Polymerization-Depolymerization of Tobacco Mosaic Virus Protein. II. Theory of Protein Hydration*

MAX A. LAUFFER

From the Department of Biophysics,† University of Pittsburgh, Pittsburgh 13, Pennsylvania

Received January 21, 1964

To pave the way for demonstrating experimentally that water is released when tobacco mosaic virus protein is polymerized endothermically, it is necessary to formulate a definition of hydration amenable to exact thermodynamic analysis. For a three-component system, solvent, solute (low molecular weight), and colloid, "hydrate" is defined as that amount of solvent which, when added simultaneously with anhydrous colloid, produces identical changes in the logarithms of the activities of solvent and solute. This operation is equivalent to adding "hydrated" colloid. Upon adding hydrated colloid thus defined, the changes in the logarithms of the mole fractions and of the activity coefficients are also identical for, respectively, "free" solvent and solute on the "free" basis. This treatment is extended to multicomponent systems by taking into account "binding" of other constituents. If the colloid is osmotically ideal, the activity coefficients of "free" solvent and of solutes on the "free" basis are identical with those of solvent and solutes before addition of "hydrated" colloid, and even for concentrated osmotically nonideal solutions they differ only very slightly. With this definition of "hydration" one can obtain with great accuracy the relationship between mole fraction and activity for all components on the "free" basis in a multicomponent system. It is therefore possible to solve with precision for the distribution of all components at equilibrium on the two sides of an ideal membrane impermeable only to colloid. Thus, one can measure "hydration" by weighing at equilibrium a protein solution in an ideally semipermeable sack suspended in a multicomponent solution.

I. INTRODUCTION

It was shown by Lauffer *et al.* (1958) that the polymerization of tobacco mosaic virus protein (A protein) into rodlike polymers resembling original tobacco mosaic virus (TMV)[1] rods is an endothermic process, and it was postulated that the increase in entropy required for this reaction to be spontaneous was derived from the liberation of water. There are many endothermic processes in biology and the possibility exists that the same explanation can account for many or all of them. It is therefore of the utmost urgency that experimental methods be devised to determine whether or not water molecules are, indeed, released. C. L. Stevens and M. A. Lauffer (paper in preparation) have carried out an experiment in which the polymerization takes place inside a cellophane sack suspended by a delicate quartz helix in a medium containing buffer, glycerol, and water. The approximate idea behind this experiment is that buffer, glycerol, and water will come to equilibrium and therefore will not be weighed, but that protein and any water associated with it will be weighed, corrected for buoyancy (Lauffer, 1962; Stevens, 1962). If water is released during polymerization, it will then come to equilibrium and will no longer be weighed. This apparatus, therefore, is capable in principle of determining directly the amount of water released during this endothermic polymerization process.

There are, however, many complications involved in such an experiment. In order to determine precisely what it is one weighs, it is necessary to begin with a definition of hydration and with an exact treatment of the problem of membrane equilibrium involving multicomponent systems, and this must be followed by a precise analysis of the way in which each of the components at equilibrium affects the weight of the whole system.

II. DEFINITION OF HYDRATION

Attempts have been made in the past to differentiate in terms of thermodynamic concepts between hydrate or "bound" solvent and "free" solvent in systems containing proteins or other colloidal solutes. In general, total solvent has been considered as the sum of "free" solvent with the same solvent action, vapor pressure, etc. as pure solvent, and the remaining solvent defined as being "bound." The experimental work outlined in the preceding paragraphs has forced us to seek a definition of hydration amenable to exact thermodynamic treatment. As the subsequent argument will show, we have found one which differs slightly in a quantitative sense but importantly in a conceptual sense from previous thermodynamic definitions and which has the advantage of leading to very simple relationships between the "free" solvent in a system containing "hydrated" colloid and solvent in a system which does not contain such colloid.

Our definition of hydration can be stated most simply in reference to a three-component system; solvent, colloid, and (low-molecular-weight) solute. For such a system we define "hydrate" solvent as that amount of solvent which must be added along with anhydrous colloid to a solution so that the *ratio* of the activity of the solvent in the colloidal solution to that in the original solution is equal to the *ratio* of the activity of the solute in the colloidal solution to that in the original solution. It must be emphasized that this definition does not involve equality of activities before and after. The activities of both solvent and low-molecular-weight solute are changed by adding "hydrated" colloid, but our definition involves equal proportional change of both. It can be proved rigorously that, when this definition obtains, the *activity coefficients*

* The first paper in this series was Ansevin, A. T., and Lauffer, M. A. (1963), *Biophys. J. 3*, 239.

† Publication No. 104 of the Department of Biophysics. Work was supported by a U. S. Public Health Service grant (913-425-360).

