Isopiestic Compositions as a Measure of Preferential Interactions of Macromolecules in Two-Component Solvents. Application to Proteins in Concentrated Aqueous Cesium Chloride and Guanidine Hydrochloride[1]

E. P. Kirby Hade[2] and Charles Tanford

Contribution from the Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27706. *Received April 21, 1967*

Abstract: Equilibrium isopiestic compositions of solutions which contain a volatile and a nonvolatile solvent component, and varying amounts of a macromolecular substance, directly yield values for the preferential binding of either solvent component to the macromolecules. It is shown that native serum albumin in 2.5 *m* CsCl is preferentially hydrated, in agreement with previous observations by other methods. Unfolded proteins in concentrated guanidine hydrochloride, on the other hand, preferentially bind guanidine hydrochloride, though only to a small extent. The relation of these results to the determination of protein molecular weights in concentrated guanidine hydrochloride is discussed.

In a solution which contains a macromolecular solute and two components of low molecular weight, each macromolecule will in general preferentially "bind" one or the other of the solvent components. We use the term "binding" here in a pragmatic sense, to include all solvent molecules which come within the sphere of influence of the macromolecule. Preference for one component or the other may result not only from the existence of specific binding sites for that component, but also from the exclusion of the other component from the vicinity of the macromolecule, as, for example, for steric reasons.[3]

The measurement of such preferential interactions is an interesting subject for its own sake. The present study was undertaken, however, primarily because preferential interactions influence the determination of the molecular weights of macromolecules in three-component systems, when sedimentation equilibrium, sedimentation velocity, or light scattering are used as the basis for the molecular weight measurements.[4-8] A measure of the preferential interactions is essential if the correct molecular weight of the unsolvated macromolecule is to be extracted from the data.

The most precise measurements of preferential interactions of this kind (at least for proteins and nucleic acids) have been based on the buoyant behavior of the macromolecules in the two-component solvent in the ultracentrifuge.[9-11] Other precedures which have been used have depended ultimately on redistribution of the solute components across a membrane impermeable to the macromolecule. For example, comparisons of the partial specific volume or refractive index increment, before and after such redistribution, have been employed.[12-15] As far as we are aware, the method used in the present paper, which is more direct and simpler to carry out than either of the foregoing procedures, has been used only once heretofore, in a study of the preferential hydration of DNA.[15b] A somewhat related use of vapor pressure measurements has been reported by Schormuller and Laskowski.[16]

This paper will briefly describe the theory which underlies the use of the method of isopiestic compositions for determination of preferential interactions. Results will be reported for the preferential interactions of a number of proteins in 6 *M* guanidine hydrochloride (GuHCl), this being a solvent system which we and others have frequently used for molecular weight measurements. As a control, results are also reported for the behavior of serum albumin in 2.5 *m* cesium chloride, since Ifft and Vinograd[11] have published accurate data for that system by measurements of buoyant behavior in the ultracentrifuge.

An additional reason for choosing 6 *M* GuHCl (which is 10 *m* on the molal scale) and 2.5 *m* CsCl as solvents for investigation is that proteins are random coils in 6 *M* GuHCl, with the polypeptide backbone and all side chains accessible to the solvent.[17-20] In 2.5 *m* CsCl, on the other hand, proteins are native, and only the generally hydrophilic surfaces of their compact structures are in contact with solvent. One would expect proteins to be preferentially hydrated in 2.5 *m*

(1) This work was supported by grants from the National Science Foundation and from the National Institutes of Health, U. S. Public Health Service.

(2) Fellow of The Arthritis Foundation.

(3) V. N. Schumaker and D. J. Cox, *J. Am. Chem. Soc.*, **83**, 2445 (1961).

(4) R. J. Goldberg, *J. Phys. Chem.*, 57, 194 (1953).

(5) J. W. Williams, K. E. Van Holde, R. L. Baldwin, and H. Fujita, *Chem. Rev.*, **58**, 715 (1958).

(6) R. L. Baldwin, *J. Am. Chem. Soc.*, 80, 496 (1958).

(7) G. Scatchard in "Ultracentrifugal Analysis," J. W. Williams, Ed., Academic Press Inc., New York, N. Y., 1963, p 105.

(8) E. F. Casassa and H. Eisenberg, *Advan. Protein Chem.*, **19**, 287 (1964).

(9) J. Vinograd and J. E. Hearst, *Fortschr. Chem. Org. Naturstoffe*, 20, 372 (1962).

(10) D. J. Cox and V. N. Schumaker, *J. Am. Chem. Soc.*, **83**, 2433, 2439 (1961).

(11) J. B. Ifft and J. E. Vinograd, *J. Phys. Chem.*, 70, 2814 (1966).

(12) A. Vrij and J. T. G. Overbeek, *J. Colloid Sci.*, 17, 570 (1962).

(13) E. F. Casassa and H. Eisenberg, *J. Phys. Chem.*, **65**, 427 (1961).

(14) W. W. Kielley and W. F. Harrington, *Biochim. Biophys. Acta*, **41**, 401 (1960).

(15) (a) M. E. Noelken, *Federation Proc.*, **26**, 777 (1967); (b) J. E. Hearst, *Biopolymers*, **3**, 57 (1965).

(16) M. L. Schormuller and M. Laskowski, Jr., *Federation Proc.*, **23**, 473 (1964).

(17) C. Tanford, K. Kawahara, and S. Lapanje, *J. Am. Chem. Soc.*, **89**, 729 (1967).

(18) Y. Nozaki and C. Tanford, *ibid.*, **89**, 742 (1967).

(19) C. Tanford, K. Kawahara, S. Lapanje, T. M. Hooker, Jr., M. H. Zarlengo, A. Salahuddin, K. C. Aune, and T. Takagi, *ibid.*, **89**, 5023 (1967).

