

STATE OF SOLUTION OF COLLOIDAL ELECTROLYTES

By G. S. HARTLEY, D.Sc.

AN electrolyte is a conducting solution of an ionised compound. A colloidal solution is one where the constituent particles of one component are so large that markedly peculiar properties arise directly due to their great size or, in a way which would not arise were the particles small, to their shape. In the generally accepted meaning of the term colloidal electrolyte, these definitions have not simply been united, but emphasis has at the same time been laid on the simple reversible, thermodynamically stable nature of simple electrolyte solutions. Colloidal electrolytes have come to mean, therefore, true reversible solutions which have some colloidal properties and in which the electric charge is an integral part of the dissolved substance. Except where anomalous viscous and gelling properties occur, they stand at the very small end of the range of particle sizes usually considered to be colloidal.

In a ferric hydroxide sol, although the dispersion as such is not stable, as can be shown by the irreversible coagulation resulting from addition of excess electrolyte, yet it changes so slowly that many properties dependent on it, for example, activity of hydrogen chloride in the sol, can be treated as in equilibrium. In a colloidal electrolyte, on the other hand, it is the state of dispersion itself which must be at equilibrium. In general, the system will be fully in equilibrium. If irreversible chemical changes occur which affect the dispersion, this will accommodate itself rapidly to the state appropriate to the instantaneous composition.

Quantitatively perhaps the most important colloidal electrolyte in industry is the solution of cellulose xanthate in alkali (viscose). Its physical properties change very profoundly with time. A ferric hydroxide sol is, in comparison, apparently far more stable. We know, however, that, in the latter case, the dispersion is not in an equilibrium state. It can readily be changed in an irreversible manner. The instability of viscose, however, arises from the irreversible chemical reactions of the xanthate groups. By analogy we can be sure that, could the xanthate reactivity be "frozen" at any desired stage, the physical properties would become constant and truly reflect the reversible equilibrium of the solution.

We will limit the term "colloidal electrolytes" in this way to include only physically reversible solutions where the solute is colloiddally dispersed and where the electrolytic properties are due to ionic groups forming an integral part of the solute molecules. We still include solutions of diverse types which cannot adequately be treated only in terms of the properties they have in common.

They fall into two main groups. The first we might term "sub-ultimate" solutions of ionised compounds of fairly small molecular weight: "sub-ultimate" because the equilibrium dispersion is not into the ultimate molecules or ions but falls far short of this. These electrolytes are colloidal

as a result of some physical process of aggregation of molecules which does not proceed, however, beyond some stable limit. A soap solution is the most familiar example. The second group is formed by compounds which are (on dissolution) dispersed to the ultimate molecules but where these are themselves so large that colloidal properties result. Gum arabic solution is perhaps the simplest example on which a large amount of information has been obtained, but the group includes the numerous and important proteins as well as many substances on which little work has been done but of which a further study would be most fruitful, in salts derived from synthetic linear macromolecules—for example, acrylic acid and its salts.¹

We will first attempt some description of the distinguishing properties of colloidal electrolytes which are common to both groups, *i.e.*, properties in which both water-soluble soap-like salts and gums are to be distinguished from salts of simple organic acids.

The Size of the Units in the Dissolved State

The direct result of the great molecular weight (real or, in the first group, effective, as a result of aggregation) on the classical properties of dilute solutions—osmotic pressure, depression of freezing point, etc.—is largely obscured, when the salts are pure, by the compensating effect of the numerous small ions of opposite sign whose presence is demanded by electrostatic forces. We cannot confine the large ions, R , of a salt behind a membrane without having z univalent (or $z/2$ bivalent, etc.) small ions in company, where z is the valency of the large ion or aggregate. To a first approximation therefore the salt will, in its osmotic properties, behave as though its molecular weight were $1/(z + 1)$ of its stoichiometric value.

The small ions are not held by the membrane itself, but by the electrostatic field resulting from the direct retention of the large ones. If other ions are present in the solution outside the osmotic membrane, a partial exchange will result: the well-known Donnan equilibrium will be set up. If the outside solution contains NaCl at a constant concentration c_2 and the inside solution has R ions (z -valent) in molar concentration m_1 , then, if c_1 is the concentration of Cl ions at equilibrium inside the membrane, that of Na^+ ions must be $zm_1 + c_1$. The Donnan equation, in its simplest form, states that the ionic product of NaCl must be the same inside and out, *i.e.*,

$$(zm_1 + c_1)c_1 = c_2^2 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The excess molecular concentration, effective osmotically, is

$$m_1 + 2c_1 - 2c_2$$

and substitution of equation (1) in this shows that the osmotically effective concentration falls from $(z + 1)m_1$ for $c_2 = 0$ towards m_1 as $c_2/zm_1 \rightarrow \infty$.