[1] Abbreviation used in this work: TMV, tobacco mosaic virus.

of "free" solvent and of solute are changed proportionately the same by addition of "hydrated" colloid, that in all cases this change is extremely small, and that in an important special case it is exactly zero. In the following paragraphs this definition is spelled out mathematically and these claims are justified by precise thermodynamic deductions. The treatment has been extended to multicomponent systems, but in this case the "binding" of other constituents must be taken into account (Lauffer, 1963a).

A. Consider two vessels, I and II, each initially containing n_1 moles of 1 (water), n_3, n_4, etc. moles of solutes 3, 4, etc. Let single primes refer to vessel I and double primes to vessel II. The mole fraction, N, the activity, a, and the activity coefficient, f, of each constituent will be identical in the two vessels when they are at the same temperature and pressure. One can therefore write equation (1):

$$\frac{a'_{1a}}{a''_1} = \frac{a'_{3a}}{a''_3} = \frac{a'_{4a}}{a''_4} = \frac{a'_{ia}}{a''_i} \qquad (1)$$

The numerical subscripts refer to the various components and the alphabetical subscripts refer to the conditions described in the paragraphs designated by the same letter in capitals.

B. Now add n_2 moles of pure colloidal solute, 2 (anhydrous protein), to vessel I but keep vessel II constant. From the definition of mole fraction, one can write equation 2:

$$\frac{n_1 + n_3 + n_4 + \ldots}{n_1 + n_2 + n_3 + n_4 + \ldots} = \frac{N'_1}{N''_1} = \frac{N'_3}{N''_3} = \frac{N'_4}{N''_4} = \frac{N'_i}{N''_i} \qquad (i \gtrless 2) \qquad (2)$$

In the usual case, component 2 will affect differently the relative escaping tendencies of the several components, 1, 3, 4, etc., and thereby destroy the activity ratio equality expressed by equation (1). This inequality can be expresed by nonequations (3).

$$\frac{a'_{1b}}{a''_1} \gtrless \frac{a'_{3b}}{a''_3} \gtrless \frac{a'_{4b}}{a''_4}, \quad \text{etc.}$$
$$\frac{f'_{1b}N'_1}{f''_1N''_1} \gtrless \frac{f'_{3b}N'_3}{f''_3N''_3} \gtrless \frac{f'_{4b}N'_4}{f''_4N''_4}, \quad \text{etc.} \qquad (3)$$

C. The following argument can be presented best by first considering three components, 1, 2, and 3, and by then extending to multicomponents. If, in the three-component case, the amount of component 3 is held constant in vessel I and the amount of component 1 is altered, and if f'_1 and f'_3 are never zero or infinity and are both single valued for every value of $(n_1 + \Delta n_1)$, a value Δn_1 moles (positive or negative) can be found which will restore the activity ratios for components 1 and 3 to equality. This can be expressed by the central equality of equation (4).

$$\frac{f'_1N'_1}{f''_1N''_1} \equiv \frac{f'_{1c}N'_{1c}}{f''_1N''_1} \equiv \frac{a'_1}{a''_1} = \frac{a'_3}{a''_3} \equiv \frac{f'_{3c}N'_{3c}}{f''_3N''_3} \equiv \frac{f'_3N'_3}{f''_3N''_3} \qquad (4)$$

In this equation, N'_{1c} and N'_{3c} are defined in terms of the total amounts of all three components of the system ($n_1 + \Delta n_1$, n_2, and n_3). The total amount of solvent in vessel I is $n_1 + \Delta n_1$ moles. If we now define Δn_1 as "hydrate" solvent, one can describe the contents of vessel I by saying it contains n_1 moles of "free" solvent, n_2 moles of "hydrated" colloid, and n_3 moles of component 3. In terms of this definition or this mode of description, the mole fraction of "free" solvent in vessel I is therefore $n_1/(n_1 + n_2 + n_3)$, which is exactly equal to N'_1 of paragraph *A*. Similarly, the mole fraction of component 3 in vessel I computed on the "free"-solvent basis is exactly equal to N'_3 of paragraph *A*. The activity coefficients f'_1 and f'_3 of equation (4) are those which, when multiplied by the respective mole fractions calculated on the "free"-solvent basis, give the actual activities of components 1 and 3 in vessel I after the equality of activity ratios has been restored. They refer to the "free" state. From equation (2) and equation (4) one can derive the first equality in equation (5).

$$\frac{f'_1}{f''_1} = \frac{f'_3}{f''_3} = \frac{f'_4}{f''_4}, \quad \text{etc.} \qquad (5)$$

It is appropriate to make several observations at this point. Equations (1), (2), (4), and (5) are all exact. "Hydrate" has been defined in terms of operations; absolutely no model for a "hydrated" particle is involved. The definition of "hydrate" satisfies both equation (4) and equation (5); it is sufficient, therefore, to specify that either be satisfied and it is redundant to specify that both be.