(20) S. Lapanje and C. Tanford, **89**, 5030 (1967).

CsCl, but at least a tendency in the opposite direction is expected in salt solutions in which the native form becomes unstable, by virtue of the principle of linked functions and reciprocal effects.[21]

Theory

We shall adopt the convention of Scatchard,[22a] and use odd-numbered subscripts (1 and 3) to designate solvent components, and an even-numbered subscript (2) for the macromolecular component. The method of isopiestic compositions is applicable only when only one of the solvent components is volatile. We shall designate this component as component 1.

Preferential interactions are commonly expressed in terms of appropriate partial derivatives.[22b] If a solvent component is preferentially bound, its chemical potential in the solution will fall as the concentration of component 2 is increased. Preferential binding may thus be measured in terms of the amount of the solvent component which must be added to the solution to keep the chemical potential constant. If component 1 is taken as the principal solvent, and concentrations are expressed as moles/kilogram of component 1 (m_i) or as grams/kilogram of component 1 (g_i), the preferential binding of component 3 may be expressed in terms of the derivatives $(\partial m_3/\partial m_2)_{\mu_3}$ or $(\partial g_3/\partial g_2)_{\mu_3}$

$$\left(\frac{\partial g_3}{\partial g_2}\right)_{\mu_3} = \frac{M_3}{M_2}\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_3} \tag{1}$$

where M_i is the formula weight of a component.

We shall be interested in obtaining values for either of the derivatives of eq 1, or the related quantities to be considered below, at the limit of infinite dilution of the macromolecular component, *i.e.*, as $m_2 \rightarrow 0$. The reason for this is that we are interested in the interaction of an isolated macromolecule with the solvent, uncomplicated by interactions of macromolecules with each other. The molecular weight equations to which these studies will be applied are also valid only at the limit of $m_2 \rightarrow 0$.

If the derivatives of eq 1 are negative, it means that component 3 must be removed from the solution to maintain its chemical potential constant, and this implies that component 1 is being preferentially bound. To obtain a direct expression for the binding of component 1, we employ the device of considering component 3 as the principal solvent. Where m_i' and g_i' represent amounts of the other components per kilogram of component 3, we have, in analogy to eq 1

$$\left(\frac{\partial g_1'}{\partial g_2'}\right)_{\mu_1} = \frac{M_1}{M_2}\left(\frac{\partial m_1'}{\partial m_2'}\right) \tag{2}$$

Since ratios of the quantities of each component must be the same, regardless of the manner in which the composition is expressed, we have

$$(1000/M_1):m_2:m_3 = m_1':m_2':(1000/M_3) \tag{3}$$

(21) J. Wyman, *Advan. Protein Chem.*, 19, 223 (1964).

(22) (a) G. Scatchard, *J. Am. Chem. Soc.*, 68, 2315 (1946). (b) The derivatives used in this paper contain an ideal contribution, which is however negligibly small in comparison with the contribution ascribable to preferential interactions. For a thermodynamically ideal aqueous solution (water activity = mole fraction), $(\partial g_3/\partial g_2)_{\mu_1} = -\nu_2 M_3/\nu_3 M_2$, where ν_2 and ν_3 are the moles of particles per mole of components 2 and 3, respectively. With $\nu_2 = 1$, $\nu_3 = 2$, $M_3 = 100$, and $M_2 = 10^4$–10^5, $(\partial g_3/\partial g_2)_{\mu_1} = -5 \times 10^{-3}$ to -5×10^{-4}, which is smaller than the experimental error in the determination of the derivative.

This equation may be used to obtain a relation between the derivatives of eq 1 and those of eq 2. We first obtain directly from eq 3

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_1} = -\frac{\dfrac{M_1 m_3}{1000}\left(\dfrac{\partial m_1'}{\partial m_2'}\right)_{\mu_1}}{1 - \dfrac{M_1 m_2}{1000}\left(\dfrac{\partial m_1'}{\partial m_2'}\right)_{\mu_1}} \tag{4}$$

Since we are interested in the derivatives of eq 1 and 2 only as $m_2 \rightarrow 0$, we can delete the second term in the denominator of eq 4. Actually this term is probably always negligible compared to unity, even at finite values of m_2, because m_2 must be very small at any reasonable working concentration if component 2 is a macromolecule. Thus we can quite generally replace eq 4 by

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_1} = -\frac{M_1 m_3}{1000}\left(\frac{\partial m_1'}{\partial m_2'}\right)_{\mu_1} \tag{5}$$

It is easy to obtain a relation between $(\partial m_3/\partial m_2)_{\mu_1}$ and $(\partial m_3/\partial m_2)_{\mu_3}$. This can be done, for example, by writing

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_1} = -\left(\frac{\partial \mu_1}{\partial m_2}\right)_{m_3} \Big/ \left(\frac{\partial \mu_1}{\partial m_3}\right)_{m_2} \tag{6}$$

and then applying the Gibbs–Duhem equation

$$(1000/M_1)\mathrm{d}\mu_1 + m_2\mathrm{d}\mu_2 + m_3\mathrm{d}\mu_3 = 0 \tag{7}$$

to both numerator and denominator. The general result is cumbersome, but, for the limit of $m_2 \rightarrow 0$, it reduces to the simple expression

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_3} = \left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_1} + \frac{1}{1 + m_3(\partial \ln \gamma_3/\partial m_3)_{m_2=0}} \tag{8}$$

where γ_3 is the activity coefficient of component 3 on the molal scale. Equations 5 and 8 can be combined to convert a *negative* derivative for preferential binding of one component to a *positive* derivative for preferential binding of the other.

