[CONTRIBUTION FROM THE DEPARTMENT OF CHEMISTRY, CORNELL UNIVERSITY]

Thermodynamic Considerations of Protein Reactions.[1,2] I. Modified Reactivity of Polar Groups

BY MICHAEL LASKOWSKI, JR.,[3] AND HAROLD A. SCHERAGA

RECEIVED MAY 6, 1954

The thermodynamics of protein reactions are considered in terms of a model containing intramolecular hydrogen bonds between the side chain polar R groups of the amino acid residues of the protein. A quantitative treatment of modified reactivity of polar groups is given for this model involving the use of equilibrium constants for the formation of hydrogen bonds. Approximate values of these constants are obtained by evaluation of the entropy and enthalpy of formation of an internal hydrogen bond in a protein. The entropy of formation is estimated by considering the effect of the hydrogen bond upon the rotational degrees of freedom of the side chain donor and acceptor groups. The enthalpy of formation is estimated by considering the generally accepted values to be modified by effects due to these rotational degrees of freedom. This treatment is applied here to the ionization of polar groups involved in several different kinds of hydrogen bonding situations. It is shown that the ionization constants can be increased or decreased by the presence of the hydrogen bonds to an extent quite different from that normally attributed to electrostatic effects. It is thus possible to explain the abnormal spreading and steepening of some titration curves without introducing such concepts as inaccessibility of reactive groups, or gross unfolding of the protein molecule. Rather, the reduced, or enhanced, reactivity may be attributed to thermodynamic effects arising from the presence of the hydrogen bonds. Another application of this treatment is made to the problem of the binding of small molecules or ions to proteins by considering that the small molecule or ion can compete with, say, a polar R group acceptor for the polar R group donor as a binding site. These competitive effects considerably modify the behavior normally attributed to statistical and electrostatic effects. With this model it is possible to account for some of the abnormal dependence of binding constants on the amount of unbound material, and also for the phenomenon of "all or none" binding, without requiring unfolding of the protein molecule.

Introduction

In protein reactions the behavior of the various reactive groups is usually quite different from that exhibited by these groups in low molecular weight model compounds, giving rise to the notion that this different reactivity is due to intramolecular cooperation between two or more such groups. This modified reactivity of polar R groups is encountered in investigations of such familar problems as those of ion and dye binding by proteins, reactivity of SH groups and ionization. The reduced reactivity in native proteins has been attributed[4] to the unavailability of most of the groups concerned due to the folding of the protein molecule. However, from the work of Crammer and Neuberger,[5] Levy and Benaglia,[6] Tanford,[7] and Klotz,[8] it appears that the assumption of impermeability of a protein molecule to the flow of small molecules and ions is untenable, and that the lack of reactivity can more readily be attributed to hydrogen bonding between neighboring polar side groups of the amino acid residues.[9] A hydrogen bonded model for proteins has been considered by Mirsky and Pauling[10] and also by Eyring and Stearn.[11] The notion was, however, strongly criticized by Jacobsen and Linderstrøm-Lang.[12,13] In light of the considerable evidence[5-8] in favor of the existence of internal hydrogen bonds in proteins, such intramolecular coöperation will be assumed here in a quantitative treatment of the thermodynamics of protein reactions.

Two examples of modified reactivity of polar groups have been chosen for consideration: (1) ionization, and (2) binding of small molecules and ions to proteins. However, the results presented here can readily be generalized to include other types of modified reactivity. Extensions of this treatment to the problems of denaturation and proteolysis will be considered in subsequent papers of this series.[14,15]

For consideration of ionization (which is, in general, instantaneous, reversible, and easy to follow experimentally) use will be made of hydrogen bonding equilibrium constants K_{ij}, K_{lm} and K_{rs}, and intrinsic ionization constants K_1° and K_2°. In addition, for the problem of the binding of small molecules and ions, a binding constant, K_B, will

(1) This work was done in connection with projects supported at Cornell University by the National Institutes of Health (U. S. Public Health Service), and by the Office of Naval Research.

(2) Presented, in part, before the Division of Biological Chemistry at the 124th meeting of the American Chemical Society, Chicago, Illinois, September, 1953, and, in part, at the 126th meeting, New York, N. Y., September, 1954.

(3) U. S. Public Health Service Research Fellow of the National Heart Institute, 1952–1954.

(4) See, for example, F. Haurowitz, "Chemistry and Biology of Proteins," Academic Press, Inc., New York, N. Y., 1950.

(5) J. L. Crammer and A. Neuberger, *Biochem. J.*, **37**, 302 (1943).

(6) M. Levy and A. E. Benaglia, *J. Biol. Chem.*, **186**, 829 (1950).

(7) C. Tanford and G. L. Roberts, THIS JOURNAL, **74**, 2509 (1952); C. Tanford, *ibid.*, **72**, 441 (1950); *Proc. Iowa Acad. Sci.*, **59**, 206 (1952).

(8) I. M. Klotz, *Cold Spring Harbor Symp. Quant. Biol.*, **14**, 97 (1949).

(9) While such bonding has been occasionally referred to as "salt bridging" we shall use the term "hydrogen bonding" to specify such interactions regardless of whether the donor and/or acceptor carry a charge.

(10) A. E. Mirsky and L. Pauling, *Proc. Natl. Acad. of Sci., U. S.*, **22**, 439 (1936).

(11) H. Eyring and A. E. Stearn, *Chem. Revs.*, **24**, 253 (1939).

(12) C. F. Jacobsen and K. Linderstrøm-Lang, *Nature*, **164**, 411 (1949).

(13) The argument of Jacobsen and Linderstrøm-Lang,[12] that the decrease in volume accompanying ionization of carboxyl groups cannot be reconciled with extensive hydrogen bonding, appears to be incomplete. Even if their statement about presumed charge neutralization in a $C\langle^{O}_{O^-}$ ...NH_3^+ bond is correct they have neglected the possible existence of carboxyl...carboxyl and hydroxyl...carboxylate ion bonds. The ionization of a carboxyl group to form a hydroxyl–carboxylate ion hydrogen bond, as illustrated in Fig. 1, *would* lead to a volume decrease. The confusion seems partly due to the use of the term "salt bridge." Thus, even if $C\langle^{O}_{O^-}$ ...NH_3^+ bonds did not occur frequently in proteins, the argument of Jacobsen and Linderstrøm-Lang would not vitiate the hydrogen bonded model.

(14) M. Laskowski, Jr., and H. A. Scheraga, THIS JOURNAL, to be submitted.

(15) From a consideration of the molecular motions of large molecules it has also been possible to account for the thermodynamics of protein associations.

also be required. These constants, together with the net charge on the protein, determine the values of K_{obs} and $(K_B)_{obs}$, where the former is the observed equilibrium constant characterizing the ionization behavior of a *specific* group in a protein for various hydrogen bonding situations, and the latter is the corresponding observed equilibrium constant characterizing the binding properties of proteins for small molecules and ions. Approximate values of the hydrogen bonding equilibrium constants will be obtained by considering the effect of hydrogen bonding upon the freedom of the bonded groups to rotate, or rather to undergo torsional oscillations around single bonds.

The Protein Model[10,11,16]

A native, globular protein is thought of as consisting of a few helically folded[16-19] peptide chains held together by disulfide links and hydrogen bonds between polar groups of the amino acid residues. The folding of the helix is maintained by hydrogen bonds between the peptide NH and C=O groups, while the polar R groups are free to form both inter and intra-chain hydrogen bonds. It is possible that such hydrogen bonds between polar R groups may confer an additional amount of stability on any of the helical configurations which have been proposed. In accordance with the views of Bull[20] and Neurath, *et al.*,[21] that the specificity of the protein is due primarily to interactions between side chains, we shall make the arbitrary assumption that polar R groups can hydrogen bond only to other polar R groups and not to peptide bonds. For the sake of simplicity we have also assumed that the helices are held rigidly with respect to each other and that their relative positions do not change due to breakage or formation of hydrogen bonds between the R groups. It may be noted that in this hydrogen bonded model no consideration is given to non-polar forces such as van der Waals attractions.

Furthermore, it is of importance to decide whether the structure of the native protein is so rigid as to restrict a given donor group to interacting with only one acceptor (and *vice versa*) or whether the specific donor can compete for several acceptors, or coöperate by bonding to an acceptor to which another donor is already bonded. The structural evidence does not at present allow one to decide this point. Since a general solution involving extensive competition or coöperation would be both very cumbersome and difficult to obtain we have included here only several illustrative examples of both phenomena. The treatment of the non-competitive, non-coöperative type of hydrogen bond is, however, a general one.

(16) For a recent discussion of protein structure see the chapter by B. W. Low, pp. 235 ff. in "The Proteins" (Ed. by Neurath and Bailey), Vol. IA, Academic Press, Inc., New York, N. Y., 1953.

(17) L. Pauling, R. B. Corey and H. R. Branson, *Proc. Natl. Acad. of Sci., U. S.*, **37**, 205 (1951), *et seq.*; R. B. Corey and L. Pauling, *Proc. Royal Soc. (London)*, **B141**, 10 (1953); L. Pauling and R. B. Corey, *ibid.*, **B141**, 21 (1953).

(18) M. L. Huggins, THIS JOURNAL, **74**, 3963 (1952).

(19) B. W. Low and R. B. Baybutt, *ibid.*, **74**, 5806 (1952).

(20) H. B. Bull, *Advances in Enzymol.*, **1**, 1 (1941).

(21) H. Neurath, J. P. Greenstein, F. W. Putnam and J. O. Erickson, *Chem. Revs.*, **34**, 157 (1944).

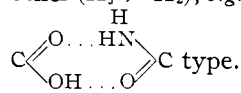
Thermodynamic Formulation

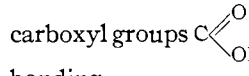
It is of interest to consider four types of hydrogen bond between polar side chain groups, which conceivably could be present in proteins.[22]

I. Single Bonds. (a) **Heterologous.**—This is a bond where the two participating groups are chemically distinct and where it can readily be decided which group will in most cases serve as a donor and which as an acceptor ($K_1 \neq K_2$, $K_{ij} \neq K_{ji}$, see below). The tyrosyl...carboxylate ion bond discovered by Tanford[7] in serum albumin is an example.

(b) **Homologous.**—The two participating groups are chemically identical. The donor differs from the acceptor by possession of an additional proton. ($K_1 = K_2$, $K_{ij} = K_{ji}$). Either group can equally well act as either donor or acceptor, *e.g.*, a histidyl...histidyl bond of the $NH^+ \ldots N$ type.

II. Double Bonds. (a) **Heterologous.**—The two groups are chemically distinct and, in general, one of them has more of the donor character than the other ($K_1 \neq K_2$), *e.g.*, a carboxyl...amide bond of the



(b) **Homologous.**—The two groups are chemically identical. Furthermore, we cannot distinguish at all between a donor group and an acceptor group, since both groups are donor–acceptor hybrids and maximum bonding is obtained when both groups are in the same ionization state. The acetic acid dimer type bond between two free carboxyl groups 

is an example of such bonding.

Let us now consider a specified i^{th} donor (DH) capable of forming a heterologous single ij^{th} bond (DH...A) with a j^{th} acceptor in a simple manner (*i.e.*, no competition, no coöperation). Since both the donor and the acceptor can also ionize to lose or gain a proton, respectively, we are able to distinguish between six species of the protein molecule on the basis of the behavior of the ij^{th} pair. The hydrogen bonded species are denoted by $P_{(DH...A)}$ and $P_{(D...HA)}$, while those involving no bond are denoted by $P_{(DH,A)}$, $P_{(DH,HA)}$, $P_{(D,HA)}$ and $P_{(D,A)}$ We shall choose as the standard state for the protein species a hypothetical, ideal, one molar solution, at a *p*H such that the net charge[23] on the molecules is zero. We shall call the quantities $P_{(DH...A)}$, etc., the *concentration* fractions of the species they denote, *i.e.*, the total amount of the species present divided by the total amount of protein in the system. Making the usual assumption that at high enough ionic strength the activities of all protein species are proportional to their concentrations we define two hydrogen bonding constants K_{ij} and K_{ji}, and two ionization con-

(22) No definite proof of the presence of these bond types is provided. The various cases are included for the sake of generality.