Equation (4) could have been satisfied by adding both component 1 and component 3 to vessel I. An infinite number of combinations is possible. By inspection of equation (4) it can be seen that when specific values for Δn_1 and Δn_3 which satisfy this equation are available, within the range over which f'_{1c} and f'_{3c} are constant, all other values of Δn_1 and Δn_3 which are arrived at by merely adding additional amounts of components 1 and 3 to vessel I in the proportion in which they are already present will also satisfy the equation. A convenient set of solutions is the one in which either Δn_1 or Δn_3 is zero and the other is positive. When the solution is protein and some relatively inert solute such as glycerol or sucrose in water, it is frequently found that Δn_1 is positive when Δn_3 is zero.

In multicomponent systems one can restore the inequalities in component 3 to equalities by adding Δn_1 moles of component 1, Δn_3 moles of 3, Δn_4 moles of 4, etc. When one adheres to the same definitions as above, equation (5) is obtained in its full or general form. In this case also it is convenient to choose the set of Δn_1 values in which one is zero and the others are all positive.

III. Membrane Equilibrium

D. We now have in vessel I n_2 moles of colloid which has "bound" Δn_1, Δn_3, Δn_4, etc. moles of 1, 3, 4, etc., respectively (one of the Δn_i values is zero and the others are all positive), plus n_1, n_3, n_4, etc. moles of "free" components 1, 3, 4, etc. Everything has been adjusted so that equation (4), and therefore equation (5), has been satisfied. We now close vessels I and II, equip each with pistons so that alterations in pressure and volume can be made, and connect them with an ideal semipermeable membrane impermeable only to component 2. According to Gibbs, the condition for equilibrium is $\mu'_2 \gtrless \mu''_2$, $\mu'_i = \mu''_i$, when $i \gtrless 2$, $T' = T''$, and $P' \gtrless P''$. Equation (6) expresses the relationship between μ_i, P, and a_i.

$$d\mu_i = \bar{V}_i\, dP + RT\, d \ln a_i \qquad (6)$$

In this equation, μ_i is the chemical potential of ith constituent, P is the pressure, $\bar{V}_i$ is the partial molal volume of the ith constituent, R is the gas constant, and T is the absolute temperature. At equilibrium, equation (7) and therefore equation (8) must be obeyed by every constituent except component 2. In equation (8),

$$\int_{P''}^{P'} \bar{V}_i\, dP + RT \int_{a''_i}^{a'_i} d \ln a_i = 0 \qquad (7)$$

$$\ln \frac{a''_{id}}{a'_{id}} = \frac{(P' - P'')[\bar{V}_i]}{RT} \equiv h_i \qquad (8)$$

$[\bar{V}_i]$ is an average partial molal volume equal to $\int_{P''}^{P'} \bar{V}_i\, dP/(P' - P'')$. When attention is focused on solvent (component 1) and any one solute, j, other than colloid, equations (9), (10), (11) follow directly from equation 8. In equation (11) m_j is the molality

$$\frac{a'_{jd}}{a'_{1d}} = \frac{a''_{jd}}{a''_{1d}} e^{(h_1 - h_j)} \tag{9}$$

$$\frac{N'_{jd}}{N'_{1d}} = \frac{f'_{1d} f''_{jd}}{f'_{jd} f''_{1d}} \frac{N''_{jd}}{N''_{1d}} e^{(h_1 - h_j)} \tag{10}$$

$$m'_j = \frac{f'_{1d} f''_{jd}}{f'_{jd} f''_{1d}} m''_j e^{(h_1 - h_j)} \tag{11}$$

of the jth component on the "free" basis. The exponent in these three equations is equal to the real colloidal osmotic pressure multiplied by the difference between the average partial molal volumes of solvent and component j, all divided by RT. It is relatively easy to evaluate this term; therefore equation (9) can be solved readily for any of the j components.

Equations (9), (10), and (11) were derived in this same manner for a ternary solution by Adair (1937). They are all exact. Adair, however, did not solve the problem of evaluating the ratio of activity coefficient products found in equations (10) and (11), and therefore did not solve for the ratio of the molalities of permeable solutes at equilibrium. From equation (5) it follows that, before equilibrium is established, when the definitions in paragraphs *C* are adhered to, these activity coefficient product ratios in equations (10) and (11) are identically equal to 1. During equilibration, however, there is a slight shift in the composition of the "free" portion, and to the extent that activity coefficients are concentration dependent there will also be slight shifts in activity coefficients. It will be shown in a subsequent paragraph that highly accurate evaluations of these activity terms can be obtained, at least in principle.

