference in energy of activation for α- and β-substitution, but is exclusively dependent on the probability of collision of the bromine molecule (or the bromine atom) and an α- or β-position of the naphthalene molecule.

Theoretical considerations lead to a formula which accounts for the influence of the temperature on the ratio of α- and β-bromonaphthalenes formed in the whole temperature range of 215–650°C.

Quite different results have been obtained in the *catalysed bromination of naphthalene, using ferric-bromide or ferric-chloride* as a catalyst. In the catalytical bromination of liquid naphthalene at 150°C as much as 60% of β-bromonaphthalene may be formed. The bromination of naphthalene is reversible under the catalytic influence of ferric-chloride.

As a result of the reversible reaction: $C_{10}H_8 + Br_2 \rightleftarrows HBr + C_{10}H_7Br$ (α or β) an equilibrium α-bromonaphthalene $\rightleftarrows$ β-bromonaphthalene is established. At 150°C the equilibrium mixture contains 62·3% β-bromonaphthalene. The equilibrium values vary only slightly with the temperature.

In the catalytic bromination of naphthalene vapour (catalyst ferric-bromide on pumice) considerable amounts of β-bromonaphthalene are formed in the temperature range 250–400°C. Here again the formation of β-bromonaphthalene is the result of the reversible bromination-debromination of naphthalene.

The Internal Structure of Protein Molecules[1]

By FELIX HAUROWITZ[2], Bloomington, Ind.

I. *Constellation and internal structure of globular protein molecules*

The protein molecule consists, doubtlessly, of one or of several chains of amino acids linked to each other by peptide bonds. Measurements of the osmotic pressure, the diffusion rate and of the sedimentation equilibrium indicate molecular weights of 50,000 to 100,000 for serum proteins and for similar soluble proteins. If a mean value of 115 is assumed for each of the amino acid residues, a protein of the molecular weight of 100,000 must contain about 870 amino acid molecules.

According to x-ray analyses the length of each of the amino acid residues in the peptide chain is equal to 3·67 Å[1]. Hence, the length of a straight peptide chain of 870 amino acids corresponds to 3,200 Å. Solutions of thread-like molecules of such a length are highly viscous or gelatinous and show a strong flow-birefringence[2]. Since neither the serum proteins nor ovalbumin or hemoglobin have these properties, one has to conclude that their molecules are not threadlike, but that they have a sphere-like or ellipsoidal shape[2]. The classification of proteins according to which they consist either of sphere-like or of more or less thread-

[1] Presented in a colloquium of the Department of Chemistry of the Eidgenössische Technische Hochschule in Zürich (June 7, 1948).

[2] Present address: Dept. of Chemistry, Indiana University, Bloomington, Ind., U.S.A.

[1] R. COREY, J. Amer. Chem. Soc. *60*, 1598 (1938).

[2] G. BOEHM and R. SIGNER, Helv. chim. acta *14*, 1370 (1931).

like molecules was confirmed and developed[1]. The proteins with sphere-like or ellipsoidal shape have been designated as sphero-proteins[2], or as corpuscular proteins.

The diameter of sphero-protein molecules can be determined by viscosimetry, by the measurement of light-scattering or by electron-microscopy. Diameters from 50 to 150 Å, corresponding to a straight chain of 14 to 41 amino acids have been found. Since the protein molecule contains several hundred amino acids one must conclude that the peptide chains, which form the protein molecule, are either ramified or closely folded.

Contrary to many synthetic plastics, proteins are not built up of units linked to each other by three-dimensional covalent bonds. For, the molecules of globular proteins are spread almost instantaneously on the surface of water. They form there films whose height of 7 Å corresponds to the transverse diameter of a *single* peptide chain[3]. From the facility and the speed of spreading it is evident that the globular protein molecule cannot consist of three-dimensional covalent lattices and that their peptide chains cannot be wound up like the threads of a coil, because uncoiling would require a certain time.

The folding of the peptide chains in the original protein molecule and their unfolding, when a surface film is formed, are due to the ability of the peptide chains to rotate about the axes of the C–C and the C–N bonds. It is well known that also other thread-like macromolecules, such as those of rubber, fold or curl in their solutions so that their "constellation" undergoes a permanent alteration[4]. Contrary to these macromolecules *the constellation of protein molecules is specific and remains unchanged* even if the protein is dissolved, salted out, dialysed, cautiously dried and again dissolved. The maintenance of the original specific constellation is proved by the unchanged serological specificity of proteins treated in this way.

The serological specificity of a protein is demonstrated by its property to act as an antigen when injected parenterally into animals of an other species and to induce the formation of antibodies which react specifically with the injected protein. The specificity of the reaction between antigen and antibody is due to a structural adaptation of the surface of the antibody molecule to that of the antigen molecule[5]. This adaptation is highly specific so that even very similar atomic groups such as leucyl-glycyl or glycyl-leucyl residues can be differentiated from each other by means of serological methods[6]. In the same way the species-specific proteins of different animals can be differentiated, even if no differences are revealed by chemical methods.