(23) Whenever the term "net charge" is used it is to be regarded as the "average net charge" over all the protein species present. The fact that the *p*H of standard solutions of various protein species would differ slightly due to their different ionization states is neglected.

stants K_1 and K_2 for the *non-hydrogen bonded* donors and acceptors,[24,25] respectively.

$$K_{ij} = \frac{P_{(DH...A)}}{P_{(DH.A)}} \quad \text{(I-1)}$$

$$K_{ji} = \frac{P_{(D...HA)}}{P_{(D.HA)}} \quad \text{(I-2)}$$

$$K_1 = \frac{P_{(D.A)}[H^+]}{P_{(DH.A)}} = \frac{P_{(D.HA)}[H^+]}{P_{(DH.HA)}} \quad \text{(I-3)}$$

$$K_2 = \frac{P_{(DH.A)}[H^+]}{P_{(DH.HA)}} = \frac{P_{(D.A)}[H^+]}{P_{(D.HA)}} \quad \text{(I-4)}$$

$[H^+]$ denotes the hydrogen ion activity.

When the system is at the isoelectric point, where the net charge Z is zero, the ionization constants K_1 and K_2 are equal to the intrinsic ionization constants K_1° and K_2°. When, however, the pH departs from isoelectric, K_1 and K_2 are no longer equal to K_1° and K_2° but depend on Z, and, therefore, on pH

$$K_1 = K_1^\circ e^{2wZ} \text{ and } K_2 = K_2^\circ e^{2wZ} \quad \text{(I-5)}$$

where w is the electrostatic interaction factor which is customarily determined from the titration curve of the protein. For a uniformly charged sphere, w is given by[26]

$$w = \frac{N\epsilon^2}{2DRT}\left(\frac{1}{b} - \frac{\kappa}{1+\kappa a}\right) \quad \text{(I-6)}$$

where N, R and T have their usual meaning, D is the dielectric constant, ϵ is the electronic charge, κ is the inverse of the radius of the ionic atmosphere, and b and a are the hypothetical radii of the molecule and of exclusion, respectively. Since this model appears rather improbable and since the data available at present on the dimensions of proteins are subject to criticism,[27] it is more convenient to think of w as simply an *experimental* electrostatic interaction factor, required to fit the data.

The nomenclature for homologous single bonds will be the same as above except that, as stated before, $K_{ij} = K_{ji}$ and $K_1 = K_2$. For double bonds we introduce another hydrogen bonding constant K_{lm}, characterizing the double interaction; we still retain the constant K_{ij} (and, if necessary, K_{ji}) to describe single interactions in such systems.

It is useful to define the thermodynamic parameters of formation of the ijth bond, which are simply functions of K_{ij} and of its dependence on temperature and pressure.

$$\Delta F^\circ_{ij} = -RT \ln K_{ij} \quad \text{(I-7)}$$

$$\Delta S^\circ_{ij} = -\left(\frac{\partial \Delta F^\circ_{ij}}{\partial T}\right)_P \quad \text{(I-8)}$$

$$\Delta H^\circ_{ij} = \Delta F^\circ_{ij} + T\Delta S^\circ_{ij} \quad \text{(I-9)}$$

$$\Delta V^\circ_{ij} = \left(\frac{\partial \Delta F^\circ_{ij}}{\partial P}\right)_T \quad \text{(I-10)}$$

Analogous expressions hold for the standard free energy, entropy, enthalpy and volume changes for the formation of the jith bond and of the double lmth bond.

We can further obtain similar expressions for ionization of *non-hydrogen bonded* polar groups

$$\Delta F^\circ_1 = -RT \ln K_1 = (\Delta F^\circ_1)^\circ - 2RTwZ \quad \text{(I-11)}$$

$$\Delta S^\circ_1 = -\left(\frac{\partial \Delta F^\circ_1}{\partial T}\right)_{P,[H^+]} = (\Delta S^\circ_1)^\circ + 2R\left[\frac{\partial (TwZ)}{\partial T}\right]_{P,[H^+]} \quad \text{(I-12)}$$

$$\Delta H^\circ_1 = \Delta F^\circ_1 + T\Delta S^\circ_1 = (\Delta H^\circ_1)^\circ + 2RT\left\{-wZ + \left[\frac{\partial (TwZ)}{\partial T}\right]_{P,[H^+]}\right\} \quad \text{(I-13)}$$

$$\Delta V^\circ_1 = \left(\frac{\partial \Delta F^\circ_1}{\partial P}\right)_{T,[H^+]} = (\Delta V^\circ_1)^\circ - 2RT\left[\frac{\partial (wZ)}{\partial P}\right]_{T,[H^+]} \quad \text{(I-14)}$$

Here the terms $(\Delta F^\circ_1)^\circ$, $(\Delta S^\circ_1)^\circ$, $(\Delta H^\circ_1)^\circ$ and $(\Delta V^\circ_1)^\circ$ are the standard free energy, entropy, enthalpy and volume changes for ionization of a *non-hydrogen bonded* donor group on the protein at a pH where the protein has zero net charge, while ΔF°_1, ΔS°_1, ΔH°_1 and ΔV°_1 are the ionization parameters at any other pH under consideration. Analogous expressions can be written for ionization of acceptors.

Evaluation of K_{ij}

In order to assess the influence of hydrogen bonding on the reactivity of polar groups, it is necessary to have information about the approximate magnitude of the equilibrium constant K_{ij}. Since K_{ij} is determinable from ΔF°_{ij} the problem resolves itself into an estimation of ΔS°_{ij} and ΔH°_{ij}.

Entropy.—Stearn and Eyring[28] suggested that the entropy loss in intramolecular hydrogen bonding is due to a decrease in the amount of internal rotation. More recently Mizushima[29] demonstrated that the entropy change due to intramolecular hydrogen bonding in some model compounds was attributable almost entirely to the decrease in torsional oscillations around single bonds. The evaluation of the entropy change accompanying the formation of hydrogen bonds between polar R groups may be made on the basis of Mizushima's findings.

Consider, as an illustration, a heterologous single hydrogen bond between the hydroxyl group of a tyrosyl residue and the carboxylate ion of a glutamyl residue as represented in Fig. 1. When such an internal hydrogen bond is formed the mass of the whole protein molecule remains constant and the over-all moments of inertia do not change appreciably. Therefore, the entropy change due to translation and rotation of the molecule as a whole may be considered as zero. The entropy change must then be due to a change in the internal degrees of

(24) The assumption that both members of eq. (I-3) and (I-4) are equal is clearly an approximation. The charge on the species involved differs by one and thus the two terms differ by at least a factor of e^{2w}, which, however, is quite close to unity for large proteins since w is small. The difference might be much greater if near neighbor interactions were involved. However it must be remembered that the ionizing groups are relatively free to rotate and thus the charges may be much more than 2.8 Å. apart (approximate length of a hydrogen bond). The approximation that the ionization constant of a *non-hydrogen bonded* group is independent of the ionization state of its neighboring groups will be made throughout this paper. For a somewhat more rigorous treatment of the charge problem see the third paper in this series.[14]

(25) Only three of the four equations listed as (I-3) and (I-4) are independent. All four are listed for the sake of generality.

(26) See, for example, Tanford.[7]

(27) H. A. Scheraga and L. Mandelkern, THIS JOURNAL, **75**, 179 (1953).

(28) A. E. Stearn and H. Eyring, *J. Chem. Phys.*, **5**, 113 (1937).

(29) S. Mizushima, "Internal Rotation," Reilly Lectures, University of Notre Dame, Vol. V, 1951; S. Mizushima, T. Shimanouchi, M. Tsuboi, T. Sugita, E. Kato and E. Kondo, THIS JOURNAL, **73**, 1330 (1951); S. Mizushima, T. Shimanouchi, K. Kuratani and T. Miyazawa, *ibid.*, **74**, 1378 (1952).

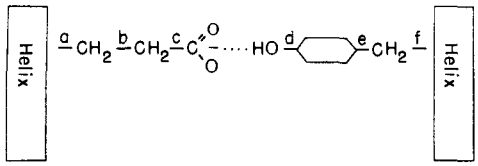


Fig. 1.—Heterologous single hydrogen bond between the hydroxyl group of a tyrosyl residue and the carboxylate ion of a glutamyl residue between two rigid, helical peptide chains. The symbols a–f indicate specific bonds to which reference is made in the text.

freedom of the two participating side chain groups. These degrees of freedom are (a) bond stretching, (b) bending, and (c) oscillation or restricted rotation. As shown by Mizushima, *et al.*, the formation of an internal hydrogen bond in ethylene chlorohydrin does not appreciably affect the bond stretching and bending vibrations. Since the amplitudes of these vibrations are small we may expect that they will not prevent the formation of the hydrogen bond. One must still consider the possibility that the primary OH bond of Fig. 1 may be grossly distorted in the formation of a hydrogen bond thus leading to a large entropy change. However, Mizushima, *et al.*, showed that the entropy change associated with a changed OH stretching frequency in the hydrogen bonded form of ethylene chlorohydrin is very small. Therefore, the only degrees of freedom which may cause large entropy changes ΔS°_{ij} are the internal rotations (or torsional oscillations about single bonds) which become restricted when a hydrogen bond is formed. The amplitudes of these motions are quite high at room temperature and, if the motion continued appreciably in the hydrogen bonded species, the hydrogen bond would be ruptured. Thus, we shall assume that in hydrogen bonded species the internal rotation around single bonds is almost completely "frozen in" and that the entropy loss on formation of a hydrogen bond is essentially equal to the total entropy associated with this rotational motion in non-hydrogen bonded side chains.[30]

A knowledge of the potential energy functions for rotation is necessary for the evaluation of this entropy change. For rotation in hydrocarbons the potential barriers are quite high[31,32] compared to RT at room temperature. Therefore, we can replace the rotational potential functions by those for harmonic oscillators; *i.e.*, we can approximate the motion involved in "internal rotation" as a harmonic torsional oscillation around the potential minimum.[33] Further, since the potential curves for internal rotation in higher hydrocarbons are, in general, asymmetric and the various potential minima differ appreciably,[31b,32] we can assume that the majority of non-hydrogen bonded molecules are confined to the lowest minimum and that the entropy associated with the distribution among various possible minima is small.[30]

The frequency ν of a torsional harmonic oscillator is given by

$$\nu = \frac{1}{2\pi}\sqrt{\frac{\tau}{I_r}} \qquad (I\text{-}15)$$

where τ is the torsional force constant and I_r is the reduced moment of inertia around the single bond. From a consideration of potential functions for various carbon–carbon single bonds in hydrocarbons[31,32] it is found that τ is approximately 5 kcal./rad.^{2}. If we make the further approximation that the oscillations around various single bonds are independent, and if we assume that the moment of inertia of the whole protein molecule (for these rotational motions) is infinite, we can calculate the reduced moment of inertia from the molecular dimensions.[35] Using such values of τ and I_r the entropy associated with torsional oscillations around a single bond can then be obtained from any table[34] of the thermodynamic properties of a harmonic oscillator as a function of $u = h\nu/kT$, where ν is the characteristic frequency given by eq. (I-15) and h, k and T have their usual significance. Approximate calculations based on single bonds in all possible polar side chain groups in proteins yielded values varying from 2.5 to 7.5 e.u. per single bond for the entropy associated with torsional oscillations. It is felt that a value of 5 e.u./bond represents a crude average, which can be used in further approximate considerations. This value is much higher than those usually quoted[34] for entropies of internal rotation (approximately 2 e.u.) in simple model compounds since the moments of inertia I_r of interest here are much higher than those in ethane, for example.

It must be noted that in a few cases the formation of a hydrogen bond need not "freeze in" the internal rotation in *all* the single bonds in the interacting groups. In Fig. 1 the oxygen and the carbon attached to the aromatic ring both lie in the plane of the ring (bonds (d) and (e) lie on the same straight line). The torsional oscillation of the ring around (e) will not affect the position of the oxygen atom relative to bond (f) and thus it will not tend to disrupt the hydrogen bond.