IV. Deductions from the Gibbs-Duhem Relationship

From equation (6) and the Gibbs-Duhem relationship, equations (12), one can derive equations (13) and (14) for a multicomponent system. Except where

$$\begin{aligned} \Sigma N_i\, d \ln a_i &= 0 \\ \Sigma N_i\, d \ln N_i &= 0 \\ \Sigma N_i\, d \ln f_i &= 0 \end{aligned} \tag{12}$$

$$n_2 RT\, d \ln a_2 = v\, dP \tag{13}$$

$$d \ln f_2 = (g - 1)\, d \ln N_2 + \frac{1}{n_2}\left(\int_0^{n_s} n_2\, d \ln N_2\right) dg \simeq (g - 1)\, d \ln N_2 + dg \tag{14}$$

otherwise indicated, these equations are exact. In equation (13), v is the volume of all components except "hydrated" 2. The definition of g in equation (14) is the ratio of the real pressure difference at equilibrium to the ideal osmotic pressure. These equations are exactly the same as the thoroughly familiar ones for binary solutions. Equation (14) shows that over the range of "hydrated" colloid concentration for which g does not deviate measurably from one, $d \ln f_2 = 0$.

Equation (5) shows that when dn_2 moles of colloid which have "bound" Δn_1, Δn_3, Δn_4, etc. moles of 1, 3, 4, etc. per n_2 moles of 2 are added to a solution containing n_1, n_3, n_4, etc. moles of "free" 1, 3, 4, etc., $d \ln f_1 = d \ln f_3 = d \ln f_4$, etc. Equation 12 can therefore be rewritten $(N_1 + N_3 + N_4 + \cdots)\, d \ln f_1 = -N_2\, d \ln f_2$. When the concentration of "hydrated" colloid is low enough that it behaves ideally with respect to osmotic pressure this latter term is zero, and therefore the derivatives of the logarithms of all the other activity coefficients are also zero. Therefore, for this case, the activity coefficient ratios of equation (5) are all equal to 1. This means that the *activity coefficient* of "free" components in the colloidal solution (vessel I) are the same as in the reference solution (vessel II).

For protein solutions too concentrated to be osmotically ideal, it is usually found that the second derivative of pressure with respect to N_2 is zero or positive. Therefore one can write, $(g - 1) \leq kN_2$, where k is evaluated for the maximum N_2 used. Equations (15) follow, therefore, from equation (14).

$$\begin{aligned} N_2\, d \ln f_2 \leq 2kN_2\, dN_2 &= 2(g - 1)\, dN_2 \\ \int_0^{N_2} N_2\, d \ln f_2 \leq kN_2^2 &= (g - 1)N_2 \end{aligned} \tag{15}$$

Even for the most concentrated solutions of the most nonideal proteins for which osmotic pressure data exist, g rarely exceeds 4 and N_2 rarely exceeds 10^{-4}. Thus the integral in equation (15) is, in general, $\leq 3 \times 10^{-4}$. Therefore in the general case one can say that the integral of $d \ln f_i$ for all constituents except 2 in going from vessel II to vessel I $\leq 3 \times 10^{-4}$. In the worst cases, the activity coefficients on the "free" basis of the various components in the protein solution differ from those in a solution without protein by 3 parts in 10,000 or less. This is usually beyond the precision of measurements of activity coefficients.

We are now in a position to attempt to solve equations (10) and (11) when the solution for equation (9) is known. From equations (12) and (15) one can derive equations (16). The second of the equations (16)

$$\begin{aligned} N_1\, d \ln a_1 + N_3\, d \ln a_3 + N_4\, d \ln a_4 + \cdots &= -N_2\, d \ln a_2 \leq -(2g - 1)\, dN_2 \\ N_1\, d \ln N_1 + N_3\, d \ln N_3 + N_4\, d \ln N_4 + \cdots &= -N_2\, d \ln N_2 = -dN_2 \\ N_1\, d \ln f_1 + N_3\, d \ln f_3 + N_4\, d \ln f_4 + \cdots &= -N_2\, d \ln f_2 \leq -(2g - 2)\, dN_2 \end{aligned} \tag{16}$$

can be transformed into equation (17). It can be

$$\frac{dN_1}{dN_1} + \frac{dN_3}{dN_1} + \frac{dN_4}{dN_1} + \cdots = -\frac{dN_2}{dN_1} \simeq 0 \tag{17}$$