[1] See by G. BOEHM, Handb. biol. Arbeitsmethoden (E. ABDERHALDEN), Abt. II, Teil 3, p. 3939 ff. (1939).
[2] F. HAUROWITZ, Kolloid-Z. *74*, 208 (1936).
[3] E. GORTER and F. GRENDEL, Biochem. Z. *201*, 391 (1928).
[4] W. KUHN, Z. angew. Chem. *49*, 858 (1936).
[5] F. BREINL and F. HAUROWITZ, Z. physiol. Chem. *192*, 45 (1930). – F. HAUROWITZ, ib. *245*, 23 (1936); Lancet *152*, 149 (1947).
[6] C. LANDSTEINER, *The Specificity of Serological Reactions* (Cambridge, 1946).

It is obvious that the serological specificity of a protein is due to a unique unchangeable arrangement of the peptide chains. One is entitled, in other words, to assume an *unchangeable internal structure* of globular protein molecules. There are two possible causes for the maintenance of the internal structure of the protein molecules: (I) Chain-branching of the peptide chains by covalent bonds and (II) mutual association of the peptide chains by means of van der Waals forces. The possible importance of these two factors for the maintenance of the internal structure of proteins, the alteration of this structure during the denaturation of proteins and the biological significance of the internal structure will be discussed in the following sections.

II. *Chain-branching and formation of rings in protein molecules*

There is no doubt that the bulk of the amino acids is bound in the protein molecule according to formula I.

$$\ldots.NH\cdot CHR_1\cdot CO{-}NH\cdot CHR_2\cdot CO{-}NH\cdot CHR_3\cdot CO\ldots. \qquad \text{I}$$

If R in formula I is an apolar group such as in alanine, valine, leucine, proline or phenylalanine, no other linkage of amino acids than that shown in formula I is possible. If, however, the residue R contains reactive groups such as –COOH, $-NH_2$, –SH or –OH, they could react with each other, forming peptide, ester or thioester bonds[1]. Such bonds would cause a ramification of the main peptide chain (Fig. 1) or the formation of closed rings or networks (Fig. 2)[2]. During the last years such a ramification and formation of rings has been found in glycogen and other polysaccharides[3]. It is necessary, therefore, to examine the question, whether the same has to be assumed for proteins.

Fig. 1.

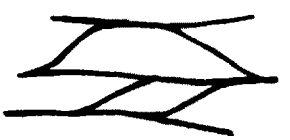


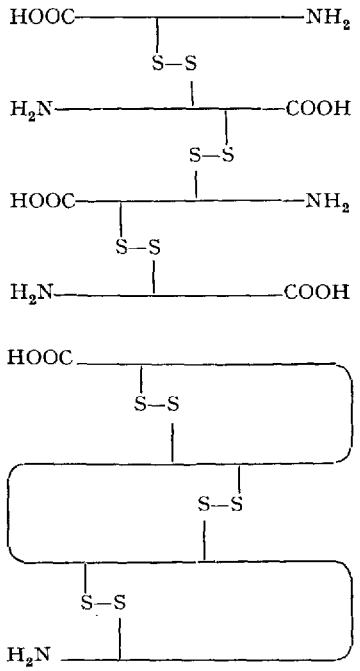
Fig. 2.

(a) *Chain-branching by cystine molecules.* As cystine is a diamino-dicarboxylic acid, it is evident that two parallel peptide chains are interlinked by the dithio group of the cystine molecule (Formula II).

$$\begin{array}{l}\ldots.NH\cdot CHR\cdot CO{-}NH\cdot CH\cdot CO{-}NH\cdot CHR\cdot CO\ldots. \\ \qquad\qquad\qquad\quad \dot{C}H_2 \\ \qquad\qquad\qquad\quad \dot{S} \\ \qquad\qquad\qquad\quad \dot{S} \\ \qquad\qquad\qquad\quad \dot{C}H_2 \\ \ldots.CO\cdot CHR\cdot NH{-}CO\cdot CH\cdot NH{-}CO\cdot CHR\cdot NH\ldots.\end{array} \qquad \text{II}$$

[1] A. CHIBNALL, Proc. Roy. Soc., Ser. B *131*, 136 (1941).
[2] H. MOSIMANN and R. SIGNER in: *The Svedberg 1884–1944* (Uppsala/Stockholm, 1945), p. 464.
[3] W. HAWORTH, E. HIRST, and D. ISHERWOOD, J. Chem. Soc. *1937*, 577. – K. H. MEYER, Naturwissenschaften *29*, 287 (1941).

Many proteins contain more than one cystine molecule so that structures as those shown in formulæ III and IV are to be expected.