Thus, the problem of obtaining a rough estimate of the entropy lost in formation of an intramolecular hydrogen bond in a very large molecule where this bond cannot appreciably affect the over-all moments of inertia reduces simply to an estimation

(30) We are indebted to Prof. K. S. Pitzer for pointing out that such internal rotations are not *completely* "frozen in," and that the torsional oscillations in the *hydrogen bonded* groups probably contain an amount of internal freedom corresponding to about 1 e.u./bond. However, he also suggests that this value is cancelled in the ΔS calculation by an entropy of approximately 1 e.u./bond corresponding to the distribution, in the *non-hydrogen bonded* form, among various potential energy minima.

(31) (a) E. N. Lassettre and L. B. Dean, Jr., *J. Chem. Phys.*, **17**, 317 (1949); (b) N. W. Luft, *Trans. Faraday Soc.*, **49**, 118 (1953).

(32) K. Ito, THIS JOURNAL, **75**, 2430 (1953).

(33) For a discussion of the mathematical techniques used in the study of internal rotation see Pitzer.[34]

(34) K. S. Pitzer, "Quantum Chemistry," Prentice–Hall, Inc., New York, N. Y., 1953.

(35) Such calculations are only very approximate, since the moment of inertia around say bond (a) of Fig. 1 will depend upon the orientation of the $-CH_2-C(=O)-O^-$ group around bond (b) and upon the orientation of the $-C(=O)-O^-$ group around (c). Since the potential functions for these groups are not known, assignments of orientation of these groups with respect to each other cannot be made. However, for the purpose of the highly approximate solution for I_r sought here any arbitrary orientation will yield satisfactory results since the entropy is not a very sensitive function of τ/I_r.

of the number of single bonds, whose torsional oscillations are restricted by the presence of hydrogen bonds. In the glutamyl...tyrosyl bond (Fig. 1) five single, torsional oscillations are "frozen in" by hydrogen bond formation; thus the entropy of formation of this hydrogen bond, ΔS°_{ij}, is approximately -25 e.u. If similar computations are carried out for polar R groups of amino acid residues other than those illustrated in Fig. 1 it is found that from four to six single bonds are "frozen in" and ΔS°_{ij} is -20 to -30 e.u.

Enthalpy.—The enthalpy of formation of a hydrogen bond of the OH...O, OH...N, NH...O and NH...N type is generally quoted[36,37] as -6 kcal./mole. For an internal hydrogen bond, such as the ones under consideration here, ΔH°_{ij} may be larger or smaller than -6 kcal./mole. It may happen, in a few cases, that the hydrogen bond can form between the two side chain groups only if they are rotated into positions of higher energy (either within the lowest potential well or even in one of the higher potential wells). This will make a positive contribution to ΔH°_{ij} but will not affect ΔS°_{ij} since the system would still be essentially confined to a single potential well, though not necessarily the lowest. On the other hand, a negative contribution to ΔH°_{ij} arises from the fact that the torsional oscillations are "frozen in," *i.e.*, one degree of freedom is removed. Since I_r is large, making $u = h\nu/kT$ very small, the energy of the oscillators approaches the classical mechanical value which, from equipartition, is $RT \sim 600$ cal./bond for a one-dimensional oscillator. There is thus a contribution of $-nRT$ included in ΔH°_{ij} where n is the number of single bonds "frozen in" per hydrogen bond formed. From these considerations one would expect ΔH°_{ij} to vary from about -5 to -9.6 kcal./mole (bonds with ΔH°_{ij} much less negative than -5 kcal./mole would have K_{ij} so small that they would not be observed). For the bond illustrated in Fig. 1 ($n = 5$), $\Delta H^\circ_{ij} \sim -9$ kcal./bond if the orientation of the side chains is extremely favorable.

Free Energy.—Combining a value of $\Delta S^\circ_{ij} = -20$ e.u. with a value of $\Delta H^\circ_{ij} = -6$ kcal./mole, a value of ΔF°_{ij} equal to zero at room temperature is obtained. Thus the values of K_{ij} should be of the order of magnitude of unity.

For the double carboxyl...carboxyl bond interactions ΔS°_{lm} should be expected to vary from about -30 to -40 e.u. (note that in the carboxyl group there is an extra torsional oscillation around the C–OH bond not present in a carboxylate ion), and ΔH°_{lm} may range from -10 to -16 kcal./mole. Thus, K_{lm}'s should be expected to vary from say 1 to 100 or more.

Effect of Water.—The major approximation in the above treatment is the neglect of the influence of water upon intramolecular hydrogen bonding of proteins in solution. Water would be expected to decrease the amount of the hydrogen bonding calculated here.[38,39] However, it is known[5-8] that polar groups in proteins are appreciably hydrogen bonded. Further, Benesch and Benesch[40] find extensive intramolecular hydrogen bonding of the SH...N type in sulfhydryl peptides in water solution, where the hydrogen bonds must be intrinsically much weaker than the OH...O, OH...N, NH...O and NH...N bonds of primary interest here. Also, McDaniel and Brown[41] postulate extensive hydrogen bonding in water solution of some dicarboxylic acids to explain their ionization behavior: and appreciable dimer formation among monocarboxylic acids in water solution was observed by Katchalsky, Eisenberg and Lifson.[42]

Favorable compensating effects[38] may make such hydrogen bonds possible in aqueous solutions.

Value of K_{ij}.—It is of interest to see how well the estimated range of values of K_{ij} agrees with experiment. Tanford accounted for his data[7] on carboxyl and tyrosyl ionization in serum albumin by assuming that a tyrosyl...carboxylate ion hydrogen bond is present.[43] If no hydrogen bond existed, then the ionization constants of the carboxyl and phenolic groups should have the values expected from low molecular weight model compounds, with a pH dependence given by eq. (I-5). The difference between the expected and the observed ionization constants, K_1 (or K_2) and K_{obs}, respectively, is attributed to hydrogen bonding, and serves as a basis for computing K_{ij}. These data are shown in Table I, where K_{ij} is computed from the modified reactivity of both the carboxyl and phenolic groups. It is seen that good agreement is obtained between the values of K_{ij} computed from both groups. Further, using the average value of $K_{ij} = 2.5$, together with Tanford's value of $\Delta H^\circ_1 = 6$ kcal./mole and $\Delta H^\circ_{obs} = 11.5 \pm 1.0$ kcal./mole for the phenolic ionization, one obtains a value of $\Delta H^\circ_{ij} = -7.7$ kcal./mole (from eq. (I-39) below). Combining the average value of $\Delta F^\circ_{ij} = -0.55$ kcal./mole with this value of ΔH°_{ij}, one obtains $\Delta S^\circ_{ij} = -24$ e.u. The result for the entropy change can be expressed in another way

(36) M. Davies, *Ann. Repts. Prog. Chem.*, **43**, 5 (1946).

(37) C. A. Coulson, "Valence," Oxford University Press, London, 1952.

(38) It seems reasonable to expect that *small* entropy and enthalpy changes are involved in the transfer of water from the protein to the solvent, a process involving breakage of water...protein hydrogen bonds and formation of water...water hydrogen bonds. While the size of such hydration effects is impossible to calculate they should render both ΔS°_{ij} and ΔH°_{ij} somewhat less negative. This would give rise to a compensating, and therefore smaller, effect on ΔF°_{ij}.

(39) Two specific effects of the solvent on the magnitude of K_{ij} should be considered. First, the solvent may offer a resistance to the rotational motion of the R groups. From the work of Mizushima and co-workers[29] it appears that the effect of such frictional resistance on the thermodynamic parameters of intramolecular hydrogen bonding is very small. A very large (>15 e.u.) entropy difference was observed between the extended, non-hydrogen bonded (E) form and the hydrogen bonded, coiled (B) form of acetylglycine-N-methylamide in carbon tetrachloride solutions. Such a large entropy difference seems attributable to great differences in the internal rotations (torsional oscillations) between the B and E forms. The thermodynamic parameters for the gauche–*trans* equilibrium in ethylene chlorohydrin were also not changed by transferring the material from the gas phase to CCl_4 solution. The second effect arises from possible hydration of the polar group. Such hydration would tend to increase I_r.

(40) R. E. Benesch and R. Benesch, THIS JOURNAL, **75**, 4367 (1953).

(41) D. H. McDaniel and H. C. Brown, *Science*, **118**, 370 (1953).

(42) A. Katchalsky, H. Eisenberg and S. Lifson, THIS JOURNAL, **73**, 5889 (1951).

(43) The pH range in which such a bond may exist is about 4.5 to 10.5. As will be shown later, this bond seems to be adequately described in this pH range as a heterologous single bond, but appears to be complicated by other types of hydrogen bonding outside this pH range, and even, perhaps, within this range.

by computing ΔS°_{obs} from eq. (I-41) below, using Tanford's value of $\Delta S^\circ_1 = -26 \pm 4$ e.u. and the above values for K_{ij} and ΔH°_{ij}. The computed value of ΔS°_{obs} is -10 e.u. compared to Tanford's experimental value of -8 ± 4 e.u. All of these thermodynamic parameters are in surprisingly good agreement with those estimated from theory, these data being summarized in Table II. A further discussion of the hydrogen bond parameters is presented below, following eq. (I – 41).

TABLE I

EVALUATION[a] OF K_{ij} FROM TANFORD'S DATA[7] FOR THE TYROSYL...CARBOXYLATE ION BOND IN SERUM ALBUMIN IN THE *p*H RANGE 4.5 TO 10.5

Group	ΔF°_1 (or ΔF°_2), cal./mole	ΔF°_{obs}, cal./mole	ΔF°_{ij}, kcal./mole	K_{ij}
Carboxyl	6250	5460	−0.61	2.8
Phenolic	13,000	13,700	− .48	2.2
			Av. = −0.55	Av. = 2.5

[a] As will be shown below
$\Delta F^\circ_{obs} = \Delta F^\circ_1 + RT \ln (1 + K_{ij})$ for the ionization of donors (phenolic groups) at high *p*H (9.5–10.5).
$\Delta F^\circ_{obs} = \Delta F^\circ_2 - RT \ln (1 + K_{ij})$ for the ionization of acceptors (carboxyl groups) at low *p*H (4.5 to 6.0).

TABLE II

COMPARISON OF CALCULATED AND OBSERVED THERMODYNAMIC PARAMETERS FOR A HETEROLOGOUS SINGLE TYROSYL...CARBOXYLATE ION BOND

	ΔF°_{ij}, kcal./mole	K_{ij}	ΔH°_{ij}, kcal./mole	ΔS°_{ij}, e.u.
Calculated	~0	~1	~−6 to −9	~−20 to −25
Observed	−0.55	2.5	−7.7	−24

Ionization

Having obtained information about the approximate size of K_{ij} we can now see its effect on ionization; *i.e.*, the acidity of a specific group of a given kind having an ionization constant, say K_1, will be modified by the presence of the hydrogen bond so that the observed ionization constant, K_{obs}, will, in general, differ from K_1.