The method of isopiestic compositions provides a direct measure of the foregoing derivatives. The essential feature of the method is to allow a number of solutions of known initial composition, all containing the same single volatile solvent component, to come to equilibrium with each other with respect to the activity of that component. The equilibrium composition is established simply by weighing, as changes in composition can result only from transfer of the volatile component. To use the method for preferential interaction studies one uses a series of solutions, all containing component 3 and varying amounts of component 2. The equilibrium compositions measure directly the change in m_3 or m_1' (depending on one's choice of principal solvent), as a function of m_2, at constant μ_1, *i.e.*, they are a direct measure of $(\partial m_1'/\partial m_2')_{\mu_1}$ or $(\partial m_3/\partial m_2)_{\mu_1}$. The latter derivative can be converted to $(\partial m_3/\partial m_2)_{\mu_3}$ by means of eq 8. The desired quantities, as given by eq 1 and 2, are thus obtained without performing any operation other than to weigh out reagents, to obtain the initial composition of each test solution, and a final weighing to obtain the equilibrium composition.

It should be pointed out that the term $(\partial \ln \gamma_3/\partial m_3)_{m_2=0}$ which occurs in eq 8, and which must be known in order to express the final results in terms of the derivatives of eq 1, depends only on the activity coefficients of solutions of component 3 in component 1,

in the absence of component 2, *i.e.*, this term will normally be a known quantity for solutions of interest, such as CsCl–H_2O or sucrose–H_2O. The term is normally quite small, in any case, and the difference between $(\partial m_3/\partial m_2)_{\mu_1}$ and $(\partial m_3/\partial m_2)_{\mu_3}$, in the experiments to be described here, is within the experimental error of the measurements.

Ionic Nature of Third Component. In the experiments to be reported in this paper, the third component is a 1+,1− electrolyte, present at high concentration. No modification in the theoretical treatment is required. The radius of the ionic atmosphere is very small at high ionic strength. Even if one ion of the electrolyte is specifically bound to the macromolecule, and no intrinsic attraction for the other exists, the other ion will necessarily be in close proximity, and with be physically indistinguishable from a tightly bound counterion. The isopiestic method is simply a measure of the excess concentration of the natural third component in the vicinity of the macromolecule, and can give no information regarding the chemical forces which lead it to be present at an excess concentration.

Definition of Macromolecular Component. In dealing with proteins, which may carry electrostatic charges, there is some choice in the manner in which the macromolecular component is defined.[22] In this paper the protein component is defined so as to consist of uncharged isoionic molecules, and we shall seek experimentally to add the protein component in that state. This is not always possible, however, and it is sometime necessary to add the protein component as a salt, containing protein ions with a small average net charge $(\pm Z)$. In that case Z univalent counterions[23] will be added to the solution along with each protein molecule and they will naturally exert an independent effect on water activity. As a rough estimate, one may consider the effect to be equal to that of $Z/2$ moles of component 3. If m_3 is the measured concentration of component 3, the *effective* concentration will thus be $m_3 + (Z/2)m_2$, and the quantity which truly expresses preferential salt binding or negative preferential hydration is $\partial m_3/\partial m_2 + Z/2$.

As Table IV will show, the experimental uncertainty in determining $\partial m_3/\partial m_2$ is approximately 1 mole of salt for each 10,000 g of molecular weight. Thus proteins may be treated as isoionic, within experimental error, as long as Z is less than ± 2 charges per 10,000 molecular weight.

Experimental Section

Materials. The following proteins were used: ribonuclease (type II A) and ovalbumin, purchased from Sigma Chemicals, Inc., aldolase, purchased from Boehringer Mannheim Corp., bovine serum albumin and egg white lysozyme, purchased from Pentex, Inc., and β-lactoglobulin, donated by Dr. R. Townend, of the Eastern Utilization Research and Development Division, U. S. Department of Agriculture.

The ribonuclease, serum albumin, and ovalbumin were desalted by first dialyzing a 4% solution against several changes of distilled water and then passing the dialyzed solution through a mixed-bed resin column prepared according to the procedure of Dintzis.[24] Some of the serum albumin was simply dialyzed exhaustively against several changes of glass-distilled water. The results obtained with the sample treated in this manner were identical with those obtained with the sample which was passed through the column.

The lysozyme was dissolved in glass-distilled water to make a 4% solution, and the β-lactoglobulin and aldolase were dissolved in 0.1 M NaCl. These solutions were put in acetylated dialysis bags and dialyzed for 72 hr against a constant flow of deionized glass-distilled water.

All of the proteins were lyophylized after desalting. The amount of water in each sample was measured by taking about 0.05 g of each protein to constant weight in a circulating air oven at 107°. Samples of the lyophilized proteins were put into solution at the time of measurement, and pH measurements were carried out. These measurements indicated that all of the proteins were at or close to their isoionic points, except lysozyme, which gave a pH of 5.5, indicative of a molecular charge of about +9.

The guanidine hydrochloride used in this study was prepared as previously described.[25] Before use it was dried in a continuously evacuated vacuum chamber over silica gel and phosphorus pentoxide. The cesium chloride was obtained from Harshaw Chemical Co. Before use it was dried at 300° for 24 hr.