shown that $dN_2/dN_1 = n_2/n_1$. For a hypothetical 1% solution of protein with usual molecular weight and a value of "g" not greater than 1.5, dN_2/dN_1 must be of the order of magnitude of 10^{-5}. The first term on the left of equation (17) is +1 and the sum of the other terms on the left is approximately −1. It is therefore a reasonable approximation to set the right-hand member of equation (17), whose actual value rarely exceeds 10^{-5}, equal to zero. This is also true of the second of equations (16) and, since the first and third are, in the example chosen, of the same order of magnitude, they are also almost exactly equal to zero. In this form equations (16) are the same as for a solution without "hydrated" colloid but with composition equal to the composition of "free" constituents of the protein solution. For the ternary system, 1, 2, 3, equation (16) reduces to approximately that of the binary system, 1, 3. In this case it is very easy to evaluate the change in N associated with a change in activity from data such as freezing-point depressions for the binary solution. Such data are available for many solutes in water; for example, glycerol. One can therefore readily evaluate equations (10) and (11) from equation (9). The solution of the equations is more difficult when additional permeable components are present, but it is possible, in principle, to obtain the

requisite data by studies on solutions of the permeable components. For example, if components 1, 3, and 4 were water, glycerol, and sodium chloride, respectively, one could obtain the required information from combining vapor-pressure or freezing-point-depression data with electrochemical data. If the ionic strength of the electrolyte is 0.1 or less, it ought to be reasonably satisfactory to evaluate $d \ln f_4$ from the Debye-Huckel theory.

V. The Buoyant Weight of "Hydrated" Macromolecules

Consider a "weightless" container, initially with impermeable walls, of volume v', filled with a solution containing components 1 (water), 3 (such as glycerol), and 4 (such as sodium chloride), and suspended in a much larger solution of the same composition. Now add to the content of the container γ_2 g of component 2 (such as Na,Pr) in the pure "anhydrous" state. A volume of solution (1, 3, and 4) will have to be removed from the container exactly equal to $\gamma_2(V_2)$. (V_i) refers to the average partial specific volume, $\int V_i\, d\gamma_i / \int d\gamma_i$, of the ith component. Therefore the displaced volume is $\gamma_2(V_2)$, and the displaced weight is $\gamma_2(V_2)\rho$, where ρ is the density of the solution, 1, 3, 4. This term encompasses all volume changes resulting from introducing 2, the volume "occupied" by the molecules of 2, any contraction or expansion of any other components of the solution when "bound" by the molecules of 2, and any volume changes which take place at a distance as a result of introducing 2. As has already been explained, this introduction of "pure anhydrous" component 2 will create the inequality in activity-coefficient ratios described by inequality (3). The activity ratios and therefore the activity-coefficient ratios expressed in equation (5) can be restored to equality by adding Δn_1 moles of 1, Δn_3 moles of 3, and Δn_4 moles of 4. For convenience we will follow the convention explained previously and adopt the set with $\Delta n_3 = 0$. Accordingly, in order to satisfy equation (5) we must add Δn_1 moles or $G\gamma_2$ g of component 1 and Δn_4 moles or $G'\gamma_2$ g of component 4 to the container. Since the volume, v', must remain constant, $G\gamma_2(V_1) + G'\gamma_2(V_4)$ ml must be displaced from the container. G refers to the weight of "hydrate" solvent and G' to the weight of "bound" component 4 per gram of 2. The net or buoyant weight of the content is given by equation (18).

$$W = \gamma_2 G[1 - (V_1)\rho] + \gamma_2 G'[1 - (V_4)\rho] + \gamma_2[1 - (V_2)\rho] \quad (18)$$

If the walls of the container are now transformed into ideal semipermeable membranes and are equipped with a piston to allow changes in pressure and volume without altering the property of weightlessness, adjustments in pressure and composition will take place to satisfy equations (9), (10), and (11). If the volume is adjusted to keep the amount of "free" component 1 constant at γ_1 g, there will be slight changes in the amounts of "free" component 3 and 4 to give final amounts inside the container of γ_3 and γ_4 g, respectively. These adjustments will contribute $(\gamma_3 - \Gamma_3)[1 - (V_3)\rho]$ and $(\gamma_4 - \Gamma_4)[1 - (V_4)\rho]$ to the net or buoyant weight, where Γ_3 and Γ_4 are the amounts of 3 and 4 respectively per γ_1 g component of 1 in the external solution and originally in the internal solution. The final net weight is given by equation (19).

$$W = \gamma_2 G[1 - (V_1)\rho] + \gamma_2 G'[1 - (V_4)\rho] + \gamma_2[1 - (V_2)\rho] + (\gamma_3 - \Gamma_3)[1 - (V_3)\rho] + (\gamma_4 - \Gamma_4)[1 - (V_4)\rho] \quad (19)$$

Equation (19) is exact for an incompressible system. In equation (19), $(\gamma_3 - \Gamma_3) = (M_3\gamma_1/1000)(m'_3 - m''_3)$. A formally identical relationship exists for component 4. Equation (11) can be written $m'_3 = H_3 m''_3$, where H_3 is defined as the remaining terms in equation (11). Thus $(\gamma_3 - \Gamma_3) = (M_3\gamma_1 m''_3/1000)(H_3 - 1)$.