Heterologous Single Bond.—For any given bonding situation we can calculate the concentration fractions of all possible species present subject to the normalizing condition that the sum of all the concentration fractions must be equal to unity. For example, for a heterologous single bond

$$P_{(DH...A)} + P_{(D...HA)} + P_{(DH.A)} + P_{(D.HA)} + P_{(DH.HA)} + P_{(D.A)} = 1 \quad \text{(I-16)}$$

Considering the ionization of the *donor* group in such a bond we have

$$K_{obs} = \frac{(P_{(D...HA)} + P_{(D.HA)} + P_{(D.A)})[H^+]}{(P_{(DH...A)} + P_{(DH.A)} + P_{(DH.HA)})} \quad \text{(I-17)}$$

A simultaneous solution[44] of equations (I-1), (I-2), (I-3), (I-4), (I-16) and (I-17) yields

$$K_{obs} = K_1 \left(\frac{1 + [H^+]/K_2 + K_{ji}[H^+]/K_2}{1 + [H^+]/K_2 + K_{ij}} \right) \quad \text{(I-18)}$$

In general, we are concerned with the ionization of donors only in the *p*H range where essentially all acceptors are ionized (*e.g.*, carboxyl groups are already ionized in the *p*H region where the ionization of the tyrosyl hydroxyl is of interest), *i.e.*, in the region where $[H^+] \sim K_1$ (which is $\ll K_2$). In such a case (keeping in mind that K_{ij} is of the order of unity) equation (I-18) reduces to

$$K_{obs} = \frac{K_1}{1 + K_{ij}} \quad \text{(I-19)}$$

Thus, in the protein the ionization of the donor is made more difficult, since the *non-hydrogen bonded* constant K_1 is divided by $(1 + K_{ij})$; *i.e.*, the hydrogen bonded donor cannot ionize as easily as when it is not hydrogen bonded. Thus, as an example, the acidity of tyrosyl OH groups in serum albumin in the *p*H range 9.5 to 10.5 is reduced by hydrogen bonding.[7]

Similarly, for ionization of an *acceptor* group, we have, instead of equation (I-17)

$$K_{obs} = \frac{(P_{(DH...A)} + P_{(DH.A)} + P_{(D.A)})[H^+]}{(P_{(D...HA)} + P_{(D.HA)} + P_{(DH.HA)})} \quad \text{(I-20)}$$

By the method of Appendix I we get

$$K_{obs} = K_2 \left(\frac{1 + K_{ij} + K_1/[H^+]}{1 + (1 + K_{ji})K_1/[H^+]} \right) \quad \text{(I-21)}$$

But in the case of bonds where $K_1 \ll K_2$ (*e.g.*, the carboxylate ion...tyrosyl bond) the range of interest for ionization of the acceptor is usually one where $[H^+] \sim K_2$ (which is $\gg K_1$), *i.e.*, in the range where the acceptor groups (COOH) ionize the vast majority of donors (tyrosyl OH) will be un-ionized. Thus eq. (I-21) becomes

$$K_{obs} = K_2(1 + K_{ij}) \quad \text{(I-22)}$$

and the ionization of the acceptor group is *enhanced* by virtue of the factor $(1 + K_{ij})$. In other words, the COOH group can ionize more easily if there is a tyrosyl OH nearby to form a hydrogen bond with the carboxylate ion. *It is noteworthy that such enhancement would be difficult, if not impossible, to reconcile with the steric hindrance theory of modified reactivity.* This enhancement has been observed by Tanford[7] in the ionization of COOH in the *p*H region 4.5–6.0 and was attributed by him to hydrogen bonding.

Homologous Single Bond.—For homologous single bonds $K_1 = K_2$ and $K_{ij} = K_{ji}$. Thus equation (I-18) or (I-21) can be rewritten as

$$K_{obs} = [H^+] \left(\frac{1 + K_{ij} + K_1/[H^+]}{1 + K_{ij} + [H^+]/K_1} \right) \quad \text{(I-23)}$$

For this case K_{obs} is strongly dependent on *p*H in a manner other than that due to the variation of K_1 with *p*H. We shall use the term "trivial"[45] to refer to the dependence of K_1 (and K_2) on *p*H as given by eq. (I-5), reserving the term "non-trivial" for the increased *p*H dependence of K_{obs} caused by hydrogen bonding. Figure 2 illustrates the effect of hydrogen bonding on the *p*H dependence of pK_{obs} (eq. (I-23)) for a homologous single bond in a hypothetical protein. It is seen that in the low *p*H region ionization is greatly *enhanced* by hydrogen bonding. In fact eq. (I-22) is the low

(44) It is possible to avoid the laborious computations involved in solving so many simultaneous equations by approaching the problem from another point of view. See Appendix I for details. The method of Appendix I will be used throughout this paper.

(45) The "trivial" dependence is not to be regarded as less significant or smaller in magnitude than the "non-trivial" one. The relative magnitudes of the two dependences will be determined by $2w\, dZ/dpH$ in the given *p*H region and by the magnitude of the hydrogen bonding constants.

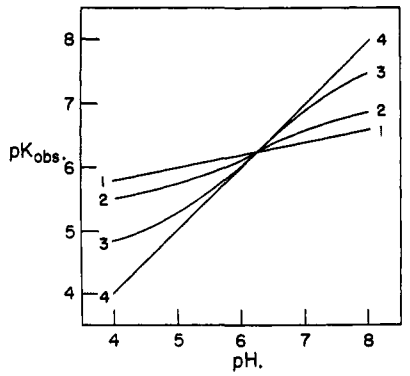


Fig. 2.—A plot of pK_{obs} *vs.* pH for a specified group involved in a homologous single bond of the histidyl... histidyl type; isoelectric point 5.00, $2w(dZ/dpH) = -0.46$, $pK_1° = 6.00$. 1, $K_{ij} = 0$; 2, $K_{ij} = 1$; 3, $K_{ij} = 10$; 4, (limiting case) $K_{ij} = \infty$.

pH limit of this dependence. This enhancement can be understood by consideration of an $NH^+ \ldots N$ type of bond. At low pH both nitrogens possess a proton which is relatively easily ionized since its ionization would lead to the formation of an acceptor group for the hydrogen bond. However, as the pH increases, ionization becomes more and more difficult and as the pH becomes greater than pK_1 the constant K_{obs} is smaller than K_1, with equation (I-19) becoming the high pH limit, *i.e.*, at high pH the presence of the hydrogen bond reduces the tendency of the donor group to lose its proton. At pH = pK_1 (the crossing point[46] of all curves in Fig. 2) equation (I-23) reduces to

$$K_{obs} = K_1 = K_2 \qquad \text{(I-24)}$$

and the ionization is neither enhanced nor hindered due to hydrogen bonding. Since at $pK = p$H half of the groups are ionized, and thus maximum hydrogen bonding can result, any displacement from this pH would lead to a decrease in the amount of bonding. It may be recalled that the condition where pH = pK is the region in a titration curve, where the most exact data are obtained. However, at this pH it can be seen from equation (I-24) that hydrogen bonding of the homologous single type has no effect on K_{obs} or, therefore, on any of the derived thermodynamic functions based on it. Thus, it is clear that the study of titration curves will not be able to yield significant information about hydrogen bonds of this type. If the histidyl ... histidyl bonds are common in proteins, their presence could account for the great difficulty in obtaining unambiguous titration curve data in the histidyl ionization region.[47a,b]

Homologous Double Bond.—We shall now consider a case of a homologous double bond of the carboxyl...carboxyl (or acetic acid dimer) type. Here we are dealing with the possibility of either double (C<O...HO / OH...O>C) or single (C<O,HO / OH...O>C) bond type interactions, characterized by constants K_{1m} and K_{ij}, respectively. Both participating carboxyl groups are assumed to have the same *non-hydrogen bonded* ionization constant K_2. The list of all possible protein species of interest here and their concentration fractions are given in Appendix II. From these we obtain for the ionization of a specified carboxyl group

$$K_{obs} = K_2\left(\frac{1 + K_{1j} + K_2/[H^+]}{1 + 2K_{1j} + K_{1m} + (1 + K_{1j})K_2/[H^+]}\right) \qquad \text{(I-25)}$$

with a sharp, non-trivial dependence on pH.

It is profitable to discuss this equation by imposing a set of restricting conditions, and then later removing these restrictions. Since, as already pointed out, $K_{1m} \gg K_{ij}$ let us first consider eq. (I-25) on the basis of the assumption that $K_{ij} = 0$; *i.e.*, a single hydrogen bond, under this assumption, is not sufficient to couple two carboxyl groups. Before examining real situations it is of interest to see how K_{obs} depends on pH for the unreal case of no electrostatic interactions (*i.e.*, K_2 independent of Z and, therefore, of pH). In such a case, illustrated by curve 1 of Fig. 3, if the pH is low (*i.e.*, $K_2/[H^+] \ll K_{1m}$) eq. (I-25) approaches the low pH limit

$$K_{obs} = \frac{K_2}{1 + K_{1m}} \qquad \text{(I-26)}$$

In this region (pH < 4.00) virtually all the carboxyl groups are present in the un-ionized, doubly bonded form, and the *double* bond has to be broken before a *specified* carboxyl group can ionize. In the intermediate pH region ($K_2/[H^+] \sim K_{1m}$) the probability that the *other* carboxyl group is ionized increases as the pH rises. This correspondingly lowers the probability that the *specified* un-ionized carboxyl group is involved in a double hydrogen bond, and thus the value of K_{obs} increases; *i.e.*, ionization becomes easier as the pH rises. At high pH ($K_2/[H^+] \gg K_{1m}$) virtually all double bonds are broken and the ionization constant K_{obs} closely approaches its high pH limit of simply K_2. Curve 2 of Fig. 3 illustrates the trivial dependence of pK_{obs} on pH (eq. (I-5)), where both K_{1m} and K_{ij} are zero. The "real" situation, illustrated in curve 3 of Fig. 3, is a composite one (however, not strictly a product of curves 1 and 2, but given by

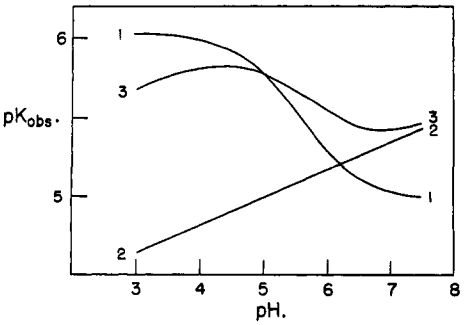


Fig. 3.—A plot of pK_{obs} *vs.* pH for a specified carboxyl group involved in a carboxyl...carboxyl interaction in a hypothetical protein; isoelectric point 5.00, $pK_2° = 5.00$, $K_{ij} = 0$. 1, $K_{1m} = 10$, $2w(dZ/dpH) = 0.00$; 2, $K_{1m} = 0$, $2w(dZ/dpH) = -0.40$; 3, $K_{1m} = 10$, $2w(dZ/dpH) = -0.40$.

(46) The crossing point in Fig. 2 occurs at pH 6.25 rather than 6.00 because of the factor e^{2wZ} (see eq. (I-5)).

(47) (a) R. K. Cannan, A. C. Kibrick and A. H. Palmer, *Ann. N.Y. Acad. Sci.*, **41**, 243 (1941); (b) R. K. Cannan, A. H. Palmer and A. C. Kibrick, *J. Biol. Chem.*, **142**, 803 (1942).

eq. (I-25)). At low pH the slopes of curves 2 and 3 are the same, while at high pH curve 3 approaches curve 2 asymptotically, the electrostatic effect predominating at the extreme pH ranges. It is seen that the composite curve 3 exhibits both a maximum and a minimum, the existence of which will be observed only if the trivial (*i.e.*, electrostatic) dependence is not too great. It can be shown (see Appendix III) that eq. (I-25), for $K_{ij} = 0$, will exhibit both a maximum and a minimum only if

$$\frac{2w}{2.303}\frac{dZ}{dpH} > \frac{1}{2\sqrt{1 + K_{1m}}} - \frac{1}{2} \qquad \text{(I-27)}$$

in the pH region of interest, *i.e.*, for $K_2/[H^+] \sim K_{1m}$.

The behavior of the carboxyl...carboxyl bond is very interesting because of the presence, in a given pH range (see curve 3, Fig. 3), of a region where pK_{obs} decreases with pH. Such a behavior will lead to phenomena which may be loosely termed "all or none ionization," "ionization explosion," or "triggered ionization." The presence of carboxyl...carboxyl double bonds in proteins may well explain the abnormal, low pH steepening of titration of curves of hemoglobin[48] and of serum albumin.[7,49]

If we now remove the restriction that $K_{ij} = 0$, then the curves of pK_{obs} *vs.* pH, given by eq. (I-25), are somewhat more complicated. The nature of the changes will depend on the relative sizes of K_{ij} and K_{1m}. However, it must be kept in mind that $K_{1m} \gg K_{ij}$ and, thus, the presence of a finite K_{ij} will contribute only as a minor perturbation except at very high pH values (where $K_2/[H^+] \gg K_{1m}$), where the limit of eq. (I-25) will not be simply K_2, but rather eq. (I-19).

Heterologous Double Bond.—The ionization behavior of the heterologous double bond of the carboxyl...amide type does not differ greatly from that of a heterologous single bond given in equation (I-19) with K_{1m} essentially replacing K_{ij}. The dependence of the ionization constant on pH is trivial.