Isopiestic Method. The isopiestic technique has been described by Robinson and Stokes.[26] The method is used for solutions in which only one component, usually water, is volatile. If aqueous solutions are put in a closed chamber and are in good thermal contact with one another they will reach concentrations at equilibrium in which all solutions have the same water activity. The technique is most often used to measure the water activity of a solution by comparing it to a standard with a known water activity. In the work described here we were concerned only with the change in concentration of one component, either salt or water, as a function of protein concentration. Therefore our reference vessels contained the same aqueous salt solution as was used to make up the protein samples.[27]

The apparatus consists of glass vacuum desiccators which were used as the isopiestic chambers. The glass sleeve was removed from the tops and a small brass tube with a short length of rubber hose attached to it was fastened into each inlet hole with epoxy cement. Each chamber contained a silver-plated copper block 15 cm in diameter and 2.5 cm thick. Twenty-one flat-bottomed holes, 2 cm deep and 2.8 cm in diameter were milled into each block. For isopiestic vessels we used scintillation vials which fit snugly into the holes in the copper block. To increase stirring of the solutions during the equilibration period each isopiestic vessel contained a small glass bead.

To prepare for a run, a salt solution of an accurately known concentration was made. Most of it was used to make protein solutions of various concentrations, and the remainder was used as a reference solution in the chamber. Each protein solution and the reference solution were put into two separate vessels. The solution put into each vessel contained about 5 g of water. The vessels were weighed, and four drops of water were put into one member of each pair of solutions so that each pair differed in concentration by about 4%. The isopiestic vessels were then put in the holes of the copper block in a chamber. Care was taken that each vessel was set as far as possible from the other member of its pair to ensure that equality of concentrations in each pair at the end of a run was an adequate criterion of equilibrium. After the solutions were put into a chamber it was evacuated for several hours with an aspirator pump. To prevent air or vapor bubbles from splashing in the solutions the entire chamber was precooled and set in a pan of ice water during the first part of the evacuation.

After evacuation the rubber tube on the inlet was clamped shut. The chamber was then put in a 25° water bath for a period of a week or two. Temperature control in the bath was ±0.005°, and the chambers were rocked through an angle of 20° once every 3 sec. At the end of a run the chamber was removed from the water bath, the vacuum broken, and the chamber top removed, and the vessels were quickly closed with tight-fitting polyethylene stoppers. The isopiestic vessels were weighed on a balance with their tops off while on the balance pan. Tests indicated that weight changes in the solutions due to condensation or evaporation during the time they were open would be no greater than 0.2 mg.

(23) We have used only 1+,1− electrolytes in this study and are assuming that the protein counterions are also univalent. The parameter Z is the absolute value of the charge, regardless of sign.

(24) H. Dintzis, Ph.D. Thesis, Harvard University, 1952.

(25) Y. Nozaki and C. Tanford, *J. Am. Chem. Soc.*, **89**, 736 (1967).

(26) R. A. Robinson and R. H. Stokes, "Electrolyte Solutions," 2nd ed, Revised, Butterworths and Co. (Publishers) Ltd., London, 1965.

(27) A standard of known water activity was present in some series of experiments because absolute values of the osmotic coefficient of GuHCl are being determined by us as a parallel project. Some of the results will be used below.

All weighings were corrected to vacuum. The density used for crystalline guanidine hydrochloride was 1.35 g/cc, that for cesium chloride was 3.327. The densities of cesium chloride and GuHCl solutions were taken from published tables.[9,28]

Experimental Precision. Experimental data will be reported to five significant figures in the concentrations. The *absolute* concentrations are probably not known to quite this accuracy. As will be seen, however, the final results depend on *differences* in water content at equilibrium and on the resulting *differences* in concentration between reference solutions and those which contain protein. These differences depend only on the accuracy of the final weighings, and, on that basis, are correct to within $^1/_{2500}$ of the concentrations themselves. The number of significant figures in the tables is intended to reflect the precision of the differences, rather than that of the absolute concentrations.

Results

All the data reported were obtained at 25°, in 2.5 *m* CsCl, and in GuHCl solutions which ranged from 9.60 to 10.60 *m*. A molal concentration of 10.5 *m* GuHCl corresponds to a molar concentration of 6.0 *M*.

As explained in the Experimental Section, each solution containing given amounts of salt and protein was placed into two separate vials. About 0.2 g of H_2O was added to one of the vials, none to the other. This had the effect of making the salt concentration of one member of each pair initially *higher* than the equilibrium concentration, while that of the other member was initially *lower*. Thus one member of each pair gained water during equilibration, while the other lost water. Equalization of the final salt concentrations of the two members of a pair thus served as a criterion for the attainment of equilibrium. Since the rate of water transfer decreases as the solutions approach equilibrium, it is not practical to wait until complete equilibrium (within the error of weighing) is achieved. We found it convenient to equilibrate solutions for 1 to 2 weeks, at the end of which time the concentrations of duplicate vials were always within about 0.1% of each other, or closer. Variations in closeness of approach to equilibrium, between one set of experiments and another, reflect differences in the extent of evacuation of the isopiestic chambers.

It was found, in almost every pair, that the solution which was originally more dilute remained more dilute at the end of the equilibration period.[29] This is the expected result if the major cause of the residual difference is due to a residual deviation from equilibrium. Under these conditions, the mean composition of each pair of solutions is clearly closer to the true equilibrium composition than either member of each pair. The fact that the two members of a pair were placed as far apart as possible in the equilibration chambers should assure that solutions of different composition are at least as close to equilibrium as the two members of a pair.

Typical results, showing initial and final compositions, are given in Tables I and II. Table II represents one of the largest deviations between duplicates observed in any experiment. It may be noted that the difference between duplicates in Table II is of the order of 0.4 g/kg of H_2O in g_3, which corresponds to a difference of about 0.02 in $\Delta g_3/\Delta g_2$. The standard deviation in the value of $\partial g_3/\partial g_2$, based on the means of duplicates, will be seen (in Table IV) to be significantly smaller than this difference. This is the expected result if duplicate members of pairs bracket equilibrium compositions, as explained in the preceding paragraph.

(28) K. Kawahara and C. Tanford, *J. Biol. Chem.*, 241, 3228 (1966).