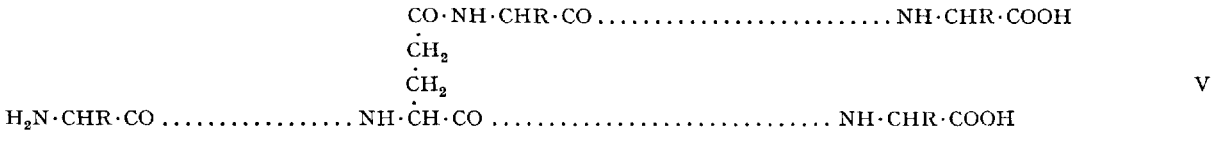


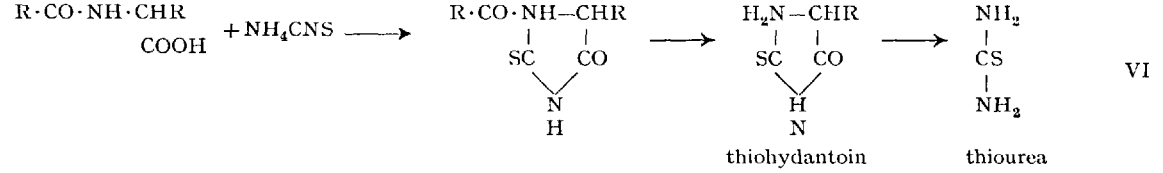
The dithio bond S—S can be split by oxidation with performic acid[1] or by reduction with thioglycollic acid[2] without any apparent diminuition of the molecular weight[2]. The viscosity of solutions of ovalbumin or of the serum proteins treated with performic acid is not different from that of proteins treated with formic acid; likewise the treatment with thioglycolate can be replaced by a treatment with acetate without any effect on the viscosity of the protein solutions[3]. One has to conclude, therefore, that the internal structure of these proteins is maintained predominantly by other atomic groups than the dithio groups of the cystine molecules and that the formulæ III and IV are not quite adaquate.

(b) *Chain-branching by amino-dicarboxylic acids.* This kind of chain-branching is suggested by the formula of glutathion, in which the γ-carboxylic group forms a peptide bond with the terminal amino group of the dipeptide cysteyl-glycine. An analogous linkage of aspartic or glutamic acid in the protein molecule would give rise to chain-branching as shown in formula V.

In spite of chain-branching the number of free COOH groups remains unchanged in such a case, because the new side chain has a terminal COOH group (see formula V). But this group is an α-carboxy group and is separated from the next NH group only by one C atom, while the terminal β- and γ-carboxy groups of aspartic and glutamic acid are separated from the NH group by 2 and 3 C atoms respectively. In order to differentiate the terminal α-carboxy groups from β- and γ-carboxy groups we have condensed proteins with ammonium thiocyanate in the presence of acetic anhydride. It is evident from formula VI that only α-carboxy groups are able to give rise to the formation of a thiohydantoin ring, because the α-imido group is indispensable for the closure of the ring (formula VI).

If the condensation products of proteins with thiocyanates are treated with N—NaOH, thiourea is formed, which can be determined by its color reaction with aquoferricyanide[1] or by its conversion into urea[2]. The amount of thiourea obtained from casein, ovalbumin, zein, beef serum albumin and beef globin, after their treatment with thiocyanate, did not exceed one molecule of thiourea per protein molecule[3]. Since the yield of thiourea from thiohydantoins is 25 to 30%, more than one terminal α-carboxy group might be present in each of the examined protein molecules. But it is evident that their number cannot be very large,





We have attempted to prove the presence of γ-glutamyl bonds by the partial hydrolysis of different proteins with trypsin and the oxidation of the digestion product. Since α-peptide bonds are split by trypsin at

[1] F. SANGER, Nature *160*, 295 (1947).
[2] M. ANSON, J. Gen. Physiol. *25*, 355 (1942).
[3] F. HAUROWITZ, F. BURSA, and A. TÜMER, unpublished experiments.

[1] I. W. GROTE, J. Biol. Chem. *93*, 25 (1931).
[2] R. KITAMURA, J. pharm. Soc. Jap. *55*, 72 (1935); quot. Chem. Zbl. 1935, II, 1347.
[3] F. HAUROWITZ, F. BURSA, and S. LISIE, 17. Int. Congr. Physiol. 79 (1947) and unpublished experiments.

$$\begin{array}{l} \qquad\qquad\qquad\qquad\qquad\qquad\quad CH_2O\cdot CO\cdot CHR\cdot NH \ldots\ldots\ldots\ldots\ldots\ldots CO\cdot CHR\cdot NH_2 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad | \\ H_2N\cdot CHR\cdot CO \ldots\ldots\ldots\ldots NH\cdot CH\cdot CO \ldots\ldots\ldots\ldots\ldots NH\cdot CHR\cdot COOH \end{array} \qquad \text{VII}$$

a higher rate than the γ-glutamyl bonds of glutathion[1], it was expected that γ-glutamyl peptides would be found after the treatment of proteins with trypsin. The cautious oxidation of γ-glutamyl peptides with H_2O_2, NaOBr or with chloramine T furnishes succinyl peptides which, on total hydrolysis with HCl, give succinic acid:

$$\underset{\gamma\text{-glutamyl peptide}}{R\cdot NH\cdot CO\cdot CH_2\cdot CH_2\cdot CHNH_2\cdot COOH} \longrightarrow \underset{\text{succinyl peptide}}{R\cdot NH\cdot CO\cdot CH_2\cdot CH_2\cdot COOH}$$