The situation with respect to component 4 is more complex because we have taken it to be a uni-univalent electrolyte. The chemical potential, therefore, of component 4 is the sum of the chemical potentials of the positive ion and of the negative ion, and the condition for equilibrium as stated by Gibbs is that these two sums be equal for the two sides of the membrane. Equations (9), (10), and (11), therefore, take the form of equations (20), (21), and (22) for this special case.

$$\frac{a'_{4+}}{a'_1}\frac{a'_{4-}}{a'_1} = \frac{a''_{4+}}{a''_1}\frac{a''_{4-}}{a''_1} e^{(2h_1 - h_4)} \quad (20)$$

$$\frac{N'_{4+}N'_{4-}}{N'_1N'_1} = \frac{f'_1f'_1}{f'_{4+}f'_{4-}}\frac{f''_{4+}f''_{4-}}{f''_1f''_1}\frac{N''_{4+}N''_{4-}}{N''_1N''_1} e^{(2h_1 - h_4)} = H_4\frac{N''_{4+}N''_{4-}}{N''_1N''_1} \quad (21)$$

$$m'_{4+}m'_{4-} = H_4 m''_{4+}m''_{4-} \quad (22)$$

These three equations are exact. When $H_4 = 1$, i.e., when $P' - P'' = 0$, equation (22) reduces to the familiar approximate equation for Donnan equilibrium.

Because of the necessity to preserve zero net charge, m'_{4+} of equation (22) is equal to m'_4, the molality of sodium chloride inside the box, plus q, the molality of positive ion from the ionization of sodium proteinate. The molality of the negative ion inside the box is equal to m'_4, and the molalities of both the positive and the negative ions on the outside are equal to m''_4. When these values are substituted into equation (22) and the quadratic is solved for m'_4, one obtains equation (23).

$$m'_4 = \sqrt{H_4(m''_4)^2 + \frac{q^2}{4}} - \frac{q}{2} \quad (23)$$

Equation (24) then follows.

$$(\gamma_4 - \Gamma_4) = \frac{M_4\gamma_1}{1000}\left[\sqrt{H_4(m''_4)^2 + \frac{q^2}{4}} - \frac{q}{2} - m''_4\right] \quad (24)$$

Therefore, equation (19) can be transformed into equation (25).

$$W = \gamma_2 G[1 - (V_1)\rho] + \gamma_2 G'[1 - (V_4)\rho] + \gamma_2[1 - (V_2)\rho] + \frac{M_3\gamma_1}{1000} m''_3(H_3 - 1)[1 - (V_3)\rho] + \frac{M_4\gamma_1}{1000} m''_4\left[\sqrt{H_4 + \left(\frac{q}{2m''_4}\right)^2} - \frac{q}{2m''_4} - 1\right][1 - (V_4)\rho] \quad (25)$$

This is an exact equation for an incompressible system.

All of the terms on the right side of equation (25) except G and G' can be evaluated independently. When $\rho = 1/(V_1)$, G' can therefore be evaluated from a determination of the buoyant weight of the material inside the container. When the density of the solution is subsequently raised, G can be evaluated from W provided only that G' does not change appreciably. This is a likely situation if the density is changed primarily by altering the amount of component 3 in the system, keeping component 4 constant. In any event, G' is apt to be very much smaller than G, so that even if this assumption is not completely valid no great error in the evaluation of G results.

When one deals with an isoelectric protein, component 4 can be omitted entirely and equation (25)

becomes very much simpler. In this instance G can be evaluated directly from one determination of W. We have therefore, in principle, a simple method, solidly grounded on thermodynamics, for determining protein hydration as defined in section II (Lauffer, 1963b).

VI. Relationship to Other Methods of Determining Hydration

Newton and Gortner (1922) defined "bound" water as that fraction of the water in a system which is not available to act as a solvent for solutes added to the system. They measured this amount of water by determining the freezing-point depression of the sample containing colloidal material to which enough sucrose was added to make the solution molar in sucrose and compared this with the calculated value of the freezing-point depression if the total water in the solution had been available to dissolve sugar. Hill (1930) used an accurate vapor-pressure apparatus to study "bound" water content of muscle and blood. He defined "free" water as the weight of water in 1 g of fluid or tissue which can dissolve substances added to it with a normal depression of the vapor pressure. In both these methods the difference between "free" water and total water is taken to be "bound" or "hydrate" water. If, as a refinement of the method of Hill, one defines as hydrate, $\Delta' n_1$, that amount of solvent one must add to the colloid solution to make $d \ln a'_1/dn_3 = d \ln a''_1/dn_3$ or $d \ln (P'/P_0)/dn_3 = d \ln (P''/P_0)/dn_3$, it can be shown that this definition of hydration amounts to one solvent molecule per colloid particle less than our value, Δn_1.[2]

Briggs (1932) showed that the amount of "bound" water in several systems, whether determined by the method of Hill (1930) or by that of Newton and Gortner (1922), depended on the activity of the water in the system. The same relationship must hold for "hydrate" water determined by our method.