Simple Competition.—It is of interest to examine a few cases of limited competition. First, consider a simple case where an ionizable donor (DH) competes for two *already ionized* acceptors (A_1) and (A_2). Here we can distinguish between four protein species $P_{(A1,DH,A2)}$, $P_{(A1...HD,A2)}$, $P_{(A1,DH...A2)}$ and $P_{(A1,D,A2)}$. If the interaction constants of (DH) with (A_1) and (A_2) are $K_{ij}^{(1)}$ and $K_{ij}^{(2)}$, respectively, then, by the method of Appendix I, the concentration fractions of these species are: $1/y$, $K_{ij}^{(1)}/y$, $K_{ij}^{(2)}/y$, $K_1/[H^+]y$ where $y = 1 + K_{ij}^{(1)} + K_{ij}^{(2)} + K_1/[H^+]$. The observed ionization constant of the donor (DH) is then

$$K_{obs} = \frac{(P_{(A_1,D,A_2)})[H^+]}{(P_{(A_1,DH,A_2)} + P_{(A_1...HD,A_2)} + P_{(A_1,DH...A_2)})} \qquad \text{(I-28)}$$

Putting in the concentration fractions, eq. (I-28) reduces to

$$K_{obs} = \frac{K_1}{1 + K_{ij}^{(1)} + K_{ij}^{(2)}} \qquad \text{(I-29)}$$

Thus, the ionization is *hindered* because K_1 is divided by the factor $(1 + K_{ij}^{(1)} + K_{ij}^{(2)})$, *i.e.*, the presence of two possible hydrogen bond acceptors reduces the acidity of a given ionizable donor. Further, K_{obs} depends on pH only in a trivial manner. From symmetry considerations we can write for one *ionizable* acceptor (HA) competing for two un-ionizable donors (D_1H) and (D_2H)

$$K_{obs} = K_2(1 + K_{ij}^{(1)} + K_{ij}^{(2)}) \qquad \text{(I-30)}$$

Complex Competition.—Hydrogen bonding may, however, lead to a non-trivial pH dependence of K_{obs} in more complex competitive situations. For a case of two ionizable donors (D_1H and D_2H), with ionization constants $K_1^{(1)}$ and $K_1^{(2)}$ competing for one already ionized acceptor (A) with interaction constants $K_{ij}^{(1)}$ and $K_{ij}^{(2)}$, respectively, an argument very similar to that given above yields for K_{obs} for the first donor, say D_1H

$$K_{obs}^{(1)} = K_1^{(1)}\left(\frac{1 + K_{1j}^{(2)} + K_1^{(2)}/[H^+]}{1 + K_{1j}^{(1)} + K_{1j}^{(2)} + (1 + K_{ij}^{(1)})K_1^{(2)}/[H^+]}\right) \qquad \text{(I-31)}$$

The non-trivial dependence of $K_{obs}^{(1)}$ on pH is clear. Figure 4 illustrates the $pK_{obs}^{(1)}$ *vs.* pH curve for a few hypothetical cases. It is seen both from the figure and from the model itself that the ionization of D_1H will become progressively more difficult as the pH rises. At low pH, eq. (I-31) reduces to

$$K_{obs}^{(1)} = K_1^{(1)}\left(\frac{1 + K_{1j}^{(2)}}{1 + K_{1j}^{(1)} + K_{1j}^{(2)}}\right) \qquad \text{(I-32)}$$

which, for $K_{ij}^{(1)} = K_{ij}^{(2)} \gg 1$ reduces simply to $K_1^{(1)}/2$. In other words, at low pH both donors possess a proton and either can form a hydrogen bond with the single acceptor, so that the acidity of D_1H is reduced by the statistical factor $1/2$. As the pH rises (and, therefore, D_2H will tend to be ionized), D_1H will become relatively more difficult to ionize since, being the only available donor, it will tend to be hydrogen bonded. In fact, if the K_{ij}'s are large, ionization of D_2H will cause a comparatively small loss in the amount of hydrogen bonding. The high pH limit is, of course, eq. (I-19). It might be noted that a similar non-trivial dependence on pH, where the ionization becomes increasingly more difficult, was observed in the case of a homologous single bond (Fig. 2, eq. (I-23)).

An interesting problem arises in the interpreta-

(48) E. J. Cohn, A. A. Green and M. H. Blanchard, THIS JOURNAL, **59**, 509 (1937). This problem was extensively studied by J. Steinhardt and E. Zaiser, *J. Biol. Chem.*, **190**, 197 (1951), and will be discussed in more detail in a subsequent paper.[14] A titration curve is said to be steepened if $(pK)_{obs}$ falls as the pH rises. It is said to be spread out if $(pK)_{obs}$ rises with pH.

(49) The steepening of the titration curve of serum albumin in the pH range 3–4.5 was attributed by Tanford[7] in 1950 to excessive chloride binding. Keeping w constant the number of chloride ions bound per mole of albumin at each pH was adjusted to make the computed titration curve fit the experimental data. As Tanford pointed out, such a marked increase in chloride binding was greater than that found by G. Scatchard, I. H. Scheinberg and S. H. Armstrong, Jr., THIS JOURNAL, **72**, 535 (1950), and he suggested the possibility that this discrepancy could be accounted for by assuming that w changed with pH. This suggestion was elaborated on in 1952 by Tanford,[7] using the data of Scatchard, Scheinberg and Armstrong to determine Z as a function of pH. While Tanford's explanation of the steepness of the titration curve in terms of a changing w could be correct, it is seen that carboxyl...carboxyl double bonds can also account for the observed behavior. At higher pH values, where these double bonds no longer exist, the carboxyl group ionizes and forms a heterologous single bond with tyrosyl residues,[7] it being implied that a reversible transformation takes place between the two bond types at low pH.

tion of titration curves of groups involved in competitive bonding with high K_{ij}'s. If the pK_1°'s of such groups are high (as, for example, $pK_1^\circ \sim 10$ for tyrosyl groups) the pH region where pK_{obs} shows the maximum effect of hydrogen bonding may be around pH 14 (instead of pH 10–12, as shown in curve 3 of Fig. 4) and, therefore, entirely inaccessible to experiment. In such a situation the experimenter would completely miss the large effect of hydrogen bonding. By misinterpreting the situation he would, instead, observe simply a value of pK_{obs} which would be larger (by an amount log 2 = 0.3, since $K_{obs}^{(1)}$ would equal $K_1^{(1)}/2$ in this region) than the pK_1 he would expect in the absence of hydrogen bonding. Also, he would notice some unexplained spreading of the titration curve.

A somewhat more complex case is that of two donors (D_1H) and (D_2H) competing for two already ionized acceptors (A_1) and (A_2), subject to a restriction that the first donor (D_1H) can be bonded only to the first acceptor (A_1), while the second donor (D_2H) can be bonded to either acceptor. This seems to be a sterically realistic possibility. The possible species and their concentration fractions are listed in Appendix IV. On the basis of these we have

$$K_{obs}^{(1)} = K_1\left(\frac{1 + 2K_{ij} + K_1/[H^+]}{1 + 3K_{ij} + K_{ij}^2 + (1 + K_{ij})K_1/[H^+]}\right) \quad \text{(I-33)}$$

and

$$K_{obs}^{(2)} = K_1\left(\frac{1 + K_{ij} + K_1/[H^+]}{1 + 3K_{ij} + K_{ij}^2 + (1 + 2K_{ij})K_1/[H^+]}\right) \quad \text{(I-34)}$$

Both functions are non-trivially dependent on pH and both show qualitatively the same behavior as eq. (I-31) of ionization becoming more difficult as the pH rises.

Coöperative Bonding.—Another type of hydrogen bonding situation, which may be of some interest in protein systems is that of coöperative bonding. This can be illustrated, by an example of two ionizable donors (D_1H) and (D_2H), both with the same ionization constant K_1, capable of forming hydrogen bonds with an already ionized acceptor (A) in such a manner that in some cases both hydrogen bonds can exist simultaneously. Such bonding is stereochemically very unlikely if the acceptor group is a single atom. However, if the acceptor group, *e.g.*, carboxylate ion, contains in it two atoms both capable of functioning as acceptors, the steric objections to the bonding. of say a tyrosyl . . . carboxylate ion . . . tyrosyl type, are somewhat less. If such cases exist we could distinguish here among nine protein species on the basis of the behavior of the three groups involved, *i.e.*, $P_{(D_1H\ \ldots\ A\ \ldots\ HD_2)}$, $P_{(D_1H\ \ldots\ A.\ HD_2)}$, $P_{(D_1H\ A\ \ldots\ HD_2)}$, $P_{(D_1,\ A\ \ldots\ HD_2)}$, $P_{(D_1.\ A.\ HD_2)}$, $P_{(D_1H\ \ldots\ A.\ D_2)}$, $P_{(D_1H.\ A.\ D_2)}$, $P_{(D_1H.\ A.\ D_2H)}$, and the doubly ionized $P_{(D_1.\ A.\ D_2)}$.

Defining a coöperative interaction constant K_{rs} as

$$K_{rs} = \frac{P_{(D_1H\ldots A\ldots HD_2)}}{P_{(D_1H.\ A.\ D_2H)}} \quad \text{(I-35)}$$

and assign to the single interactions the usual constants K_{ij} (the same for both donors), we can obtain for K_{obs} for the ionization of a specified donor[50]

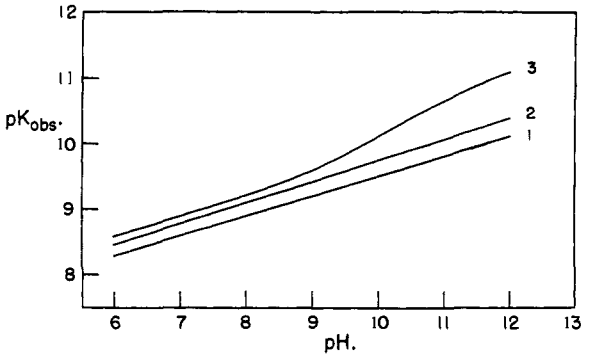


Fig. 4.—A plot of pK_{obs} *vs.* pH for a hypothetical, specified donor competing with another donor for an already ionized acceptor (eq. (I-31)); isoelectric point 5.00, $2w(dZ/dpH) = -0.69$, $(pK_1^{(1)})^\circ = (pK_1^{(2)})^\circ = 8.00$, $K_{ij}^{(1)} = K_{ij}^{(2)}$. 1, $K_{ij} = 0$; 2. $K_{ij} = 1$; 3, $K_{ij} = 10$.

$$K_{obs} = K_1\left(\frac{1 + K_{ij} + K_1/[H^+]}{1 + 2K_{ij} + K_{rs} + (1 + K_{ij})K_1/[H^+]}\right) \quad \text{(I-36)}$$

The value of the constant K_{rs} would be expected to be much larger than that of K_{ij}^2 since the enthalpy of the double interaction is expected to be roughly twice larger than that of a single one ($\Delta H^\circ_{rs} \sim 2\Delta H^\circ_{ij}$), while the entropy change is due mainly to "freezing in" of rotational oscillations in three rather than four interacting groups ($\Delta S^\circ_{rs} \sim {}^3/_2\Delta S^\circ_{ij}$). Thus we may expect that $K_{lm} > K_{rs} > K_{ij}^2$.

It is seen that if $K_{rs} > K_{ij}^2$ (*i.e.* the bond is a coöperative one) the ionization becomes easier as the pH increases, the low pH limit of eq. (I-36) being

$$K_{obs} = K_1\left(\frac{1 + K_{ij}}{1 + 2K_{ij} + K_{rs}}\right) \quad \text{(I-37)}$$

and the high pH limit being given by equation (I-19). Since the ionization of the second group becomes much easier when the first group ionizes the coöperative hydrogen bonding would lead to "coupled" or "triggered" ionization already discussed in the case of homologous double bonds of the carboxyl . . . carboxyl type, with a similar non-trivial dependence of pK_{obs} on pH (see Fig. 3). Coöperative bonding would lead to steepening of titration curves in the high pH region and thus may account for the abnormally steep tyrosyl (OH) titration curve observed by Tanford[7] in serum albumin in the pH range 10.5 to 11, without the necessity of involving variable w values.