We may note that it is not possible for the outside liquid to be completely free of ions, since water in its pure state is slightly dissociated. There will therefore be a tendency for sodium ions, in the case quoted, to diffuse through the membrane and build up a low concentration of sodium hydroxide

¹ H. Staudinger and E. Trommsdorff, in H. Staudinger "Die hochmolekularen organischen Verbindungen" (Berlin, 1932), p. 333.

outside, leaving the inside liquid correspondingly acid. The concentration of alkali built up at equilibrium when the outer liquid has the same volume as the inner and was initially pure water is given by $(zm_1K_w)^{1/3}$ where K_w is the ionic product for water. While this does not represent a serious error under the simplest conditions, and can be corrected for, it becomes much larger if the free acid HR is a weak one. If this acid is insoluble, hydrolysis will be carried yet further. Moreover, if the outer liquid is of much larger volume or, as is the case in purification by dialysis, is repeatedly changed, hydrolysis will proceed to completion unless alkali of the calculated concentration is used in place of pure water as the outside liquid. This, and many other applications of the Donnan equilibrium, are described in T. R. Bolam's book.²

The effect of interionic forces between free ions in a solution has a depressant effect on the osmotic pressure and cognate properties. A salt which we should expect to be completely dissociated behaves at any one concentration as though it were only partly dissociated, but the degrees of dissociation calculated from osmotic pressure at various concentrations do not conform to the mass law to enable a dissociation constant to be obtained. Pure interionic effect on osmotic pressure is calculable for the simplest electrolytes at very low concentrations by the Debye-Hückel theory and approximately under more practical conditions from extensions of that theory. Even for simple electrolytes, however, it is not possible to use the calculations to estimate reliably any residual degree of real association of ions. Other methods must be resorted to.³ All we can with any confidence say about the effect of interionic forces in our solution of z -valent ions (where z may be of the order of 100) with their associated univalent ions, is that this effect will be much larger than in simple electrolytes and certainly not exactly, or even approximately, calculable.

Osmotic measurements on a pure solution of the colloidal electrolyte will therefore tell us very little about the effective molecular weight, because (a) the osmotic pressure is mainly due to the much more numerous small ions and (b) the relationship between osmotic pressure and effective concentration will be obscured by a large and uncertain effect of interionic forces.

In the case, however, where a "swamping excess" of a simple electrolyte is used in the osmotic-pressure measurement, not only are the simple ions brought to approximate equality in total concentrations on either side of the membrane, but it can be shown that they differ from complete equality in exactly the right amount to compensate for the effect of interionic forces.⁴ The colloidal ions now behave, not only as though they were in solution by themselves, but as though in ideal solution, except in so far as non-ionic deviations from ideality, which are important in all solutions of macromolecules, are operative.

Osmotic measurements, therefore, applied to a solution in a swamping

² "The Donnan Equilibrium" (London, 1932).

³ C. W. Davies, *J.*, 1938, 271; K. Fajans, *Trans. Faraday Soc.*, 1927, **23**, 357.

⁴ G. S. Hartley, *Trans. Faraday Soc.*, 1935, **31**, 106, in comment on paper by F. G. Donnan, *ibid.*, p. 80.

excess of simple electrolyte, can give us information about the size of the colloidal ion which is as precise as that obtainable in the case of macro-molecular non-electrolytes dissolved in organic solvents.

It cannot be said that the same remarks apply to related measurements such as freezing-point or vapour-pressure depression, because these are not in principle applicable in the case of the solution in the swamping excess of simple electrolyte. Even were there no restriction in principle, a degree of accuracy would be demanded which is quite outside practical possibility for freezing-point measurements. Direct osmotic measurements therefore stand out as of supreme importance.

Other means of determining the size of the colloidal ions, corresponding to means of determination of molecular weight in non-ionic solutions, are sedimentation velocity, diffusion velocity, and sedimentation equilibrium. Of these, any one quantity is related by strict thermodynamic argument to the other two. Only two independent measurements can therefore be made, the third being only a check on the experimental accuracy, desirable though this may be.

For the detailed application of sedimentation methods, reference must be made to the specialised literature,⁵ but the essence of the argument is as follows. Transfer of a small amount of dissolved substance from one depth to another against the centrifugal field requires work exactly calculable from the partial molar volume, the molecular weight, and the strength of the field. At equilibrium this is exactly counterbalanced by the decrease in free energy on moving from a higher to a lower concentration. The equilibrium state is thus determined, and if the concentration gradient, the strength of the centrifugal field, and the partial specific volume are measured or calculated, the molecular weight may be calculated.