(29) Exceptions occurred, as in the last entry of Table I, only when the difference between members of a pair was very small.

Table I. Comparison between Pairs of Solutions (Serum Albumin in 2.5 *m* CsCl)[a]

Salt, Initial	Salt, Final	Protein final
420.28	412.96	. . .
409[b]	412.91	. . .
419.74	412.31	8.40
404[b]	412.25	8.40
419.27	411.79	16.06
405[b]	411.75	16.06
418.83	411.19	23.26
405[b]	411.24	23.25

[a] Concentrations in g/kg of H_2O. [b] These solutions were obtained by adding *ca.* 0.2 g of H_2O to a duplicate of the final member of each pair. The CsCl and protein contents were known exactly, but the exact amount of added water was not determined, and concentrations are for this reason given to only three significant figures.

Table II. Comparison between Pairs of Solutions (Aldolase in 10.5 *m* GuHCl)[a]

GuHCl, Initial	GuHCl, Final	Protein final
999.74	985.75	0
958[b]	985.29	0
998.67	986.68	10.57
959[b]	986.44	10.57
997.77	987.37	18.76
960[b]	987.06	18.76
996.86	987.77	27.66
932[b]	987.33	27.66

[a] Concentrations in g/kg of H_2O. [b] See footnote *a* of Table I.

Equilibrium compositions (mean values of duplicate pairs) for all solutions of proteins in aqueous GuHCl are shown in Table III. The mean compositions for

Table III. Isopiestic Compositions[a]

	g/kg of H_2O		g/kg of GuHCl	
	Protein	GuHCl	Protein	H_2O
Aldolase	. . .	985.52	. . .	1014.69
	10.57	986.56	10.71	1013.62
	18.76	987.22	19.00	1012.94
	27.66	987.55	28.01	1012.60
β-Lactoglobulin	. . .	985.52	. . .	1014.69
	11.34	986.55	11.49	1013.63
	22.02	987.28	22.30	1012.88
	29.44	988.16	29.79	1011.98
Lysozyme	. . .	916.22	. . .	1091.44
	12.15	917.16	13.25	1090.32
	18.16	917.80	19.79	1089.56
Ovalbumin	. . .	1002.63	. . .	997.37
	12.58	1003.88	12.53	996.13
	22.84	1005.23	22.72	994.80
	28.62	1005.82	28.45	994.21
Ribonuclease	. . .	1002.63	. . .	997.37
	10.39	1002.47	10.36	997.54
	18.07	1003.13	18.01	996.88
Ribonuclease	. . .	1012.91	. . .	987.25
	16.84	1012.70	16.63	987.46
	25.15	1012.54	24.84	987.62
Serum albumin	. . .	916.22	. . .	1091.44
	10.45	916.84	11.40	1090.70
	20.54	917.72	22.38	1089.65
	30.90	918.33	33.65	1088.93
Serum albumin	. . .	1012.91	. . .	987.25
	20.30	1013.82	20.02	986.37
	30.30	1014.74	29.86	985.47

[a] Protein in 9.6–10.6 *m* GuHCl.

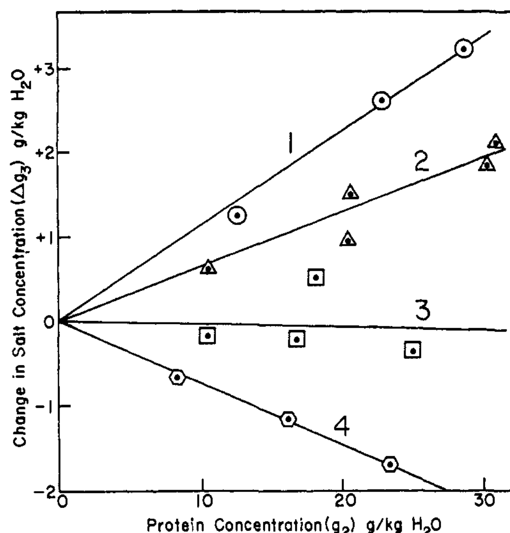

Figure 1. Experimental data of Table III, plotted to show the changes in salt concentration with protein concentration: curve 1, ovalbumin in GuHCl; curve 2, serum albumin in GuHCl; curve 3, ribonuclease in GuHCl; curve 4, serum albumin in CsCl.

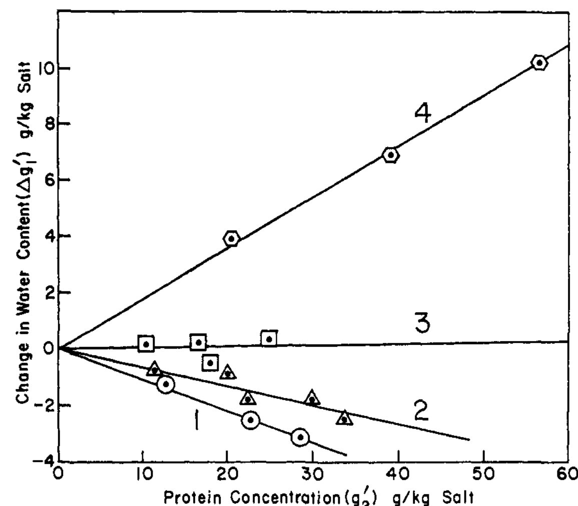

Figure 2. Experimental data of Table III, plotted to show the changes in water concentration with protein concentration. The numbers correspond to the same systems as in Figure 1.

the system serum albumin–CsCl–H_2O are not listed separately as they are readily obtained from Table I.

In Figure 1 we have plotted the change in salt concentration (Δg_3) as a function of protein concentration. It is seen that Δg_3 is positive for unfolded proteins in 6 *M* GuHCl, except for ribonuclease, for which Δg_3 is essentially zero. On the other hand, Δg_3 is negative for native serum albumin in 2.5 *m* CsCl.