Bull (1944) carried out an extensive study on sorption of water vapor by many proteins when suspended in water vapor in equilibrium with solutions with various water activities. He found, in agreement with the observation of Briggs, that the amount of water taken up per unit weight of protein increased as the activity of the water increased. In effect, protein is separated from a solution of components 1, 3, 4, etc. by a membrane permeable only to water. The condition for equilibrium is that the activity of the water in the protein phase be equal to that in the solution.

For every case in which Δn_1^* moles of solvent are sorbed by n_2 moles of protein in equilibrium with the vapor of a solution containing n_1 moles of 1 and Σn_i moles of 3, 4, $\cdots$ j, equation (26) can be written, where $'''$ refers to the protein and sorbed vapor.

$$a'''_1 = \frac{f'''_1 \Delta n_1^*}{n_2 + \Delta n_1^*} = a''_1 = \frac{f''_1 n_1}{n_1 + \sum_3^j n_i} = \frac{f'''_1 \Delta n_1^* + f''_1 n_1}{n_2 + \Delta n_1^* + n_1 + \sum_3^j n_i} \quad (26)$$

[2] Add $\Delta' n_1$ moles of 1 to I to satisfy $d \ln a'_1/dn_3 = d \ln a''_1/dn_3$. Since $f'_1 = f''_1$, exactly for an osmotically ideal solution and almost exactly for all solutions, $a'_1 = (N'_1/N''_1)a''_1$ and $da'_1/dn_3 = (N'_1/N''_1)\, da''_1/dn_3 + a''_1\, d(N'_1/N''_1)/dn_3$. Therefore $d \ln a'_1/dn_3 = d \ln a''_1/dn_3 + (N''_1/N'_1)\, d(N'_1/N''_1)/dn_3$. Since $N'_1 = n'_1/(n'_1 + n_2 + n_3 + n_4$, etc.) and $N''_1 = n''_1/(n''_1 + n_3 + n_4$, etc.), $d \ln a'_1/dn_3 = d \ln a''_1/dn_3 + N''_1 n'_1(n'_1 + n_2 - n''_1)/[N'_1 n''_1(n'_1 + n_2 + n_3)^2]$. The final term is zero when $(n'_1 + n_2) = n''_1$. Thus, $d \ln a'_1/dn_3 = d \ln a''_1/dn_3$ when $n'_1 = (n''_1 - n_2) = (n_1 - n_2) \cdots \Delta' n_1 = \Delta n_1 - n_2$.

Equation (27) is an expression for the activity of solvent after the protein with its sorbed solvent is dissolved in the binary solution.

$$a'_1 = \frac{f'_{1c}(\Delta n_1^* + n_1)}{n_2 + \Delta n_1^* + n_1 + \sum_3^j n_i} \quad (27)$$

The activity coefficient is the one in which total solvent is considered. For the case in which the activity coefficient of total solvent after mixing is equal to the weighted mean of f'''_1 and f''_1, equation (27) becomes equal to equation (26). For this case, therefore, the Δn_1^* determined by a sorption experiment satisfies the requirement that $a'_1 = a''_1$.

For the multicomponent system, 1, 2, 3, $\cdots$ j, $\Delta n_1^* = \Delta n_1 + (n_1 n_2 / \sum_3^j n_i)$. Therefore, except for solutions in which $n_1/\sum_3^j n_i$ is very large, this definition of hydration approaches ours within a few solvent molecules per colloid particle. The equation just mentioned can be derived as follows. From equations (26) and (27), one can write:

$$a''_1 = \frac{f''_1 n_1}{n_1 + \sum_3^j n_i} = a'_1 = \frac{f'_{1c}(\Delta n_1^* + n_1)}{n_1 + \Delta n_1^* + n_2 + \sum_3^j n_i} = \frac{f'_1(n_1 + k)}{n_1 + \sum_3^j n_i + n_2 + k}$$

Since $f''_1 = f'_1$,

$$\frac{n_1}{n_1 + \sum_3^j n_i} = \frac{n_1 + k}{(n_1 + \sum_3^j n_i) + (n_2 + k)}$$

or

$$\frac{n_1}{n_1 + \sum_3^j n_i} = \frac{k}{n_2 + k}$$

$$k = \frac{n_1 n_2}{\sum_3^j n_i}$$

Therefore $\Delta n_1^* = \Delta n_1 + (n_1 n_2 / \sum_3^j n_i)$.