The sedimentation velocity (at zero concentration gradient) is equal to the driving force per molecule divided by the frictional resistance. For a spherical particle this is equal, according to Stokes's law, to $6\pi\eta a$, where η is the viscosity of the solvent and a is the radius. For non-spherical particles, a may be taken as the radius of a sphere of equal volume without serious error if the dimensions are not greatly unequal. For very long rods or flat discs, however, the resistance coefficient becomes very much greater. In non-electrolytes, therefore, the combination of osmotic measurements, which are unaffected by particle shape, with sedimentation (or diffusion) velocity measurements, can give some information about the "axial ratio" of a long molecule.⁶

The diffusion velocity is derived theoretically by the following argument or its logical equivalent. Knowing the sedimentation equilibrium, we may say that the rate of transfer of solute by sedimentation, in the absence of the concentration gradient which would balance it under equilibrium conditions, must be equal and opposite to the transfer which would be produced by this concentration gradient in the absence of the sedimenting force.

⁵ The Svedberg, "The Ultracentrifuge" (Oxford, 1940).

⁶ See P. Johnson, *Ann. Reports*, 1946, **43**, 30, and references therein.

All these methods, applied to a pure solution of a colloidal electrolyte, suffer from an exactly analogous disadvantage to that operating in the case of osmotic pressure, due to the necessity for an equivalent member of small ions to accompany the colloidal ion in its movements. The free-energy change of transfer of one colloidal ion is thus $(z + 1)$ times that for a corresponding transfer of an uncharged colloidal aggregate or macromolecule.

Thus, in the case of sedimentation equilibrium, if we know the equivalent weight of the colloidal electrolyte and the other necessary quantities, the sedimenting force will be equal to Kz where K is known. The free-energy term will, however, be proportional to $(z + 1)$ so that at best our measurement gives us $z/(z + 1)$. In view of the uncertain and large errors in the simple theory due to interionic effects, it tells us therefore nothing of what we want to know, except that we can distinguish between $z = 1$ and some large number, *i.e.*, can detect aggregation in what might have been a simple electrolyte.

Similarly, in the case of diffusion. Here, for electrolytes, we may derive the resistance coefficient for the ions from their mobilities in the electric field obtained from conductance and transport-number measurements. The relating equation is

$$D = \frac{RT}{F^2 \times 10^7} \left(\frac{1}{z} + 1 \right) \cdot \frac{l_+ l_-}{l_+ + l_-} \quad (2)$$

where D is the diffusion coefficient, F the equivalent of electricity in coulombs, and l_+ , l_- are the mobilities in ordinary units (ohm cm.² g.-equiv.⁻¹). When z is large the expression is obviously very insensitive to z and, having regard to its approximate nature, based on "ideal solution" behaviour (*i.e.*, ignoring interionic effects), the diffusion measurement tells us nothing more than can be derived from the conductivity and transport-member measurements.

H. R. Bruins⁷ and the Reviewer and C. Robinson⁸ simultaneously and independently pointed out the significance of the above equation and showed that in sufficiently purified starch and gum sols diffusion was as rapid as in many simpler solutions, even more rapid than that of copper sulphate in water! This observation is of especial interest, as Graham's first distinction of a colloid from a "crystalloid" was the much lower diffusion velocity of the former. It was incidentally a colloidal electrolyte, namely glue, which gave its name to the whole class "colloid". Had Graham's glue been dialysed it would have appeared, on this basis, to be a crystalloid! That the earlier examinations of colloid electrolytes showed them to diffuse slowly was due, of course, to their containing sufficient salt impurities to swamp the effect of the high potential gradient which arises when a pure colloidal electrolyte diffuses.

The restriction of precise information to solutions containing an excess of simple salt may seem a severe limitation in the case of those colloidal electrolytes which are colloidal as a result of aggregation of simple ions. As we shall see when discussing these aggregating electrolytes in more detail,

⁷ *Kolloid-Z.*, 1931, **54**, 265, 272.

⁸ *Proc. Roy. Soc.*, 1931, **A**, **134**, 20.

however, in some cases the effect of simple electrolytes on aggregation is not very great, and in others it is the state of aggregation in excess of electrolyte which is of most importance in practice. In the case of electrolytes which are colloidal as a result of the great size of the primary ions, information obtained about molecular weight in excess of electrolyte can more confidently be applied to the pure solutions.

The most widely used method for determination, or, at least characterisation, of molecular weight of non-ionic polymers, is by measurement of viscosity in very dilute solution. The method was first introduced by H. Staudinger,⁹ who claimed that the relative increment of solvent viscosity caused by introduction of the polymer was proportional to the concentration of the latter, if small, and the constant of proportionality, termed the intrinsic viscosity, was itself proportional to the molecular length, the constant of proportionality in this case being a function of the particular repeating unit of the polymer. For a given repeating unit, once the relationship between molecular weight and intrinsic viscosity was established by comparison with osmotic-pressure measurements, the molecular weights could be obtained from viscosity data alone, with their advantage of great simplicity and speed of measurement.