The difference between native serum albumin in the CsCl solution and the unfolded proteins in GuHCl solution is shown in terms of hydration in Figure 2, in which we have plotted $\Delta g_1'$ *vs.* g_2'. A positive slope is obtained in CsCl and a negative one in GuHCl. (As was noted in the introduction, this kind of difference between the two systems is consistent with theoretical expectation.)

The slopes of Figures 1 and 2 give the values of $(\partial g_3/\partial g_2)_{\mu_1}$ and $(\partial g_1'/\partial g_2')_{\mu_1}$, respectively. The figures indicate no systematic deviations from linearity. Within experimental error, which is evidently quite large, these derivatives are clearly constants within the concentration range of the experiments, and may therefore be taken as representative of their limiting values at zero protein concentration. The actual values of the derivatives, and their standard deviations, were calculated by the method of least squares. The two separate experiments with ribonuclease and serum albumin (in GuHCl; see Table III) were each treated as a single experiment for this purpose. The concentrations of GuHCl in the two experiments with serum albumin differed by about 10% (9.60 and 10.60 *m* GuHCl, respectively), but the results showed that this difference did not lead to significant differences in the derivatives.

The derivatives can be converted to molal derivatives by eq 1 and 2, and the derivatives $(\partial g_3/\partial g_2)_{\mu_1}$ or $(\partial m_3/\partial m_2)_{\mu_1}$ can be converted to derivatives at constant μ_3 by eq 8. To carry out this conversion, it is convenient to rewrite eq 8 in terms of molal osmotic coefficients[26] (φ_3), *i.e.*

$$\frac{1}{1 + m_3(\partial \ln \gamma_3/\partial m_3)_{m_2=0}} = \frac{1}{\nu\phi_3 + \nu m_3(\partial\phi_3/\partial m_3)_{m_2=0}} \tag{9}$$

where, for an electrolyte, ν is the number of ions produced per mole of the component. The value ν is 2 for both CsCl and GuHCl.

We have used the data of Robinson and Stokes[26] to evaluate this factor for CsCl. For GuHCl, ϕ_3 has been evaluated as part of the present series of experiments.[27] We have found that $\phi_3 = 0.655$, and that it is essentially independent of m_3 between $m_3 = 9.6$ and 15.7. The factor given by eq 9 turns out to be very small, as can be seen from footnote *a* of Table IV.

Table IV. Preferential Binding at 25°

	g/g of protein		moles/mole of protein	
	$\left(\frac{\partial g_1'}{\partial g_2'}\right)_{\mu_1}$	$\left(\frac{\partial g_3}{\partial g_2}\right)_{\mu_3}$	$\left(\frac{\partial m_1'}{\partial m_2'}\right)_{\mu_1}$	$\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_2}$
Native Serum Albumin in 2.5 *m* CsCl				
Serum albumin	+0.180	−0.073	+690	−30
Randomly Coiled Proteins in 10.5 *m* (6 *M*) GuHCl				
Ribonuclease	+0.004	+0.001[a]	+3	0
Serum albumin	−0.067	+0.064	−257	+46
Aldolase	−0.082	+0.082	−182	+34
β-Lactoglobulin	−0.088	+0.090	−90	+17
Lysozyme[b]	−0.092	+0.089	−73	+13
Ovalbumin	−0.110	+0.113	−275	+53
Standard deviation[c]	±0.005	±0.007	...	...

[a] $(\partial g_3/\partial g_2)_{\mu_1}$ must necessarily have the opposite sign from $(\partial g_1'/\partial g_2')_{\mu_1}$. In this case $(\partial g_3/\partial g_2)_{\mu_1}$ is actually −0.004, and the difference between $(\partial g_3/\partial g_2)_{\mu_3}$ and $(\partial g_3/\partial g_2)_{\mu_1}$, as given by eq 8 or 9, is +0.005. The difference between $(\partial g_3/\partial g_2)_{\mu_3}$ and $(\partial g_3/\partial g_2)_{\mu_1}$ was ≤0.005 for all the data shown. [b] The figures given are subject to systematic error if the proteins introduced into the solutions were not at their isoionic pH's. Counterions would then be included as part of the protein component, and their effect on water activity must be corrected for. The correction is to increase $\partial g_3/\partial g_2$ or $\partial m_3/\partial m_2$ and to make $\partial g_1'/\partial g_2'$ or $\partial m_1'/\partial m_2'$ more negative. The correction is likely to be significant for lysozyme, but not for any of the other proteins listed. Corrected values for lysozyme are likely to be about 30% larger than the values listed. Further work with lysozyme is in progress, and results will be reported at a later date. [c] The standard deviation is an average value, obtained from the least-squares analysis of data for individual proteins. The deviation is somewhat larger than the average value in the case of ribonuclease, for which one experimental point is rather far from the rest, as seen in Figures 1 and 2, and it is less for the data for serum albumin in 2.5 *m* CsCl. The deviations in molarity units are obtained by multiplying those in mass units by M_2/M_1 or M_2/M_3, respectively, and thus differ for each protein.

The final summary of all the results obtained in this study is given in Table IV, in terms of all four of the useful derivatives of eq 1 and 2.

Discussion

The data of Table IV show that native serum albumin is preferentially hydrated in 2.5 *m* CsCl. The value of 0.18 g of H_2O/g of serum albumin agrees well with the determination by Ifft and Vinograd,[11] based on buoyant behavior in the ultracentrifuge. Results of the same order of magnitude have been obtained by Cox and Schumaker[10] for various native proteins in aqueous CsCl and in mixtures of CsCl and $(NH_4)_2SO_4$.