It must be admitted that coöperative and competitive bonding of the type discussed above appear sterically rather unlikely. However, it is possible that in order to form an ijth hydrogen bond the helically-folded peptide chains must be in certain geometric relations to each other. Several hydrogen bonds between the polar R groups may well be involved in maintaining such a structure, while the existence of several others might prevent it. In such cases we would have *distant* coöperation or competition among sets of hydrogen bonds. While

(50) Equation (I-36) is a generalized equation for the behavior of single bonds in competitive, non-competitive, and coöperative cases. Thus, if $K_{rs} = 0$, the bond is competitive and (I-36) reduces to (I-31) for $K_{ij}^{(1)} = K_{ij}^{(2)}$ and $K_1^{(1)} = K_1^{(2)}$; if $K_{rs} = K_{ij}^2$, the bond is non-competitive and (I-36) reduces to (I-19); if $K_{rs} >> K_{ij}^2$, then the situation is a coöperative one characterized by (I-36).

such a treatment is, at present, beyond the scope of our model, where for the sake of simplicity we have assumed that the helices are rigidly held to each other, it is worth noting that the behavior of ionizable groups involved in such distant competition or coöperation would be very similar to the near neighbor interactions discussed here.

It is also worth noting that the hydrogen bonding structure maintaining the helical form of the peptide chain is an extreme form of coöperative interaction. Therefore, such a hydrogen bonding situation would be expected to be one of very high stability compared to that between the polar R groups. *Thus, the hydrogen bonds between polar R groups should be more susceptible to breakage by denaturing agents than those which maintain the helical configuration of the peptide chain.* The applicability of these considerations to the problem of protein denaturation will be dealt with in a forthcoming paper.[14]

Thermodynamic Functions for Ionization.—From the calculated expressions for K_{obs} we can obtain the corresponding thermodynamic functions for ionization. Thus, for the ionization of donors in heterologous single bonds (eq. (I-19)), where the acceptors are already ionized, we have (by application of equations analogous to (I-7) — (I-10))

$$\Delta F^\circ_{obs} = \Delta F^\circ_1 + RT \ln (K_{ij} + 1) \quad \text{(I-38)}$$
$$= \Delta F^\circ_1 - \frac{K_{ij}}{K_{ij}+1}\Delta F^\circ_{ij} - RT\left[\frac{K_{ij}}{K_{ij}+1}\ln K_{ij} - \ln (K_{ij}+1)\right]$$

$$\Delta H^\circ_{obs} = \Delta H^\circ_1 - \frac{K_{ij}}{K_{ij}+1}\Delta H^\circ_{ij} \quad \text{(I-39)}$$

$$\Delta V^\circ_{obs} = \Delta V^\circ_1 - \frac{K_{ij}}{K_{ij}+1}\Delta V^\circ_{ij} \quad \text{(I-40)}$$

$$\Delta S^\circ_{obs} = \Delta S^\circ_1 - R \ln (K_{ij}+1) - \frac{K_{ij}}{K_{ij}+1}\frac{\Delta H^\circ_{ij}}{T} \quad \text{(I-41)}$$
$$= \Delta S^\circ_1 - \frac{K_{ij}}{K_{ij}+1}\Delta S^\circ_{ij} + R\left[\frac{K_{ij}}{K_{ij}+1}\ln K_{ij} - \ln (K_{ij}+1)\right]$$

where ΔF°_1, ΔH°_1, ΔV°_1 and ΔS°_1 are given by equations (I-11) — (I-14). Some writers[26] have referred to the quantities $\Delta F^\circ_1 - \Delta F^\circ_{obs}$, $\Delta H^\circ_1 - \Delta H^\circ_{obs}$, $\Delta V^\circ_1 - \Delta V^\circ_{obs}$ and $\Delta S^\circ_1 - \Delta S^\circ_{obs}$ as the free energy, enthalpy, volume and entropy changes associated with hydrogen bonding. However, these quantities are *not* equal to ΔF°_{ij}, ΔH°_{ij}, ΔV°_{ij} and ΔS°_{ij}, respectively, even in the simplest case described by eq. (I-19). Such an approximation would be valid only for the case of $K_{ij} >> 1$, as can be seen from eq. (I-38) — (I-41). This condition appears to be *nearly* satisfied in serum albumin[7] where we have estimated K_{ij} to be 2.5 for Tanford's tyrosyl . . . carboxylate ion bond, but would not be so in general.

In deriving equations of the type (I-38) — (I-41) for various hydrogen bonding situations it was always more convenient to derive an expression for K_{obs} first and then obtain the thermodynamic parameters by considering the dependence of K_{obs} on temperature and pressure. However, if equations like (I-38) — (I-41) are to be derived directly it must be remembered that the observed thermodynamic parameters are sums of three components. For example, for the free energy of ionization of the donor in a heterologous single bond (where the acceptors are already ionized) we have

$$\Delta F^\circ_{obs} = \Delta F^\circ_1 - n_{ij}\Delta F^\circ_{ij} - RT\{n_{ij}\ln n_{ij} + (1 - n_{ij})\ln (1 - n_{ij})\} \quad \text{(I-42)}$$

where (a) the first term, ΔF°_1, is the free energy of ionization in the absence of hydrogen bonding, (b) the second term is the negative of the free energy of formation of a hydrogen bond multiplied by n_{ij}, the concentration fraction of the hydrogen bonded species in the un-ionized form (since only the un-ionized form can be hydrogen bonded), and (c) the third term is the negative of the free energy of mixing of the hydrogen bonded and non-hydrogen bonded species in the un-ionized form. The concentration fractions of the un-ionized forms $P_{(DH,A)}$ and $P_{(DH\ \ldots\ A)}$ are $1/y$ and K_{ij}/y, where $y = 1 + K_{ij} + K_1/[H^+]$, so that the concentration fraction of $P_{(DH\ \ldots\ A)}$ in the un-ionized forms is $n_{ij} = K_{ij}/(K_{ij} + 1)$. Substitution of this value for n_{ij} and $RT \ln K_{ij}$ for $-\Delta F^\circ_{ij}$ leads to eq. (I-38). Thus, even if $K_{ij} = 1$ (and, therefore, $\Delta F^\circ_{ij} = 0$), there is still a hydrogen bonding contribution of RT ln 2, arising from the free energy of mixing, to be added to ΔF°_1 in eq. (I-38) or (I-42).

Binding of Small Molecules and Ions[51]

The treatment presented above for the binding of protons can also be applied to the more general problem of the binding of small molecules and ions by the *polar* R groups in proteins. No account will be taken of van der Waals interactions which probably play a role in the binding of the hydrocarbon portions of detergent and dye molecules. As in the case of ionization, we should expect to find that the observed binding behavior reflects the presence of internal hydrogen bonds, and, therefore, should depend on K_{ij}. In the binding problem the treatment will be restricted, for simplicity, to a pH range where only the donor (DH) can ionize.

Heterologous Single Bond.—As a first example, consider the case in which a donor (DH) can form an intramolecular ijth hydrogen bond with an already ionized acceptor (A); the donor is also assumed to be capable of binding a small molecule or ion (B). For this problem, all the previously used equilibrium constants will be required, together with an additional one, K_B. The latter characterizes the binding of (B) to (DH) and is defined by the expression

$$K_B = P_{(DH\ldots B,A)}/[B]P_{(DH,A)} \quad \text{(I-43)}$$

where [B] is the activity of unbound (B) in solution. If (B) is uncharged, K_B (analogous to K°_1, or rather to its reciprocal) is constant, but if (B) carries a charge K_B varies with Z, and thus with both pH and [B], similar to K_1 (see eq. (I-5)).

If we also introduce an equilibrium constant $(K_B)_{obs}$ (analogous to K_{obs} for ionization), then we obtain (by the method of Appendix I), for a heterologous single bond

$$(K_B)_{obs} = K_B/(1 + K_{ij} + K_1/[H^+]) \quad \text{(I-44)}$$

This equation describes the dependence of $(K_B)_{obs}$ on pH, the observed equilibrium constant being *independent* of [B]. Further, it is seen that the small (or even positive) entropy changes associated

(51) For a review see I. M. Klotz, "The Proteins" (Ed. by Neurath), Vol. IB, Academic Press, Inc., New York, N. Y., 1953, pp. 727 ff.

with binding[8] may be due to the balancing of an over-all entropy loss on binding with an entropy gain due to breakage of some hydrogen bonds between (DH) and (A). This notion was already introduced by Karush[52] who suggested "that part of the entropy increase in anion binding is due to the reversible conversion of the protein to a less condensed state involving the rupture of intramolecular bonds." It should be noted that breakage of hydrogen bonds between (DH) and (A) need not involve unfolding or changes in the assumed helical configurations of the peptide chains.

Competition.—The case of two *un-ionizable* donors (D_1H) and (D_2H) competing for one acceptor (A) leads to a dependence of $(K_B)_{obs}$ on [B]. Assuming that both donors have the same values of K_{ij} and K_B we obtain, for the binding of (B) to a *specified* donor

$$(K_B)_{obs} = K_B\left(\frac{1 + K_{ij} + K_B[B]}{1 + 2K_{ij} + K_B[B](1 + K_{ij})}\right) \quad \text{(I-45)}$$

This equation is completely analogous to eq. (I-31) with $K_1/[H^+]$ replaced by $K_B[B]$. From Fig. 5 it is seen that $(K_B)_{obs}$ will *diminish* with an increase in [B] and thus in the amount of material bound. Carsten and Eisen[53] found for the binding of substituted 2.4-dinitrobenzenes by serum albumin that $(K_B)_{obs}$ varied with the amount of material bound; *i.e.*, their data showed a decrease in the apparent value of $(K_B)_{obs}$. They were unable to account for such a variation on the basis of several K_B's for different kinds of binding sites and suggested that the variation may be due to competition between binding and intramolecular hydrogen bonding. It is seen that eq. (I-45) could well explain this behavior. It is, however, also worth noting that the dependence of $(K_B)_{obs}$ on [B] is not caused simply by competition between binding and intramolecular hydrogen bonding at constant *p*H (see eq. (I-44)), but results only if the intramolecular bonding itself is competitive.

This treatment of competitive bonding could be easily extended (by the method of Appendix I) to the case where the donors are considered to be ionizable. In such a case $(K_B)_{obs}$ would depend on both [B] and *p*H.

Coöperative Bond.—If coöperative hydrogen bonding is involved then $(K_B)_{obs}$ will *increase* with [B]. By analogy with (I-36) we have, for the binding of (B) to a specified donor

$$(K_B)_{obs} = K_B\left(\frac{1 + K_{ij} + K_B[B]}{1 + 2K_{ij} + K_{rs} + K_B[B](1 + K_{ij})}\right) \quad \text{(I-46)}$$

In extreme cases where $K_{rs} \gg K_{ij}^2$ this dependence on [B] should lead to an increase in the binding affinity or to "all or none binding" (see Fig. 5). A behavior of this type was observed by Karush,[52] who found a sigmoid r/c *vs.* r plot (see Klotz[51] for definition of these symbols) for the binding of the *d*-form of an optically active anionic azo dye by serum albumin. Such a curve indicates that, in a certain region of the free dye concentration, the intrinsic binding constant increased with the equilibrium amount of free dye. This effect was attributed by Karush to "opening up" of the protein

(52) F. Karush, *J. Phys. Chem.*, **56**, 70 (1952).

(53) M. E. Carsten and H. N. Eisen, THIS JOURNAL, **75**, 4451 (1953).

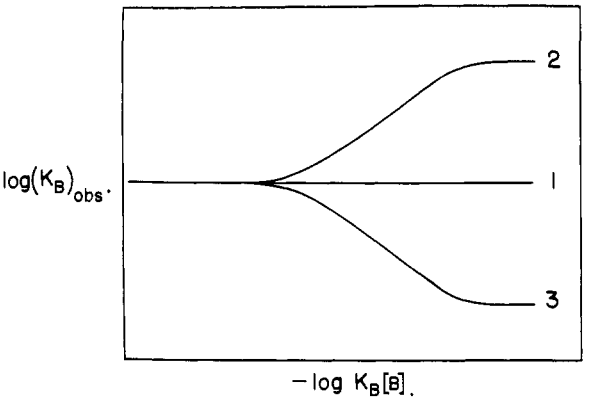


Fig. 5.—A schematic plot of log $(K_B)_{obs}$ *vs.* $-\log K_B[B]$ (at constant *p*H) illustrating binding to a donor involved in a heterologous single bond (curve 1, eq. (I-44)), to a specified donor involved in a competitive (curve 2, eq. (I-45)), and a coöperative (curve 3, eq. (I-46)) hydrogen bonding situation.

molecule resulting from breakage of secondary intramolecular linkages. It is seen that coöperative hydrogen bonding explains such "opening up" of sites without the necessity of assuming any gross geometric changes in the molecule.