Much work has been done on the subject of viscosity of dilute solutions of non-ionic polymers in recent years and the reader is referred to a recent article by R. H. Ewart¹⁰ summarising the present position. Staudinger's original views were certainly an over-simplification of the picture and we now know that the intrinsic viscosity is dependent on the solvent,¹¹ being in general higher the "better" the solvent. It is not proportional to the first power of molecular weight but more nearly to some power, intermediate between 0.5 and 1.0, which is in general greater the more rigid the structure of the molecule.

Although the viscosity of colloidal electrolyte solutions has received much attention in various connections, often because of its practical importance, no application to molecular-weight determination has been attempted. The early interest was mainly in the effect of electrolyte concentration on viscosity. It was found that, in general, the viscosity of dilute solutions of pure colloidal electrolytes was markedly reduced by the addition of simple salts—the so-called electroviscous effect.¹²

This has a similar explanation to the other effects of salt addition mentioned above but with this difference, that it is now on an amicroscopic scale that the colloidal ions are accompanied by the small ions of opposite sign. When no additional simple ions are present, the neutralising ion atmosphere is spread out further around the colloidal ion than when a high electrolyte concentration surrounds it. The viscosity of a colloidal solution is due to the effect of the large molecules in increasing the transfer of

⁹ *Op. cit.*, ref. (1).

¹⁰ In "Advances in Colloid Science" (Interscience, New York, 1946), Vol. II, p. 197.

¹¹ E. M. Frith, *Trans. Faraday Soc.*, 1945, **41**, 17.

¹² H. R. Kruyt and H. G. Bungenberg de Jong, *Z. physikal. Chem.*, 1922, **100**, 250.

momentum between adjacent layers in the sheared liquid. If these large molecules are now ionised and their neutralising atmospheres occupy considerable additional space, transfer of momentum is further increased because the atmospheres tend, by the action of Coulomb forces, to follow the movements of the ions.¹³ The more compact the atmosphere, the smaller is the region of liquid around it in which the interionic effects produce this semi-rigid behaviour.

Correspondingly, we find that if the charge on the colloidal ion itself is reduced, as by neutralisation of ionised weak acid groups, the viscosity is in general decreased. Thus in amphoteric colloidal electrolytes, *e.g.*, proteins, when both weak acid and basic groups are present, the passage of the colloidal solute through the isoelectric point is characterised, among other changes, by passage of the viscosity through a minimum.

The viscosity of colloidal electrolyte solutions becomes anomalous in character in many cases at quite low concentrations, elastic and gelling properties appearing, with evident hysteresis. Such phenomena are conveniently classed together as evidence of "structural" viscosity, it being generally considered that some more or less rigid "scaffolding" of secondary aggregates is responsible. Structural viscosity is frequently profoundly affected in a highly specific way by the ionic environment. Thus, among the paraffin-chain salts, dilute solutions of which are generally very mobile, introduction of particular ions causes highly elastic behaviour to appear: for instance, potassium or sodium oleate solutions are very mobile while that of ammonium oleate is highly elastic.¹⁴ The Reviewer has recorded¹⁵ similar differences in salts of paraffin-chain sulphonic acids which are even more striking because the effects are noticeable in very dilute solutions.

The recent work on the non-electrolyte macromolecules has proved the value of careful study of viscosity in concentrations when it exceeds that of the solvent by a few per cent. only. It seems based on sound, if not theoretically well-understood foundations, and it is difficult to see how the relationship of intrinsic viscosity to molecular length can fail to extend to sufficiently dilute solutions of colloidal electrolytes in salt concentrations where the electroviscous effect is eliminated. At least a maximum molecular length should be established, and the incidence of specific effects due to secondary aggregation should be detectable by careful examination of the effect of salt and colloid concentration.

It is well to realise that the recent enormous advances in our knowledge of organic solutions of non-electrolyte macromolecules is largely due to the examination of synthetic polymers of known and linear structure. Colloidal electrolytes of the macromolecular group which have hitherto received most study are those of natural origin and in many cases uncertain but certainly complex structure. Linear polymers of natural origin are for the most part stereochemically regular and therefore usually crystalline and, being crystalline, insoluble. Alginic acid alone of the naturally occurring carbo-

¹³ W. Krasny-Ergen, *Kolloid-Z.*, 1936, **74**, 172.

¹⁴ E. Hatschek and R. S. Jane, *Kolloid-Z.*, 1926, **38**, 33.

¹⁵ *Nature*, 1938, **142**, 161.

hydrate-type electrolytes has been shown to have a linear structure.¹⁶ Pectin, gum arabic, and others are known to be complex and branched.¹⁷ Modification of cellulose for important technological ends has proceeded almost entirely in the direction of making non-electrolyte derivatives soluble in organic solvents.