The data show, on the other hand, that unfolded proteins in 10 *m* GuHCl preferentially bind GuHCl rather than water, to the extent of 0–0.11 g of GuHCl/g of protein, depending upon the protein examined. These results, too, are consistent with previous measurements in 5 or 6 *M* GuHCl by other methods. Kielley and Harrington,[14] for example, studied myosin in 5 *M* GuHCl, and obtained preferential binding of the salt to the extent of about 0.05 g/g of protein. Indirect measurements, obtained in connection with molecular weight determination in concentrated GuHCl solutions, to be cited below, are also consistent with preferential salt binding to about the same extent. On the other hand, Schachman and Edelstein[30] have reported preferential hydration of aldolase in concentrated GuHCl, using the method of buoyant behavior in the ultracentrifuge. We do not know the reason for the large difference between their result and ours for the same protein.

Molecular Weight Determination in 6 *M* GuHCl. We pointed out in the introductory section that the motivation for this study came partly from the importance of preferential interactions in the determination of molecular weights. Molecular weight determinations in 6 *M* GuHCl, usually by sedimentation equilibrium, have become an important tool for the protein chemists, because this solvent appears to break all noncovalent attractions within and between protein polypeptide chains. If suitable steps are taken to make sure that disulfide bonds are broken, proteins in 6 *M* GuHCl appear to be always dissociated to their constituent polypeptide chains, and the chains themselves appear to be random coils devoid of long-range structure.[17–20] Molecular weight determinations under these conditions will therefore yield data for the constituent chains, and, by comparison with the molecular weight of the native protein, permit calculation of the number of polypeptide chains in the native protein molecule.[31,32]

Assuming a single protein component,[33] the equation for sedimentation equilibrium at the limit of zero protein concentration is

$$2RT\frac{\mathrm{d}\ln m_2}{\mathrm{d}r^2} = M_2\omega^2(1 - \bar{v}_2\rho)\left[1 + \frac{1 - \bar{v}_3\rho}{1 - \bar{v}_2\rho}\left(\frac{\partial g_3}{\partial g_2}\right)_{3\mu}\right] \quad (10)$$

where r is the position in the ultracentrifuge cell, ω the angular velocity of the rotor, ρ the density of the solvent, and $\bar{v}_2$ and $\bar{v}_3$ are partial specific volumes. The equation shows that M_2 can be obtained unequivocally from the experimental concentration gradient if $(\partial g_3/\partial g_2)_{\mu_3}$ can be determined by the method described in this paper, or by the alternative techniques to which we made reference in the introductory section.

It should be noted that there is still some question concerning the values of $\bar{v}_2$ for proteins in 6 *M* GuHCl.[34] There are reasons for believing that $\bar{v}_2$ will in general not differ significantly from its value in dilute aqueous salt solutions, and we shall assume that this is so far the purpose of this discussion, but precise data on this point are lacking and sorely needed. The value of $\bar{v}_3$ for GuHCl has been determined,[14] and turns out to be close to typical values for proteins, so that the factor $(1 - \bar{v}_3\rho)/(1 - \bar{v}_2\rho)$ may be placed equal to unity for the purpose of approximate calculations.

Using eq 10, the foregoing assumptions about partial specific volume, and the values of $(\partial g_3/\partial g_2)_{\mu_3}$ reported in Table IV (which we assume to be independent of pressure), it is seen that molecular weights determined in 6 *M* GuHCl, without taking preferential interactions into account, will in general be somewhat too large, typically by 5 to 10%.

It has been shown by Casassa and Eisenberg[8,13] that one can essentially avoid the problem of preferential binding of solvent components by defining the solvent component as that solution which is in osmotic equilibrium with the protein solution across a membrane impermeable to component 2. When this is done, the partial derivative of eq 10 is automatically zero,[35] but $\bar{v}_2$ must be replaced by an apparent specific volume ϕ_2', defined in terms of the same solvent component. Actual values of ϕ_2' for proteins in 6 *M* GuHCl have been measured in a number of instances,[36,37] and the observed values have been about 0.01 to 0.02 cc/g less than $\bar{v}_2$, which corresponds to about 5 to 10% preferential binding of GuHCl. In other instances, ϕ_2' values have not been measured, but have been *assumed* to be about 0.01 to 0.02 less than $\bar{v}_2$, and this practice has led to molecular weights and determinations of the number of polypeptide chains per molecule which are in agreement with information obtained by other methods.[38] Clearly, the present data are consistent with the use of the Casassa–Eisenberg procedure in the determination of molecular weight, and even the use of an assumed value of ϕ_2' (in 6 *M* GuHCl) which is 0.01–0.02 below $\bar{v}_2$, and molecular weights determined in that way should not require revision.

Differences between Proteins. We should like to comment in conclusion on the differences between individual proteins (in 6 *M* GuHCl) which are observed in the data of Table IV. Since all proteins in 6 *M* GuHCl are devoid of specific structural features, differences in behavior should as a good approximation be

(30) H. K. Schachman and S. J. Edelstein, *Biochemistry*, 5, 2681 (1966).

(31) If the polypeptide chains differ in molecular weight, then the ratio of the native molecular weight, M_0, to the number-average molecular weight of the chains, $\bar{M}_n$, gives the number of chains, regardless of size.

(32) K. Kawahara and C. Tanford, *Biochemistry*, 5, 1578 (1966).

(33) For a mixture of polypeptide chains, eq 10 gives the local weight-average molecular weight. The term in brackets would represent a similar average for the mixture.

(34) F. J. Reithel and J. D. Sakura, *J. Phys. Chem.*, 67, 2497 (1963).