While preliminary experiments with this method had furnished small amounts of succinic acid[2], later experiments with an improved method gave much lower yields[3]. Probably only very few glutamic acid molecules give rise to chain-branching of the type shown in formula V. It had been proved, before, that the α-amino groups of glutamic acid and of glutamine are not free in the protein molecule[4]. One has to assume, therefore, that glutamic acid and glutamine cannot be bound according to the formulæ $R\cdot NH\cdot CO\cdot CHNH_2\cdot CH_2\cdot CH_2\cdot CO\cdot NH\cdot R$ or $R\cdot NH\cdot CO\cdot CHNH_2\cdot CH_2\cdot CH_2\cdot CO\cdot NH\cdot CO\cdot R$. Both, glutamic acid as well as glutamine are bound according to formula I by means of the α-amino groups, and most of them also by the α-carboxy group. While the animal proteins contain only a few γ-glutamyl residues, a larger number of these was found in the capsular protein of *Bacillus anthracis*[3,5].

(c) *Chain-branching by diamino, hydroxy-amino and by thioamino acids.* The problem is here similar to that discussed in the preceding section. If the amino groups of lysine, the hydroxy groups of serine, threonine or hydroxyproline or the thio groups of cysteine combine with the terminal α-carboxy group of a peptide chain, then the other end of this peptide chain will bear a free α-amino group and the protein molecule will contain more than one terminal α-amino group. Actually many proteins, when treated with nitrous acid, give more N_2 gas than the amount corresponding to the number of their lysine molecules. The excess of N_2 formed in this way corresponds to 9 amino groups in serum albumin[1], 4 in ovalbumin[2], 4 in lactoglobulin[1], 9 in cytochrome[3] and 17 in insulin[4].

It should be borne in mind, however, that glycyl peptides give more than the calculated amount of N_2[5] and that N_2 can sometimes be formed by the action of nitrous acid on acid amides[6] or on N-free reducing substances[7]. Zein, according to careful analyses, has no excess of free α-amino groups[8].

The terminal α-amino groups of different proteins have been labeled by substituting them with phenylhydantoin[9] or with dinitrofluorobenzene[10]. By hydrolysing the dinitrophenyl-proteins the dinitrophenyl derivatives of the terminal amino acids were obtained. According to SANGER[10] the terminal α-amino groups of the insulin subunit (equiv. weight = 12,000) are formed by 2 molecules of glycine and 2 of phenylalanine, while the 3 terminal α-amino groups of lactoglobulin are those of leucine and the 6 terminal α-amino groups of horse hemoglobin those of valine[10]. It is not yet known whether the multiplicity of terminal α-amino groups is due to the presence of several parallel peptide chains according to formula III or to ramifications as shown by formula VII[2].

Formula VII shows an ester bond formed by the hydroxy group of serine. Analogous bonds could be formed by the other hydroxyamino acids. The presence of these ester bonds has neither been proved nor excluded hitherto.—A similar structure as that shown in formula VII would arise if the ε-amino groups of lysine were involved in the formation of peptide bonds. The existence of such peptide bonds is disproved, however, by the fact that all of the ε-amino groups of lysine are deaminised when proteins are treated with nitrous acid. One has to conclude, therefore, that the ε-amino groups of lysine are free in the protein molecule.

Finally, it should be mentioned that the presence of ureido bonds of the formula—NH·CO·NH—has been

[1] W. GRASSMANN, H. DYCKERHOFF, and H. EIBELER, Z. physiol. Chem. *189*, 112 (1930).

[2] F. HAUROWITZ and M. VARDAR, C. r. Soc. Turq. Sci. Nat. *11*, 33 (1944).

[3] F. HAUROWITZ and F. BURSA, Biochem. J. *44*, 509 (1949).

[4] F. HAUROWITZ and M. TUNCA, Biochem. J. *39*, 443 (1945).

[5] G. IVANOVICS and V. BRUCKNER, Naturwissensch. *25*, 250 (1937). - W. E. HANBY, and N. H. RYDON, Biochem. J. *40*, 297 (1946).

[6] A. CHIBNALL and R. WESTALL, Biochem. J. *26*, 122 (1938).

[7] H. CARTER and S. DICKMANN, J. Biol. Chem. *149*, 571 (1943).

[8] T. LAINE, Ann. Soc. Acad. Fenn. A II, no. 11 (1944).

[9] J. ROCHE, M. RAYMOND, and J. SCHILLER, C. r. Akad. Sci. *219*, 38 (1944).

[10] F. SANGER, Biochem. J. *39*, 507 (1946). – F. SANGER and R. PORTER, Biochem. J. *42*, 287 (1948).

suggested[1]. The existence of ureido bonds is supported by the fact that proteins upon hydrolysis with 20% sulphuric acid furnish several molecules of carbonic acid[2]. We have been able to show, however, that the same amount of CO_2 is obtained when the protein is heated with 0·1-*N* sulphuric acid to lower temperatures such as 80°[3]. Evidently the carbonic acid originates from bicarbonate ions bound to cationic groups of the protein. The existence of ureido groups is disproved by these experiments.