Considerable advance might be made if the resources of modern polymer chemistry were tapped to obtain definite linear colloidal electrolytes of known and simple structure. This would be of especial value in the field of viscosity, the electroviscous effect, and structural viscosity. In this connection, work on polyacrylic acid and its salts^{1, 18} and more recently on polyethyleneimine¹⁸ may introduce a new chapter. The intrinsic viscosity of the solution of the free base or acid falls steeply with increase of (very low) concentration to a minimum at about 0.07%. This is probably due to the very low degree of dissociation introducing an electroviscous effect only important at low concentrations and increasing in importance with increasing dilution. One might anticipate that in a high concentration of simple salt the minimum value might persist, or even fall, with dilution of the colloid.

A high intrinsic viscosity in dilute solutions is evidence of the presence of very long molecules. Globular, approximately spherical, molecules will confer only a very low intrinsic viscosity. Einstein's well-known treatment of the flow of a suspension of spheres predicts the increment of viscosity relative to that of the pure solvent (*i.e.*, the specific viscosity) to be equal to $2.5 \times$ the volume fraction of the whole suspension occupied by the spheres.

It is probable, as K. H. Meyer and H. Mark¹⁹ have suggested, that the increase in viscosity of proteins on either side of the isoelectric point is in part due to the straightening of crumpled chains under the influence of increasing repulsive forces between the ionised groups of the side chains, which are predominantly of one sign. This effect will be additional to the normal electro-viscous effect referred to above.¹³ A. Polson²⁰ has carried out interesting measurements on the limiting intrinsic viscosity of various globular proteins in water. He finds the values to be low in comparison with solutions of linear polymers and to show no correlation with molecular weight. No correlation would be expected for spherical particles on Einstein's equation. By comparing the intrinsic viscosity with that predicted by the Einstein equation, Polson estimates the axial ratio to be about 4 for ovalbumin, 5 for serum albumin, and 7 for serum globulin. The intrinsic viscosities for fibrillar proteins are probably much higher but no systematic investigation at adequately low concentrations seems to have been made. H. Bincer²¹ has recently made a full investigation of the viscosity of gelatin and finds a value at limiting low concentrations equivalent to that of cellulose of molecular weight 20,000 in cuprammonium solution.

¹⁶ E. L. Hirst, J. K. N. Jones, and W. O. Jones, *ibid.*, 1939, **143**, 857.

¹⁷ E. L. Hirst, *J.*, 1942, 70.

¹⁸ W. Kern in "Fortschritte der Chemie, Physik und Technik der makromolekularen Stoffe" (Berlin, 1939), Vol. I, p. 54.

¹⁹ "Der Aufbau der hochpolymeren organischen Naturstoffe" (Leipzig, 1930), p. 232.

²⁰ *Kolloid-Z.*, 1939, **88**, 51.

²¹ *Ibid.*, 1932, **59**, 82.

A property even more sensitive to molecular asymmetry is the orientation produced in a solution by shear. The extent to which a rod-shaped molecule can be orientated against the disorienting effect of thermal movement is proportional to the square of its length. Experimental information about the orientation produced can be obtained from observations on double refraction. Streaming double refraction has been a very powerful technique in the qualitative examination of many colloidal systems, a subject in which valuable pioneer work was done by H. Freundlich and H. Zocher.²² Several origins of streaming double refraction have been distinguished. It can arise purely from the elliptical deformation in a velocity gradient of otherwise spherical deformable particles, this cause being complicated by rotation of the deformed ellipsoids, and it can arise from the orientation in the velocity gradient of rigid anisodimensional particles. In the latter case the double refraction may be due both to the parallelisation of long particles themselves isotropic but having a different refractive index from the dispersion medium (pure rod double refraction) and to the long particles being themselves, as is generally the case, optically anisotropic. Pure rod double refraction may be eliminated by choice of a dispersion medium having the same refractive index as the particles.

The last-mentioned mechanism is in general much the most important and is capable of giving much higher double refraction than any other. Most of the later work has been done on this aspect and the reader is referred to the outstanding paper of J. R. Robinson²³ and the review by J. T. Edsall²⁴ for detailed information of the subject. As in the case of viscosity, recent researches have been mainly in the field of non-ionic macromolecules, and the Reviewer is not aware of any work on the complicating factors introduced by the necessary presence of small ions in the case of colloidal electrolytes, although interesting work has been done on aqueous solutions of the fibrillar proteins and especially on the giants of this class such as tobacco mosaic virus. One would expect that the electrolyte influence would be markedly different from that on the other properties mentioned.