If it is assumed that the condition for zero sedimentation rate in an ultracentrifuge experiment is "weightlessness" of the "hydrated" colloidal component,[3] one can transform equation (19) into equation (28) for a system containing 1, 3, and "hydrated" 2 at infinite dilution.

$$W = 0 = \gamma_2 G[1 - (V_1)\rho_0] + \gamma_2[1 - (V_2)\rho_0] \quad (28)$$

For equation (28) to be completely applicable, the colloid would have to be uncharged. In this equation, ρ_0 is the density of binary solution 1, 3 corresponding to zero sedimentation rate. The fourth term on the right of equation (19) vanishes at infinite dilution of colloid because the colloidal osmotic pressure vanishes.

[3] This would be an exact and sufficient condition for zero sedimentation velocity in a hypothetical system containing at least three components in which the chemical potential of each component was independent of distance from the axis of rotation. It is impossible to realize this condition in an actual sedimentation velocity experiment. To the extent that the real situation is a sufficiently close approximation to the ideal unrealizable case, "weightlessness" is an appropriate criterion for nonsedimentation.

When equation (28) is solved for $1/\rho_0$, equation (29) is obtained.

$$V_h \equiv \frac{1}{\rho_0} = \frac{G(V_1) + (V_2)}{G + 1} \qquad (29)$$

Except for differences in symbols, this equation is identical with equation (6) of Lauffer and Bendet (1954) and equation (1) of Lauffer *et al.* (1952) for the case in which the "solvate" is component 1. There is an apparent difference in definitions of ρ, but this vanishes for the case of "weightlessness." This equation has been used by many investigators to determine the "hydration" of proteins from experiments involving measurement of sedimentation velocities in sucrose or glycerol solutions of different densities. Contingent only upon the limitation expressed above, it is evident that "hydration" determined by this method is the same as that defined in section II.

While there are to my knowledge no data in the literature on the "hydration" of TMV protein, there are data on TMV itself. Schachman and Lauffer (1949) studied the sedimentation of TMV in sucrose solutions of different densities and extrapolated their results to zero sedimentation rate. This occurred at a density of 1.27. From this result and equation (29) one can calculate that 0.27 g of water is "bound" per gram of TMV. From data in the International Critical Tables one can estimate that a sucrose solution with the density of 1.27 has a relative water vapor pressure of about 0.91. Katchman *et al.* (1950, 1951) reported data on the sorption of water vapor by dry TMV in equilibrium with water vapor at various partial pressures. They obtained somewhat different results in sorption and desorption experiments. However, from their results one can estimate that something between 0.21 and 0.30 g of water is "bound" by 1 g of TMV at a relative vapor pressure of 0.91. This result is in quite good agreement with the finding of Schachman and Lauffer. In the studies of Stevens and Lauffer, to be reported later, which were carried out in glycerol solutions with slightly higher water-vapor pressures, it was found that of the order of magnitude of one-tenth of this amount of water was released when TMV protein polymerized.

Acknowledgments

The author is deeply indebted to Professor Irving J. Klotz for extensive critical discussion of earlier versions of this article and for encouragement to refine the concepts presented. This acknowledgment of gratitude in no way implies that Dr. Klotz agrees with every detail of the present version or that he is in any way responsible for shortcomings which might later become apparent.

References

Adair, G. S. (1937), *Trans. Faraday Soc. 33*, 1106.
Briggs, D. R. (1932), *J. Phys. Chem. 36*, 367.
Bull, H. B. (1944), *J. Am. Chem. Soc. 66*, 1499.
Hill, A. V. (1930), *Proc. Roy. Soc. (London) Ser. B: 106*, 477.
Katchman, B., Cutler, J., and McLaren, A. D. (1950), *Nature 166*, 266.
Katchman, B., and McLaren, A. D. (1951), *J. Am. Chem. Soc. 73*, 2124.
Lauffer, M. A. (1962), *in* The Molecular Basis of Neoplasia, Austin, University of Texas Press, p. 180.
Lauffer, M. A. (1963a), *Biol. Bull. 125*, 356.
Lauffer, M. A. (1963b), *Biol. Bull. 125*, 357.
Lauffer, M. A., Ansevin, A. T., Cartwright, T. E., and Brinton, C. C., Jr. (1958), *Nature 181*, 1338.
Lauffer, M. A., and Bendet, I. J. (1954), *Advan. Virus Res. 2*, 241.
Lauffer, M. A., Taylor, N. W., and Wunder, C. C. (1952), *Arch. Biochem. Biophys. 40*, 453.
Newton, R., and Gortner, R. A. (1922), *Botan. Gaz. 74*, 442.
Schachman, H. K., and Lauffer, M. A. (1949), *J. Am. Chem. Soc. 71*, 536.
Stevens, C. L. (1962), Ph.D. dissertation, University of Pittsburgh.