(d) *Formation of rings in protein molecules.* Each of the types of chain-branching mentioned in the preceding sections (a)—(c) can give rise to the formation of rings as soon as more than one point of branching exists in one protein molecule (Fig. 2). Hitherto the presence of such rings, formed by covalent bonds, has been proved only in the toxic peptide gramicidin[4]. Its cyclopeptide ring is formed by 5—10 amino acids.

Diketopiperazine rings are formed if proteins are treated with strong acids[5], while the heterocyclic pyrrole ring arises very easily from glutamine, even at body temperature. It is not astonishing, therefore, that a certain amount of pyrrole derivatives is obtained, if proteins are treated with acetic anhydride or with other dehydrating reagents[6]. It is not necessary to assume that other pyrrole rings than those of proline and oxyproline are present in the original protein molecule. The same is valid for the assumption that trioxy-triazine rings (so-called cyclole rings) form the scaffold of the protein molecule[7]. The large number of OH groups postulated by the cyclol hypothesis has never been found, neither by methylation, nor by acetylation[8] or by exchange with H_2O^{18} [9].

III. *Intramolecular bonding by van der Waals forces*

While the energy of covalent bonds amounts to 50—100 kcal per mole, that of bonds due to van der Waals forces varies from 1 to 10 kcal. The different types of such bonds existing in protein molecules will be discussed in the following sections.

It has been proved by different authors that the molecules of certain proteins such as hemoglobin consist of "sub-units" held together by van der Waals forces and that the cleavage of the molecule into sub-units can be brought about by the action of urea at room temperature or by similar mild procedures. These reports are confusing, to a certain degree, because true molecules, according to the classical definitions, are formed by atoms held together by covalent bonds. It is evident, therefore, that the real molecules are formed, in such a case, by the so-called sub-units and that the alleged protein molecules are micelles or aggregates of several molecules.

(a) *Dipole-dipole and dipole-ion bonds.* Polar atomic groups attract each other mutually and are attracted also by ionic groups. This attraction results, obviously, from the action of Coulomb forces.

According to the fundamental laws of electrostatics, poles of the same kind of charge repel each other, while those of the opposite kind attract each other. While the attractive power increases when the particles are approaching each other, the repulsion, obviously, decreases as soon as the particles are removed from each other. Accordingly the force of attraction will always surpass that of repulsion and a *mutual approach of the polar groups will always result.*

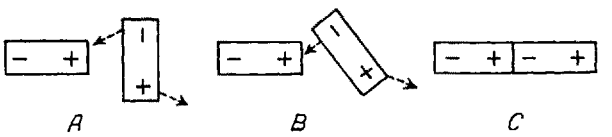
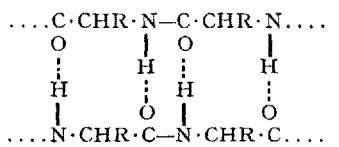


Fig. 3. – Association of dipoles.

The strongest electrostatic forces are exerted by the ionic groups COO^- and NH_3^+. But also the groups OH, NH_2, SH, COOH, the heterocyclic N-atoms and the peptide bonds have to be considered as dipoles. The polarity of these groups is due to the fact that the electron pair, which links the H atom to N, O or S, is not situated in the center of the bond, but is nearer to the N, O or S than to the H atom. Thereby the H atom becomes the positive pole, N, O or S the negative pole of the resulting dipole. Owing to its positive character the H atom is attracted by the negative poles of adjacent dipoles, mainly by O or N atoms, so that "hydrogen bonds" are formed[1]. Since peptide linkages form the bulk of the polar groups of the protein molecule, hydrogen bonds between these groups (formula VIII) have a special importance[2].



Owing to their mutual attraction the polar atomic groups of the peptide chains act as a cement, so that the peptide chains of dry proteins are very closely attached to each other and a high density of 1·30 to 1·35 results[3]. If dry proteins are dissolved in water the structural arrangement of their peptide chains remains essentially intact, as shown by the unchanged

[1] R. SYNGE, Bull. Soc. Chim. Biol. *27*, 286 (1945).
[2] M. DUNN, J. Amer. Chem. Soc. *47*, 2564 (1925).
[3] F. HAUROWITZ and M. BILEN, unpublished experiments.
[4] R. CONSDEN, A. GORDON, A. MARTIN, and R. SYNGE, Biochem. J. *41*, 596 (1947).
[5] E. ABDERHALDEN, Z. physiol. Chem. *265*, 23 (1940).
[6] N. TROENSEGAARD, *The Structure of Protein Molecules* (Kopenhague, 1944); Acta chem. Scand. *1*, 672 (1947).
[7] D. WRINCH, Proc. Roy. Soc., Ser. A *160*, 59 (1937); Nature *145*, 669 (1940).
[8] F. HAUROWITZ, Z. physiol. Chem. *256*, 28 (1938).
[9] W. MEARS and H. SOBOTKA, J. Amer. Chem. Soc. *61*, 880 (1939).