Solubility

One of the most important facts about any substance in the dissolved state is that it has dissolved. In solubility we recognise two principal factors at work. If both components of the solution are liquid the forces of interaction between the like and the unlike pairs of molecules are mainly concerned. The smaller the energy of interaction of unlike molecules compared with the arithmetic mean of the energies of interaction of like molecules the less mutually soluble will the liquids be. Dipole moment is here of importance, but may not itself be of primary importance. More specific reactions such as hydrogen bonding may predominate, and spatial or geometrical factors may considerably modify the effect of dipole moment. Specific behaviour is more frequent among substances of high dipole moment,

²² See H. Freundlich, "Kapillarchemie" (Akademische Verlagsgesellschaft, 1930).

²³ *Proc. Roy. Soc.*, 1939, A, **170**, 519.

²⁴ "Advances in Colloid Science" (Interscience, New York, 1942), Vol. I, p. 269.

and particularly hydroxylic and amino-compounds, than it is among normal organic solvents, but even here behaviour is by no means adequately reducible to a simple formula. There has recently been much use made of the quantity "cohesive energy density" ²⁵—the latent heat of evaporation at constant volume per c.c. of liquid—as the dominant factor in solubility relationships, particularly of non-polar polymers. There is danger here of misleading over-simplification if this usage is extended to more complex polymers.

Even among simple substances the rule that mutual solubility is most favoured by approximate equality of the cohesive energy densities rides roughshod over the more exact requirement that, for high solubility, the interaction energies of unlike molecules must be comparable to, preferably higher than, the mean of the interaction energies of the like molecules. The latter energies need not be similar, although, for the rule to apply, they usually are. The cohesive energy densities for ethyl ether, carbon tetrachloride, and chloroform are 58, 76, and 87 cal./c.c., respectively. Not only is water, with the much higher value of about 500, much the most soluble in ether, having the lowest value, and least in carbon tetrachloride, having the middle value, but the same is markedly true of aromatic nitro-compounds. Further, chloroform and ether, as indicated by the vapour-pressure curves, are more mutually compatible than the other pairs although having the largest difference of cohesive energy density. Among polymers, cellulose acetate is almost completely unswollen by carbon tetrachloride (76) and chlorobenzene (92), but soluble in acetone (94), and very extensively swollen by methyl ethyl ketone (81).

These specific deviations from any simple generalisation are more frequent and more marked when highly polar, and particularly hydrogen-bonding, groups are involved. In aqueous solutions, therefore, we are likely to meet them in their most extreme form.

The second important factor in solubility comes into play when one component of the solution is crystalline. Geometrical factors become much more important, since they play at least an equally important rôle to the interaction energies in determining the melting point and heat of fusion. The solubility of a crystalline substance in a liquid with which it would, if itself liquid, form an ideal solution, is determined solely by these quantities and, on the simplest assumption, that the latent heat of fusion is not dependent on temperature, the mol.-fraction solubility N is given by ²⁶

$$\log_e N = \frac{\lambda}{R} \left(\frac{1}{T} - \frac{1}{T_f} \right) \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (3)$$

where T is the absolute temperature of measurement, T_f the melting point, and λ the latent heat of fusion. In a solvent with which the supercooled melt of the crystalline substance would not form an ideal solution, the solubility of the crystals will be appropriately modified.

A crystalline substance is therefore always less soluble than a chemically similar liquid substance and will be less soluble the higher its melting point

²⁵ G. Gee, *Quart. Reviews*, 1947, **1**, 265.

²⁶ J. H. Hildebrand, "Solubility" (New York, 1936).

and the higher its latent heat of fusion. As between different solvents, however, its solubility will vary in the same way as that of a chemically similar liquid.

Applied to salts, the theory of the solubility of crystalline substances, developed notably by Hildebrand, has made little progress, on account of their usually very high melting points, their very high degrees of dissociation, and the fact that they dissolve only in the least simple solvents and with very large activation energies. Considerable study has been made in recent years of the solvation of ions in solution,²⁷ but the application to solubility is as yet undeveloped. In some respects, the salts of organic acids and bases might provide better material for this development than the typical inorganic salts.

Applied to macromolecular substances, solubility theory has explained some important facts and been subject to considerable extension and modification in a specialised direction, but is still far from complete. Many polymers are amorphous in the solid state because they are stereochemically heterogeneous. Thus alternate C atoms in the polymethyl methacrylate chain are asymmetric. Since the inanimate preparation of the polymer cannot prefer *d*- or *l*-forms at each addition to the chain, the probability of even short lengths of neighbouring chains having identical configuration is remote. Similarly, in a partly substituted cellulose, the substituents will not be regularly distributed and the structural identities necessary for crystallisation will rarely occur. Such polymers behave as liquids in their solubility and may be completely or only partly miscible with simple liquids or be almost entirely without any interaction. When partial miscibility results, however, it has this peculiarity, not found in simple liquids, that the macromolecular substance is confined to one phase, not being present in detectable concentration in the other except under carefully chosen conditions.