(35) Actually, it is $(\partial g_3/\partial g_2)_{\mu_1,\mu_3}$ (but not at constant pressure) which vanishes under these conditions, but the difference between this derivative and that of eq 10 is likely to be negligibly small, since $(\partial g_3/\partial g_2)_{\mu_1}$ has been shown above to differ only trivially from $(\partial g_3/\partial g_2)_{\mu}$ (both at constant pressure).

(36) E. Marler, C. A. Nelson, and C. Tanford, *Biochemistry*, 3, 279 (1964).

(37) P. A. Small and M. E. Lamm, Jr., *ibid.*, 5, 259, 267 (1966).

(38) L. Kanarek, E. Marler, R. A. Bradshaw, R. E. Fellows, and R. L. Hill, *J. Biol. Chem.*, 239, 4207 (1964).

entirely explicable in terms of differences in amino acid content. We believe that the differences in binding which have been observed can in fact be reasonably explained in this way. The observed differences are actually very small. A theoretical discussion of the preferential hydration of native proteins has been presented by Shumaker and Cox,[3] and they show that several hundred solvent molecules come within the sphere of influence of the surface of a native protein molecule. The number of solvent molecules associated in close contact with a randomly coiled protein molecule must be considerably larger, *i.e.*, the number is likely to exceed 1000 for most of the proteins for which we have reported data. The preferential binding in terms of moles of GuHCl/mole of protein, which lies in the range of 0 to 53 (last column of Table IV), thus reflects very small preferences indeed, and individual differences of the kind observed would seem to be entirely reasonable, since the different amino acid side chains presumably differ from each other in their preference for contact with water or GuHCl.

Acknowledgment. We wish to acknowledge the expert technical assistance of Mrs. K. B. Hade.

An Analysis of the Tyrosine Circular Dichroism Bands in Ribonuclease[1]

Norman S. Simmons and Alexander N. Glazer

Contribution from the Department of Biophysics, Laboratory of Nuclear Medicine and Radiation Biology, University of California, Los Angeles, California 90024, and the Department of Biological Chemistry, Center for the Health Sciences, University of California, Los Angeles, California 90024. *Received May 25, 1967*

Abstract: Ultraviolet circular dichroism curves have been measured for bovine pancreatic ribonuclease and for a model compound, N-acetyl-L-tyrosinamide, in neutral and alkaline solutions. The influence of ethylene glycol on the circular dichroism was also examined. The correspondence in the circular dichroism curves of ribonuclease and N-acetyl-L-tyrosinamide under various conditions is discussed in relation to the origin of the tyrosine Cotton effect exhibited by the native enzyme.

In a recent communication[2] we described near-ultraviolet Cotton effects in ribonuclease (RNase) which were sensitive to both pH and denaturing agents. The red shift and increase in amplitude of the Cotton effect in the region of the tyrosine absorption bands, which occurred on raising the pH from 6 to 11.5, led us to the conclusion that the readily ionizable tyrosine residues[3,4] made a major contribution to the observed anomalous dispersion. Recently, Simpson and Vallee[5] reported that reaction of native RNase with N-acetylimidazole led to a conversion of three tyrosine residues to the O-acetyl derivative. This reaction is associated with a large decrease in the absorbancy of tyrosine at 275 mμ and loss of optical activity associated with this tyrosine chromophore. Further, they found that whereas acetylation of native RNase led to the expected decrease in absorbancy at 278 mμ, there was *no concomitant change in optical activity at pH 6.* On the assumption that the modified residues were those showing normal ionization behavior, Simpson and Vallee[5] concluded that the major contribution to the anomalous dispersion arose from among the three "buried" residues. In view of these results, our explanation that the normally ionizing tyrosine residues contribute significantly to the total optical activity observed at neutral pH seemed no longer tenable. We have therefore undertaken an analysis of the circular dichroism (CD) of RNase in an attempt to resolve these apparently conflicting interpretations. The CD spectrum of a model compound N-acetyl-L-tyrosinamide (NACTA) was similarly examined.

(1) This work was supported in part by Grant No. GM 11061, Career Award No. K6-DE-1094 from the National Institutes of Health, U. S. Public Health Service, and Contract AT(04-1)GEN-12 between the U. S. Atomic Energy Commission and the University of California.

(2) A. N. Glazer and N. S. Simmons, *J. Am. Chem. Soc.*, 87, 3991 (1965).

(3) D. Shugar, *Biochem. J.*, 53, 142 (1952).

(4) C. Tanford, J. D. Hauenstein, and D. G. Rands, *J. Am. Chem. Soc.*, 77, 6409 (1955).

(5) R. T. Simpson and B. L. Vallee, *Biochemistry*, 5, 2531 (1966).

Experimental Section

Crystalline bovine pancreatic ribonuclease, Lot R662-ML, was obtained from Worthington Biochemical Corp. All circular dichroism measurements were performed with a Jasco Model ORD/UV5 instrument. The data displayed in the figures are direct reproductions of the chart recordings and are therefore reported as $\Delta E = E_L - E_R$. The concentrations of RNAase and of the model compound were chosen so as to give the same molar concentration of tyrosine, and hence the magnitudes of the CD curves obtained may be compared directly

Results and Discussion

As can be seen in Figure 1, RNase shows a strong negative CD band centered at 273 mμ at pH 6, ascribable to tyrosine residues. On raising the pH to 11.5 there is a gradual shift of this band to a position centered at 285 mμ, that is, in the direction expected from the shift in the absorption spectrum associated with the ionization of tyrosine residues. The curves shown are similar to those recently reported by Beychok.[6]

If the CD band centered at 273 mμ is due to the "buried" residues and does not change in magnitude or

(6) S. Beychok, *Science*, 154, 1288 (1966)