[1] W. LATIMER and W. RODEBUSH, J. Amer. Chem. Soc. *42*, 1419 (1920).
[2] A. MIRSKY and L. PAULING, Proc. Nat. Acad. Sci. *22*, 439 (1936).
[3] E. COHN and J. EDSALL, *Proteins, Amino Acids and Peptides* (New York, 1943) p. 375.

serological specificity, the unaltered shape of the protein crystals and by x-ray analysis. By means of the latter method it has been demonstrated that *the interior of globular protein molecules is very poor in water* and that the hydration of proteins is mainly due to layers of water molecules on the surface of the globular molecules (PERUTZ a. o.[1]). Apparently the hydrogen bonds between adjacent peptide chains >N–H–OC< are stronger than those between peptide linkages on the one hand and water molecules on the other, so that water molecules cannot penetrate into the interior of the globular protein molecule. The strong mutual attraction of parallel peptide chains is either due to the multiplicity of their hydrogen bonds or to the fact that also salt bridges are involved in the reciprocal linkage of these chains.

That a multiplicity of even very weak bonds may give rise to very stable structures is evident from a comparison with two magnetic bodies held together by magnetic forces. While such bodies stick to each other very loosely, if there are few points of contact, enormous forces are required to separate them, if large parts of their surfaces touch each other.

(b) *Ion-ion bonds (salt bridges)*. These are bonds between positively charged NH_3 groups of diamino acids and negatively charged COO^- groups of amino-dicarboxylic acids. Ion-ion bonds are much more stable than bonds of the type dipole-dipole.

The power of attraction operating between two ionic groups over the distance r is proportional to r^{-2}, whereas the power of attraction of two non-ionic dipoles is proportional to r^{-6}. Accordingly the power of attraction decreases very rapidly in the latter case when r increases, and the range of attraction is limited to closely adjacent dipoles. The bonding energy of salt bridges amounts to 5–10 kcal[2], a figure which is not much higher than that assumed for the attraction of non-ionic dipoles. But the power of attraction between ionic groups of an opposite charge does not decrease so rapidly with an increase of r. Therefore, salt bridges split by thermal movements will be restored again.— The force operating between an ionic and a non-ionic polar group is intermediate between that of the ion-ion and the dipole-dipole type. Thus, it has been demonstrated that the bonding energy of water molecules adsorbed to an ionic surface decreases from 6·55 kcal for the first layer of water molecules to 1·38 and 0·42 kcal for the second and third layer respectively[3].

Salt-bridges are split easily by the addition of strong acids or bases. The negatively charged COO^- groups are transformed into uncharged COOH groups in the presence of strong acids. Similarly NH_3 groups loose their charge in the presence of alkalis and are transformed into uncharged NH_2 groups. As soon as the protein looses its positive or its negative charge its peptide chains will contain only negatively or only positively charged atomic groups. The mutual repulsion operating between charges of the same sign will then lead to the unfolding of the peptide chains and to an increase of the volume occupied by the dissolved molecule. This is manifested by the increased viscosity of acid or alkaline protein solutions as compared with the viscosity of neutral solutions[1]. Since no increase in viscosity is observed, when neutral solutions of proteins are stored for several hours or days, one has to conclude that the salt bridges are resistent to the thermal movements of water molecules and of peptide chains at room temperature.

(c) *Bonding by non-polar groups*. The attractive forces which are effective between non-polar atomic groups such as the aliphatic side chains of valine, leucine or phenylalanine, are much weaker than those between polar or ionic groups. This is evident from the low melting point of paraffins and similar substances, whose molecules in the solid state are linked to each other by forces between non-polar groups. It is not likely that these weak forces have any importance for the maintenance of the internal structure of protein molecules. They are important, however, for the formation of protein-lipide complexes.

IV. *Denaturation*

If proteins are treated with strong acids or bases, with organic solvents with heat or with other chemical and physical agents, their *internal structure is changed*. This phenomenon is called denaturation. Denaturation cannot be attributed to the cleavage or to the closure of peptide bonds, because the volume of the protein solution remains unchanged during denaturation[2]; the hydrolytic cleavage of peptide bonds is accompanied by the consumption of one water molecule per peptide bond, so that the volume of the solution would diminish considerably if hydrolysis occured. For the same reason the formation of peptide bonds would be accompanied by an increase of the volume of the solution. The view that denaturation is not due to the cleavage of covalent bonds is also supported by the fact that denaturation is brought about by very mild methods such as the action of urea at room temperature.

The alteration of the internal structure of a protein molecule during its denaturation is demonstrated, schematically, by Fig. 4 *A* part of the peptide chains of the native molecule is shown by 4 A. Upon the addition of NaOH the positively charged NH_3^+ groups loose their charge according to the equation: $-NH_3^+ + \bar{O}H \rightarrow NH_2 + H_2O$, so that the salt bridges are opened and the peptide chains unfold (Fig. 4 B). The degree of unfolding depends on the time of the action of NaOH. If the displacement of the peptide chains is

[1] J. BOYES-WATSON, E. DAVIDSON, and M. PERUTZ, Proc. Roy. Soc. Ser. A *191*, 83 (1946).
[2] M. DAVIES, Ann. Reports Chem. Soc. *43*, 6 (1947).
[3] W. HARKINS, Science *102*, 294 (1945).