The simple principle underlying this characteristic behaviour, and one with far-reaching applications, was first clearly stated by J. N. Brønsted.²⁸ For one macromolecule, of degree of polymerisation n , the heat of transfer from one phase to another in equilibrium with it will be approximately n times that for the effective monomeric unit. The translational energy will, however, be the same, kT . Application of Boltzmann's law would therefore predict that the partition coefficient for the polymer between these phases will be equal to that of the monomer raised to the power n . It is evident that if n is a large number and the monomer partition coefficient differs at all appreciably from unity, that for the polymer will become effectively zero or infinity, *i.e.*, the polymer will be found wholly in one phase or in the other. Fuller analysis of the problem has shown that, although Brønsted's treatment is certainly over-simplified, the prediction is essentially correct.²⁹

²⁷ R. W. Gurney, "Ions in Solution" (Cambridge, 1936).

²⁸ *Compt. rend. Lab. Carlsberg, Sér. chim.*, 1938, **22**, 99.

²⁹ P. J. Flory, *J. Chem. Physics*, 1945, **13**, 453; M. L. Huggins, *Ann. N.Y. Acad. Sci.*, 1942, **43**, 1.

The consequent sensitivity of polymers to the composition of solvent is well illustrated by the observation that, whereas ethyl laurate is a solvent for polystyrene, propyl laurate is only a limited sweller.³⁰ Another example of the same principle has been known for a long time to paint technologists in the incompatibility of many polymers even in dilute solution in a common solvent. This behaviour has recently been quantitatively studied, but the authors³¹ have advanced what is, in the Reviewer's opinion, a quite unnecessarily specialised explanation. It is understandable, on Brønsted's principle, that addition of 1% of cellulose acetate to acetone would so alter the partition coefficient for poly(methyl methacrylate) between that solution and pure acetone as to drive the latter polymer almost entirely into the pure acetone could the solutions permit the transfer and at the same time remain separate. They do remain separate because the latter polymer has a similar effect on the first. Two phases in fact separate, both dilute solutions, the one containing nearly all the cellulose acetate, the other nearly all the poly(methyl methacrylate).

Stereochemically regular polymers are nearly always crystalline, although the melting points may vary widely and the crystallinity is considerably affected by the history of the specimen and its present physical state—*e.g.*, crystallinity in rubber is increased by stretching. Since a crystalline simple substance has always a limited solubility in a simple liquid, we should expect, by the application of Brønsted's principle, that the solubility of crystalline polymers, except near the melting point, will be effectively zero. This is in fact the case, with this important qualification, namely, that polymeric substances are never wholly crystalline. Most of the long-chain molecules will fall into line with others around different crystal nuclei, and expansion of these nuclei, to more than one of which most chains will in part belong, leads to intervening regions of entanglement which cannot be further organised. Thus a mass of very small crystals results, materially united but organisationally divided by chains in an amorphous or random (indeed probably "more than random") condition. The crystals will not dissolve except near their melting point, but liquids which would otherwise be solvents will penetrate and swell the amorphous regions. Crystalline polymeric materials are therefore rarely soluble but are capable of limited swelling.

We have said that the application of solubility theory to electrolytes is in a very rudimentary state. What, if anything, therefore, of value can be said of the application of polymer solubility theory to colloidal electrolytes?

In the first place we may note that, for a pure colloidal salt, the necessary preservation of electrical neutrality in any phase requires that an equivalent number of counter ions ("gegenionen") must be present with the colloidal ion. As with osmotic pressure, sedimentation equilibrium, etc., this multiplies by the factor $z + 1$ (z being the valency of the colloidal ion) the concentration term in the free-energy equation.

Partition between two phases is not therefore to be expected to be so

³⁰ J. N. Brønsted and K. Volqvartz, *Trans. Faraday Soc.*, 1939, **35**, 576.

³¹ A. Dobry and F. Boyer-Kawenoki, *J. Polymer Sci.*, 1947, **2**, 90.

extreme as with neutral polymers. The significance of this is, however, rather obscured by the paucity of liquids which are good solvents for any electrolyte. Even for simple salts few solvents are comparable to water in solvent power and these do not include the ordinary organic solvents. If a good electrolyte solvent incompletely miscible with water, such as nitromethane, were used, it should be possible to realise a finite distribution of a colloidal electrolyte between this and water.

Most macromolecular electrolytes of Nature—gums, starch, proteins—have in the skeleton an effective predominance of hydrogen-bond-forming groups ($\cdot\text{OH}$, $\cdot\text{NH}_2$, $\cdot\text{CO}\cdot\text{NH}\cdot$) so that, apart from the enormous preference for water as a solvent due to the ionic side groups, the molecule is otherwise hydrophilic. Substantially different behaviour might be expected if ionised macromolecules were prepared with a substantially hydrophobic skeleton. This subject will be taken up again when we come to discuss the paraffin-chain salts.