Another possible example of this phenomenon is furnished by the study of Warner and Weber[54] of metal binding by conalbumin. These authors investigated the binding of ferric, cupric and zinc ions and found that each protein molecule contains two specific sites to account for the binding of two metal ions. Further, they presented evidence to show that the phenolic groups of tyrosine (together with an unidentified R group, possibly carboxylate ion) are an essential part of the binding sites which coördinate the metal, with "interaction between the two binding sites such that the second metal ion is associated more readily than the first." It was also shown that the tyrosyl groups had to be ionized for the binding of the metal ion to take place. This increased binding affinity for the second metal ion can be explained in terms of coöperative hydrogen bonding by considering, as an example, their data for copper binding. In our nomenclature, their data indicate that $(\Delta F^\circ_B)_{obs}^{(2)}$, for binding the second cupric ion, is 2600 cal. more negative than $(\Delta F^\circ_B)_{obs}^{(1)}$, for binding the first cupric ion. Assume that the two sites are linked by a coöperative hydrogen bond (say tyrosyl$^{(1)}$ ... carboxylate ion ... tyrosyl$^{(2)}$), where tyrosyl$^{(1)}$ is part of the binding site for the first ion and tyrosyl$^{(2)}$ is part of the site for the second ion, the two sites being equivalent. In order for the first cupric ion to bind, the coöperative hydrogen bond must break and tyrosyl$^{(1)}$ must ionize. Once this has been accomplished, the coöperative interaction would no longer exist and the second cupric ion can bind more readily to tyrosyl$^{(2)}$, *i.e.*, the observed free energy for binding the second ion would be more negative than that for binding the first ion. A value of K_{rs} can be computed from this enhanced binding capacity. Assuming $K_{ij} = 0$ and $pH \ll (pK_1)_{obs}$, $(K_B)_{obs}^{(1)} = K_B/(1 + K_{rs})$ whereas $(K_B)_{obs}^{(2)} = K_B$ or

(54) R. C. Warner and I. Weber, *ibid.*, **75**, 5094 (1953).

$(\Delta F^{\circ}_{B})_{obs}^{(2)} - (\Delta F^{\circ}_{B})_{obs}^{(1)} = -RT \ln (1 + K_{rs})$ which has the observed value of -2600 cal. at room temperature if $K_{rs} = 75$. Alternatively, if two such coöperative bonds are involved between the two sites, then the free energy difference would be $-2RT \ln (1 + K_{rs})$ which has a value of -2600 cal. at room temperature if $K_{rs} = 8$. These are reasonable values for K_{rs}. Similar calculations can be made with the data of Warner and Weber for ferric ion binding where the observed free energy difference is 3600 cal. Quite clearly, several other coöperative situations, especially those involving *distant* coöperation, could account for these observations. Those cited above are quoted as examples only.

As with the competitive case, the coöperative one can also be extended to the situation where the donors are considered to be ionizable. For example, allowing the donors to ionize in the coöperative case, we have, for the binding of (B) to a specified donor

$$(K_B)_{obs} = K_B \left(\frac{1 + K_{ij} + K_B[B] + \frac{K_1}{[H^+]}}{1 + 2K_{ij} + K_{rs} + K_B[B](1 + K_{ij}) + \frac{K_1}{[H^+]}\left(2 + 2K_{ij} + K_B[B] + \frac{K_1}{[H^+]}\right)} \right) \quad \text{(I-47)}$$

The dependence of this function on pH at *constant* $K_B[B]$ will be governed by the relative sizes of the constants involved. However, it can be shown that at least in some cases $(K_B)_{obs}$ will not be a monotonically decreasing function of pH but it will exhibit a maximum.[55] This can be seen by considering the model again. The double interaction DH . . . A . . . HD is greatly preferred to binding, and thus, at low pH (*i.e.*, low pH compared to pK_1), $(K_B)_{obs}$ will be small. As ionization proceeds, however, it is seen that fewer and fewer such double interactions are possible. Since formation of a single hydrogen bond is not nearly as probable as that of a double one, the small molecules may readily compete for binding sites leading to an increase in $(K_B)_{obs}$. As ionization proceeds further the number of donors (DH) is almost completely depleted and $(K_B)_{obs}$ falls again. Klotz and co-workers[56] found that serum albumin had more tyrosyl hydroxyl sites for binding dyes at high pH than at low pH. In conjunction with titration data of Tanford[7] and binding data of Karush[52] (already discussed) it is suggested that this anomalous binding behavior may be due to coöperative bonding between the tyrosyl groups and, say, carboxylate ion.

Effect of Binding on pK_{obs}.—It is clear that while we have treated the ionization as something affecting the binding, the opposite point of view could have been taken; *i.e.*, $(K_1)_{obs}$ for the ionization of a donor depends not only on how this donor is hydrogen bonded to an acceptor, but also on whether the donor can bind some small ion. For example, if (DH) can bond to (A), and can also bind chloride ion, the value of $(K_1)_{obs}$ will also depend on K_B, its affinity for chloride ion. It is thus possible to modify our treatment of ionization where heterologous single bonds are involved, to take account of the possibility that (DH) can bind small anions. In such a case, eq. (I-19) would be replaced by

$$(K_1)_{obs} = \frac{K_1}{1 + K_{ij} + K_B[B]} \quad \text{(I-48)}$$

Since the bound (B) would change the value of Z, the electrostatic dependence of K_1 would also be affected. The extensions to competitive and coöperative cases, as well as their significance for the interpretation of titration curves, are obvious.

Discussion

Various anomalies have appeared in titration curves of proteins and in curves representing the binding of small molecules and ions by proteins. These anomalies have arisen because the reactivity of the various polar R groups is different from that expected from low molecular weight model compounds. This modified reactivity has often been "explained" in such terms as inaccessibility of reactive groups, unfolding of peptide chains, etc., whereas it has been shown here, quantitatively, that there is a thermodynamic basis for modified reactivity arising from a significant contribution of hydrogen bonding between polar R groups.

In previous interpretations of titration curves[57-59] and in binding problems[51] the only effects which have been generally considered were those due to the change of net charge on the protein (*i.e.*, electrostatic effects) and those due to the fact that many groups were capable of participating in a given reaction (*i.e.*, statistical factors). Thus, only the *trivial* dependences of pK_{obs} on pH, or of $(pK_B)_{obs}$ on [B], could be accounted for. It has been shown here that important *non-trivial* dependences could arise as a result of hydrogen bonding if K_{ij}'s, K_{lm}'s and K_{rs}'s are not vanishingly small.

Further, it is suggested that some cases where such hydrogen bonding exists have been overlooked and the data misinterpreted by a fortuitous choice of parameters pK°_1, w and Z (and, for the binding problem, K_B's and n_i's, where n_i is the number of binding sites of a given type having a given binding constant). Thus, good fits may have been obtained between experimental data and calculated curves by parameters which would have been significantly different if allowance were made for the possible existence of hydrogen bonding constants K_{ij}, K_{lm} and K_{rs}. It may be recalled that the values of Z obtained from titration curves are not in good agreement[47a,60] with those from electrophoretic mobility or membrane potential measurements. The agreement obtained between the electrostatic factor w,

(55) The condition for a maximum in $(K_B)_{obs}$ of eq. (I-47) is $K_{rs} > (1 + 2K_{ij} + 2K_{ij}^2 + 2K_B[B](1 + K_{ij}) + K_B^2[B]^2)$ for $K_B[B]$ constant.

(56) I. M. Klotz, R. K. Burkhard and J. M. Urquhart, *J. Phys. Chem.*, **56**, 77 (1952).

(57) K. Linderstrøm-Lang, *Compt. rend. trav. lab. Carlsberg*, **15**, No. 7 (1923–1925).

(58) J. G. Kirkwood, *Ann. N. Y. Acad. Sci.*, **41**, 321 (1941).

(59) For a review see R. K. Cannan, *Chem. Revs.*, **30**, 395 (1942).

(60) A. Tiselius and H. Svensson, *Trans. Faraday Soc.*, **36**, 16 (1940).

used in fitting the data, and that obtained from eq. (I-5) and "reasonable" molecular dimensions becomes even less meaningful in light of the oversimplification in the model of a uniformly charged sphere. In light of these considerations, even the excellent internal agreement of some titration curve data[47a,b] cannot be considered as proof of the lack of dependence of pK_{obs} on pH in these cases. Reference to Fig. 4 and the discussion presented there shows how easily a competitive hydrogen bonding case, for example, could be overlooked, and "good agreement" achieved with somewhat erroneous parameters.

However, some effects of hydrogen bonding between polar R groups have been recognized previously.[5-8] For example, in an outstanding investigation Tanford reported the existence of a tyrosyl . . . carboxylate ion hydrogen bond in the pH region 4.5 to 10.5 as a reflection, among other things, of the fact that carboxyl ionization was enhanced and tyrosyl ionization reduced in this pH region. This behavior is indicated by eq. (I-19) and (I-22) for a homologous single bond. However, below pH 4.5 and above pH 10.5 his observed pK's no longer had the electrostatic or *trivial* dependence on pH required by eq. (I-5). He, therefore, assumed that w changed in these regions due to gross molecular unfolding. Without disputing this assumption, we have shown here that the low pH effect (pH 3–4.5) could also be accounted for by the existence of carboxyl . . . carboxyl double bonds[61] while the high pH effect (pH 10.5 to 11) may be due to coöperative bonding of the type tyrosyl . . . carboxylate ion . . . tyrosyl. Both these low and high pH effects are manifested by decreasing values of pK_{obs} as the pH rises, and give rise to a steepening of the titration curve in the pH regions where they occur. Competitive bonding, on the other hand, gives rise to an increase in the values of pK_{obs} as the pH rises and, therefore, to a spreading out of the titration curve in this pH region. The electrostatic effect (eq. (I-5)) can account for only a spreading out effect, inasmuch as Z influences the value of pK_{obs}. Equation (I-5) certainly cannot account for the steepening of titration curves (assuming w to remain constant).

Similarly, the statistical and electrostatic effects[51] cannot account for anomalies in binding curves where $(K_B)_{obs}$ depends on [B] as illustrated in several examples cited.[53,54] On the other hand, various hydrogen bonding situations will lead to just such a behavior, as already illustrated.

It may be noted that while the treatment presented here can account for some of the bewildering observed pH dependences of protein reactions it will be a difficult problem to apply these considerations of K_{obs} to the detailed interpretation of complete titration curves in a quantitative fashion. The introduction of parameters (K_{ij}, K_{jj}, K_{lm}, K_{rs}) dependent upon the nature of local interactions invalidates the usual assumption that all groups of a given kind in a protein are characterized by the same observed ionization constant. An introduction of individual ionization constants K_{obs} for every particular group in the molecule would greatly complicate the final expressions for the titration curve. An even greater mathematical complication would arise if several groups in the molecule were involved in extensive competitive or coöperative interactions.

Thus, there are several severe limitations in the use of titration curve analysis as a method of locating the internal hydrogen bonds in proteins. These are (a) mathematical complications of the detailed theory, (b) insensitivity of the ionization behavior of some groups (*e.g.*, homologous single bond, eq. (I-24)) to the value of K_{ij}, (c) the expectation that some large effects of hydrogen bonding on titration curves (*e.g.*, competition among donors, Fig. 4) may lie outside of the experimentally accessible pH range. Some of these difficulties may be resolved by searching for hydrogen bonds by other thermodynamic procedures, namely, by considering the effect of such bonds upon the "apparent strength" of various primary bonds in proteins. This problem is treated[14] in a subsequent paper of this series.

The treatment presented here in terms of a hydrogen bonded model is applicable[14,15] to other types of protein reactions such as associations, enzyme reactions, denaturation and proteolysis.

It is recognized that there are oversimplifications in the model used here. Among these may be mentioned (a) the neglect of non-polar interactions which probably play a role in detergent and dye binding, and in insulin association, (b) the problem of hydration, already mentioned, and (c) an oversimplification of the charge problem which will be treated more extensively[14] in a subsequent paper of this series.