[1] G. ETTISCH and G. SCHULZE, Biochem. Z. *239*, 48 (1931).
[2] F. HAUROWITZ, Kolloid-Z. *74*, 208 (1936).

very large, the alteration of the internal structure becomes irreversible. Neutralization of the alcaline solution by acids leads again to the formation of positively charged groups (Fig. 4 C), so that intramolecular salt bridges are formed again. But a part of the displaced ionic groups remains unsaturated and is available for the formation of *salt-bridges between adjacent protein molecules*. In this way large aggregates of protein molecules are formed and coagulation of these aggregates occurs, as soon as their size exceeds a certain limit[1].

In an analogous manner the negatively charged groups lose their charge, if an acid is used as denaturing agent. Denaturation by heat-treatment is due to a scission of the salt-bridges by thermal movements of the atomic groups. In all these cases the essential

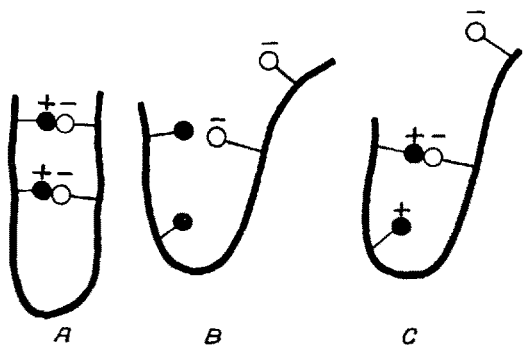


Fig. 4. – Peptide chain of a protein. *A* native, *B* in NaOH solution *C* neutralized (denatured).

process is the unfolding of the peptide chains and the simultaneous penetration of water molecules between these chains. Dry proteins cannot unfold in this manner. Therefore dry proteins are much more resistent to heat than dissolved proteins.

The assumption that denaturation consists, essentially, in an unfolding of closely folded peptide chains, has been advanced for the first time by WU[2]. This assumption is supported by the *higher reactivity of denatured proteins*. They give more intensive color reactions with nitroprussides[3], with Folin's phenol reagent[4] and with diazobenzene sulfonic acid[5] than the same proteins in the native state. The low reactivity of native proteins is due to the inaccessibility of certain atomic groups inside the closely folded peptide chains. This steric protection of reactive atomic groups is very important from the biological view-point. It is responsible for the *resistence of native proteins to the action of proteolytic enzymes*. Denatured globular proteins are hydrolysed by trypsin or by papain at a higher rate than the same proteins in the native state[1]. The well-known resistence of living tissues against digestion is due to this fact. It is not necessary to postulate a specific vital resistence or to assume the presence of specific anti-enzymes.

If the action of the denaturing agent does not last a long time the first phases of denaturation may be reversible[2]. But the renaturation is not complete. By means of sensitive methods differences between native and apparently renatured proteins have been revealed[3].

V. *Biological importance of the internal structure of proteins*

The denaturation of native proteins is accompanied not only by a loss of the original internal structure, but also by a loss of the serological specificity and, if the protein has an enzymic or a hormonal action, by a loss of this biological activity. It is obvious, therefore, that the maintenance of the specific internal structure is highly important from the biological point of view.

The mechanism by which the specific internal structure of native proteins is built up, is not yet clear. We know, however, that this mechanism is altered by the injection of heterogenous antigens, so that antibodies are formed instead of the normal proteins. These antibodies are protein molecules whose surfaces are structurally adapted to the surface of the foreign antigen molecules[4]. One can reasonably assume, therefore, that the specificity of the normal proteins is due to the adaptation of their molecules to templates present in the normal cells[5]. The biosynthesis of proteins occurs, most probably, in two steps. In the first of them a positive replica of an expanded protein template film is formed; in a second step the two-dimensional replica folds to form a three-dimensional globular particle, which is complementarily adapted to an adjacent cellular polar molecule or to a foreign antigen molecule[6]. While in the latter case a true antibody will be formed, adaptation to a normal cellular constituent furnishes "auto-antibodies". (TYLER[7]).

It is evident from the preceding sections that the specificity of the internal structure of a protein molecule is dependent on two different factors: (I) the number of the different types of amino acids and their order in the peptide chain, (II) the kind of folding of the peptide chains. It is obvious, moreover,

[1] F. HAUROWITZ, Kolloid-Z. *77*, 65 (1936).
[2] H. WU, Chin. J. Physiol. *3*, 1 (1929); *5*, 321 (1931).
[3] A. MIRSKY and M. ANSON, J. Gen. Physiol. *19*, 427, 439 (1936).
[4] M. ANSON, J. Gen. Physiol. *24*, 399 (1941).
[5] F. HAUROWITZ and S. TEKMAN, Acta biochem. *1*, 484 (1947).