Nevertheless, it is worthwhile to mention that the theory presented here is a unified one. It provides a quantitative basis for considering all types of protein reactions from a thermodynamic point of view of intramolecular coöperation. In this sense, the mathematical analysis serves in place of the vagueness of such terms as "chain unfolding" or "hydrogen bonding" (such terms usually not being specific about the nature of the interaction), and provides a basis for testing various proposed internal configurations in a quantitative manner. Further, as will be shown in subsequent papers,[14,15] it provides a valuable new experimental approach for investigating the internal structure of proteins, and for testing current hypotheses about the configurations of the peptide chains within the protein molecule.

Acknowledgment.—The authors wish to thank Prof. R. Bersohn for helpful discussions of torsional oscillations and Dr. N. Notley for suggestions related to competitive bonding.

Appendix I

Simplified Procedure for Obtaining Concentration Fractions.—The relative concentrations of all

(61) The reversible reaction in serum albumin, recently reported by Gutfreund and Sturtevant,[62] occurs in the pH region 3–4.5. This reaction may actually be a reversible conversion from carboxyl... carboxyl double bonds to tyrosyl...carboxylate ion bonds. The relation between this phenomenon and reversible changes in hydrodynamic properties is at present under investigation.[63]

(62) H. Gutfreund and J. M. Sturtevant, THIS JOURNAL, **75**, 5447 (1953).

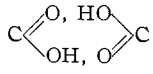
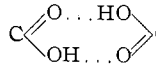
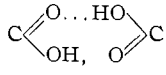
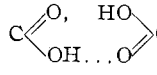
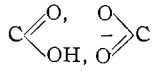
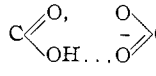
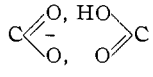
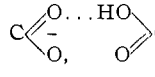
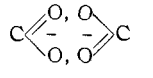
(63) H. A. Saroff, G. I. Loeb and H. A. Scheraga, unpublished results.

species given in eq. (I–16) can be obtained by arbitrarily assigning to one of the species the concentration fraction $1/y$ where y is some number (see below). Thus, *e.g.*, the species $P_{(DH.\ A)}$ is given a concentration fraction $1/y$. Then from eq. (I-1), $P_{(DH\ \ldots\ A)} = K_{ij}/y$ and from eq. (I-3), $P_{(D,A)} = K_1/[H^+]y$. Similarly, all species of eq. (I-16) are obtainable by use of eq. (I-1) to (I-4). It can be seen that y is determinable from eq. (I-16). Substitution of these concentration fractions into eq. (I–17) leads directly to eq. (I-18).

Appendix II

Carboxyl...Carboxyl Interaction.—The protein species involved in a carboxyl...carboxyl interaction are given below together with their concentration fractions x_i. The values are tabulated as $x_i y$ where

$$y = 1 + 2K_{ij} + K_{Im} + 2(1 + K_{ij})K_2/[H^+] + (K_2/[H^+])^2$$

	Species	$x_i y$
(A)		1
(B)		K_{Im}
(C)		K_{ij}
(D)		K_{ij}
(E)		$K_2/[H^+]$
(F)		$K_{ij}K_2/[H^+]$
(G)		$K_2/[H^+]$
(H)		$K_{ij}K_2/[H^+]$
(I)		$K_2^2/[H^+]^2$

Consider the ionization of a specified carboxyl group. Suppose that the right hand carboxyl group is the one whose ionization is being considered.

$$K_{obs} = \frac{(x_E + x_F + x_I)[H^+]}{(x_A + x_B + x_C + x_D + x_G + x_H)}$$

Substitution for the concentration fractions, x_i, in the above equation leads to eq. (I-25).

For the sake of simplicity no distinction has been made between the K_{ij}'s for species (C) and (H) and for the K_2's of the two carboxyl groups. While such distinctions could be introduced they would only complicate the equations unnecessarily.

Appendix III

Condition for "Ionization Explosion."—Equation (I-25) can be rewritten, for $K_{ij} = 0$, as

$$K_{obs} = K_2\left(\frac{1 + K_2/[H^+]}{1 + K_{Im} + K_2/[H^+]}\right) \qquad (1)$$

where K_2 is given by eq. (I-5). If we further assume that over the pH region of interest the net charge, Z, on the protein varies as a linear function of pH or ln $[H^+]$ we can write

$$Z = Z'_0 + \frac{dZ}{d \ln [H^+]} \ln [H^+] = \frac{Z_0}{2w} + \frac{c}{2w} \ln [H^+] \qquad (2)$$

where $c = 2w(dZ/d \ln [H^+])$ is positive. We can now rewrite eq. (1) in terms of ln $[H^+]$ using eq. (I-5) and (2).

$$K_{obs} = (K_2^\circ e^{Z_0 + c \ln [H^+]})\left(\frac{1 + K_2^\circ e^{Z_0} + (c-1) \ln [H^+]}{1 + K_{Im} + K_2^\circ e^{Z_0} + (c-1) \ln [H^+]}\right) \qquad (3)$$

Differentiation of K_{obs} with respect to ln $[H^+]$ and setting the derivative equal to zero leads to a quadratic equation in $K_2^\circ e^{Z_0 + (c-1) \ln [H^+]}$ (which is equal to $K_2/[H^+]$), whose two solutions correspond to the maximum and minimum in the pK_{obs} *vs.* pH plot (see Fig. 3, curve 3). However, $K_2/[H^+]$ must be real and positive. In order to fulfill this requirement we find that

$$c = 2w \frac{dZ}{d \ln [H^+]} < \frac{1}{2} - \frac{1}{2\sqrt{1 + K_{Im}}} \qquad (4)$$

or, in terms of pH

$$\frac{2w}{2.303} \frac{dZ}{dpH} > \frac{1}{2\sqrt{1 + K_{Im}}} - \frac{1}{2} \qquad (5)$$

This is eq. (I-27).

Appendix IV

Complex Competition.—The protein species involved in complex competition between two donors and two acceptors (subject to the conditions given in the text) are given below together with their concentration fractions x_i, tabulated as $x_i y$ where

$$y = (1 + 3K_{ij} + K_{ij}^2 + 2K_1/[H^+] + 3K_{ij}K_1/[H^+] + K_1^2/[H^+]^2)$$

The K_1's and K_{ij}'s are assumed to be the same for both donors.

	Species	$x_i y$
(A)	D_1H ... A_1 / D_2H ... A_2	1
(B)	D_1H... A_1 / D_2H ... A_2	K_{ij}
(C)	D_1H ... A_1 / D_2H... A_2	K_{ij}
(D)	D_1H ... A_1 / D_2H... A_2	K_{ij}
(E)	D_1H... A_1 / D_2H... A_2	K_{ij}^2
(F)	D_1 ... A_1 / D_2H ... A_2	$K_1/[H^+]$

	Series	x_{1y}
(G)	D_1 $D_2H\cdot\cdot\cdot A_1$ A_2	$K_{1j}K_1/[H^+]$
(H)	D_1 $D_2H\cdot\cdot\cdot A_1$ $\cdot\cdot A_2$	$K_{1j}K_1/[H^+]$
(I)	D_1H D_2 A_1 A_2	$K_1/[H^+]$
(J)	$D_1H\cdot\cdot\cdot A_1$ D_2 A_2	$K_{1j}K_1/[H^+]$
(K)	D_1 D_2 A_1 A_2	$(K_1/[H^+])^2$

From these, eq. (I-33) and (I-34) of the text are derivable in a manner similar to that shown in Appendix I.

NOTES

Metal–Amine Coördination Compounds.[1] II. The Iron(III)–1,10-Phenanthroline System in Glacial Acetic Acid[2]

By Warren W. Brandt and Wilfred B. Howsmon, Jr.

Received June 21, 1954

The nature and uses of the ferrous–1,10-phenanthroline complex (ferroin) have been studied and are reviewed by Smith and Richter.[3] The corresponding ferric system (ferriin) is, however, not well understood. When ferroin is oxidized a blue complex results,[4] $Fe(phen)_3^{+++}$. This species is not obtained when ferric ion is mixed directly with 1,10-phenanthroline. Instead, a brown precipitate having the formula $[Fe_2(phen)_4(OH)_2]Cl_4$ is reported by Gaines, Hammett and Walden.[5] Manning and Harvey[6] have reported the formation of a soluble complex, $Fe_2(phen)_3$, under the same conditions. In addition, Simon and Knauer[7] obtained a precipitate, $Fe(phen)Cl_3$, by mixing ferric ion with 1,10-phenanthroline in both methanol and diethyl ether.

In the course of the present investigation, the reactions of chelating compounds containing the =N—C—C—N= grouping with various metallic ions in glacial acetic acid were studied. The purpose of this work was to determine the effects of this non-aqueous solvent on such chelation reactions. When 1,10-phenanthroline is mixed with ferric ion in this medium, a yellow complex is formed, the formula of which has been determined to be $Fe(phen)^{+3}$. A crystalline salt, Fe phen Cl_3, is precipitated from solutions in which $[Fe^{+++}]$ exceeds 2.5×10^{-4}. When this behavior is contrasted with the analogous reaction in water, as mentioned above, it appears that the solvent effect is one of preventing the chelation of more than one 1,10-phenanthroline molecule. If this were due to coördination of the solvent, the latter would be expected to be present in the precipitate also. In this regard, it should be mentioned that the chelation of Fe(II) is apparently not altered in this solvent. A red complex is obtained, and the spectrophotometric curve is identical with the aqueous complex. The data obtained will be useful in future work on 1,10-phenanthroline systems, both from the standpoint of their analytical applications and for studies of their physical nature.

Apparatus and Reagents.—All absorbancy measurements were made on a Beckman model DU spectrophotometer. Matched one-centimeter cells were used. 1,10-Phenanthroline monohydrate was obtained from the G. Frederick Smith Chemical Company. Reagent grade glacial acetic acid and ferric chloride hexahydrate were used.

Experimental Work.—The complex forms immediately when ferric ion in glacial acetic acid is added to 1,10-phenanthroline in the same solvent. It is stable for several days. It has a wave length of maximum absorbance at 400 mμ. The blank contained the same concentration of iron as the sample but no 1,10-phenanthroline, since the absorbancy of the latter is negligible at the wave lengths used.

The formula for the complex was established by two methods, the molar ratio method of Yoe and Jones[8] and Job's method of continuous variations.[9] The result of the molar ratio method is shown in Fig. 1. For the method of continuous variations the blank was a solution containing ferric ion of a concentration equal to that in the sample in which the iron was most dilute. It was necessary to correct the absorbancy readings for the amount of free ferric ion in those solutions not cancelled by the blank.

If the concentration of iron is made greater than 2.5×10^{-4} M, an orange-yellow precipitate forms. When the solution is filtered and absorbancy measurements taken, it is found that the wave length of maximum absorbancy has shifted to about 425 mμ. This precipitate was analyzed for iron and chlorine with the results shown in Table I. The analysis shows the formula to be $Fe(phen)Cl_3$. This corresponds to the compound prepared by Simon and Knauer[7] in non-aqueous media. It is possible that this compound could be the monomer as represented by the formula, in which case iron has a coördination number of 5, or it could

(1) Previous article in the series, R. T. Pflaum and W. W. Brandt, This Journal, **76**, 6215 (1954).

(2) Taken from the thesis presented by Wilfred B. Howsmon, Jr., to the Graduate School of Purdue University in partial fulfillment of the requirements for the degree of Master of Science, June, 1952.

(3) G. F. Smith and F. P. Richter, "Phenanthroline and Substituted Phenanthroline Indicators," The G. Frederick Smith Chemical Co., Columbus, Ohio, 1944.

(4) G. H. Walden, Jr., L. P. Hammett and R. P. Chapman, This Journal, **55**, 2649 (1933).

(5) A. Gaines, Jr., L. P. Hammett and G. H. Walden, Jr., *ibid.*, **58**, 3908 (1936).

(6) D. L. Manning and A. E. Harvey, Jr., *ibid.*, **74**, 4744 (1952).

(7) A. Simon and H. Knauer, *Z. Elektrochem.*, **45**, 678 (1939).

(8) J. H. Yoe and A. L. Jones, *Ind. Eng. Chem., Anal. Ed.*, **16**, 111 (1944).

(9) P. Job, *Ann. chim.*, [10] **9**, 113 (1928).