[1] F. HAUROWITZ, M. TUNCA, P. SCHWERIN, and V. GÖKSU, J. Biol. Chem. *157*, 621 (1945).
[2] M. ANSON and A. MIRSKY, J. Gen. Physiol. *14*, 597, 725 (1931).
[3] H. NEURATH, C. COOPER, and J. ERICKSON, J. Biol. Chem. *142*, 249 (1942). – F. BERNHEIM, H. NEURATH, and J. ERICKSON, J. Biol. Chem. *144*, 259 (1942). – F. HAUROWITZ and H. WAELSCH, Z. physiol. Chem. *182*, 82 (1929).
[4] F. BREINL and F. HAUROWITZ, Z. physiol. Chem. *192*, 45 (1930). – F. HAUROWITZ, ib. *245*, 23 (1936); Lancet *152*, 149 (1947).
[5] F. BREINL and F. HAUROWITZ, Z. physiol. Chem. *192*, 45 (1930). F. HAUROWITZ, ib. *245*, 23 (1936); Lancet *152*, 149 (1947).
[6] F. HAUROWITZ, Quarterly Rev. Biol. *24*, 93 (1949).
[7] A. TYLER, Physiol. Rev. *28*, 180 (1948).

that the probability of a definite arrangement of the amino acids and of the peptide chains is much smaller than that of randomly arranged amino acids and randomly folded peptide chains. The building up of a definite unique structure instead of a mixture of randomly arranged structures is connected with a loss of entropy. The magnitude of this difference in entropy can be calculated if some simplifying assumptions are made. Thus, for a protein containing 150 molecules of each of the 20 amino acids an entropy difference of 5 cal per degree per amino acid residue results (BUTLER[1]). At 37° (T = 310°) this corresponds to a difference in energy of approximately $5 \times 310 = 1{,}550$ cal per amino acid residue. Since the average equivalent weight of these residues is about 115, the excess in energy required for the formation of a given specific arrangement is equal to $1{,}550/115 = 14$ cal per g of the protein. This is an astonishingly low figure, if account is taken of the fact that the oxidation of one g of protein furnishes approximately 4,100 calories[1]. The conclusion has to be drawn, that *the organism can build up the complicated specific structures of the protein molecules without any considerable increase in the expenditure of energy.*

The same result has been obtained by the experimental determination of the difference in energy between native and denatured proteins. The "heat of denaturation" has been determined as the difference between the heats of reaction of native and of denatured protein with KOH[2]. It amounts to 2·5 cal per g of protein, a figure which is still lower than that obtained by the above calculation.

In the present communication it has been dealt with the globular molecules of spheroproteins. It must be borne in mind, however, that the real scaffold of the cells is formed by insoluble proteins and by viscous protein solutions, which grant to the cell its characteristic semisolid consistency. The molecules of these *structural proteins* are probably elongated thread-like molecules, each of them corresponding to a single long peptide chain. Evidently one cannot speak of an internal structure of the molecule in such a case. Recent investigations on gelatin and on collagen, by means of the electron microscope and by x-rays, have revealed, however, transversal bands, spaced approximately 700 Å from each other[1]. It is suggested by this finding that the fibers of fibrous proteins consist of a chain of globular protein particles, possessing a definite internal structure[2].

It has been shown above that the denaturation of native proteins results in a loss of their resistence to proteolytic enzymes. The same is valid for the intracellular proteins. As soon as they are damaged artificially by an external agent, they are exposed to the action of the intracellular enzymes, so that the cell undergoes autolysis. One is tempted to assume that the ageing and the natural death of living organisms is due to the same phenomenon, to a gradual unfolding of the intracellular spheroproteins and, thereby, a loss of their resistence to the intracellular proteolytic enzymes.

Zusammenfassung

1. Die spezifische Konstellation der globulären Eiweißmoleküle bleibt im Gegensatz zu der Konstellation anderer Fadenmoleküle unverändert, wenn das Protein in Wasser gelöst, ausgesalzen oder vorsichtig getrocknet wird. Die große Stabilität der spezifischen Konstellation wird einer bestimmten inneren Struktur der Proteine zugeschrieben.

2. Die innere Struktur globulärer Proteinmoleküle wird vor allem durch Salzbrücken zwischen positiv und negativ geladenen Atomgruppen aufrechterhalten, ferner durch H-Bindungen, durch Dithiobrücken, vielleicht auch durch geringgradige Verzweigung der Peptidketten.

3. Denaturierung besteht in einer Änderung der spezifischen inneren Struktur eines Proteins, herbeigeführt durch eine Sprengung der Salzbrücken. Denaturierung führt zu einer Entfaltung der im nativen Protein dicht gefalteten Peptidketten. Die Resistenz nativer Proteine und lebender Zellen gegen proteolytische Fermente geht bei Denaturierung verloren, da durch die Entfaltung der Peptidketten Angriffspunkte für die proteolytischen Fermente freigelegt werden.

4. Der Mechanismus der Bildung der spezifischen Struktur nativer Proteine im lebenden Organismus wird erörtert.

[1] J. BUTLER, Nature *158*, 153 (1946).
[2] J. CONN, G. KISTIAKOWSKY, and R. ROBERTS, J. Amer. Chem. Soc. *62*, 1895 (1940); *63*, 2080 (1941).

[1] H. BAER, J. Amer. Chem. Soc. *66*, 1297 (1944). – F. O. SCHMITT, Adv. Protein Chem. *1*, 25 (1944).
[2] J. FARRANT, A. REES, and E. MERCER, Nature *159*, 535 (1947).