It is noteworthy too that, while many simple organic salts are relatively much more soluble in organic solvents than in water, particularly in acetone and alcohols, and especially if wet, the increase of valency above 1 greatly depresses the organic solubility.

In the case of osmotic pressure, etc., the masking of colloidal properties in solutions of pure salts as a consequence of the condition of electrical neutrality, *i.e.*, the Donnan effect, is eliminated in the presence of sufficient simple salt. A corresponding influence is present in the case of solubility and partition phenomena, but is less readily recognised. A related and more familiar difference is between the behaviour of charged and uncharged proteins. An isoelectric protein (note, however, that an isoelectric protein can only be obtained entirely free from other simple ions if the isoelectric pH is 7) behaves much more nearly like a neutral macromolecular substance. It is either soluble indefinitely, swollen to a limited extent, or completely insoluble. It is precipitated from aqueous solution very sharply on slow addition of alcohol or other non-electrolyte precipitant, while the protein salt is thrown out in stages.

The fact that colloidal electrolytes are of very high valency introduces a new characteristic of their solubility. Expressed, as for simple electrolytes, in terms of the ionic product, we should expect, at saturation, for the salt RNa_z ,

$$[\text{R}^{z-}] \times [\text{Na}^+]^z = \text{constant} \quad . \quad . \quad . \quad (4)$$

A high value of z makes the saturation value of $[\text{R}]$ very sensitive to the value of $[\text{Na}]$, so the solubility will be rapidly depressed from a high value to a very low one by addition of simple sodium salt. As a consequence, in a high concentration of sodium chloride, the solubilities of different colloidal electrolytes will differ by much larger ratios than in the absence of other source of sodium ions. Addition of sodium chloride to a mixture of colloidal electrolytes would be expected therefore to throw out the colloidal sodium salts successively and with almost complete separation. This is the principle of a much-used method of fractionation of native

proteins, where, however, experience has shown ammonium sulphate to be preferable as the precipitant.

One might expect that, in the presence of simple salts in concentration inadequate to cause precipitation, the colloidal electrolyte would behave towards non-electrolyte additions in a way more similar to the behaviour of an un-ionised macromolecular substance. This prediction is, however, largely obscured by the incidence of a series of new phenomena. These consist in the separation, not only under these conditions but under others also, of a second liquid (often very viscous) phase from the solution. This phenomenon, first described for the system starch-gelatin-water by Wo. Ostwald,³² was later examined by H. R. Kruyt and H. G. Bungenberg de Jong,³³ by whom it was called coacervation, and made the subject of an extended series of investigations by the latter author and his collaborators.³⁴

Coacervation may occur, instead of the separation of a solid, on the salting out of a colloidal electrolyte by a simple electrolyte referred to above. Further addition of salt solidifies the coacervate. Among the many systems giving rise to coacervation, three other main types may be mentioned :

(1) Addition of a miscible neutral non-solvent (*e.g.*, alcohol) to an aqueous solution of uncharged hydrophilic colloid or of charged colloid in the presence of a low concentration of salt.

(2) As a stage in, or alternative to, the mutual precipitation of oppositely charged colloidal ions.

(3) By the addition of small amounts of salt having a very high-valent ion opposite in charge to the colloidal ion.

It would take us too far afield to discuss these interesting processes in detail and the reader is referred to Bungenberg de Jong's review.³⁴ The apparent similarity to the separation into two phases of a solution of incompatible neutral macromolecular substances in a common solvent, previously referred to,³¹ is, however, suggestive that a common factor is at work. Although, in the most important type of coacervation, (2) above, the separated phase is richer in *both* colloids than is the equilibrium liquid, whereas incompatible neutral colloids keep each to their own phase, the Reviewer considers that there is indeed a common factor, namely the "exaggeration" of partition coefficients for colloidal substances—the Brønsted effect.

Positive gelatin and negative gum arabic form a coacervate richer in both substances than the equilibrium liquid from which it separates. One may imagine that, could both substances be shorn of their charges without other change, they would be incompatible and, when aqueous solutions were mixed, separation would occur. The tendency to separate would of course increase greatly with increase of total concentration. In a solution sufficiently dilute for the macromolecules to come very rarely in contact,

³² *Kolloid-Z.*, 1927, **43**, 131.

³³ *Ibid.*, 1930, **50**, 39.

³⁴ H. G. Bungenberg de Jong, "La Coacervation" (Hermann, Paris, 1936); H. G. Bungenberg de Jong and E. G. Hoskam, *Proc. Acad. Sci. Amsterdam*, 1942, **45**, 59; H. G. Bungenberg de Jong, E. G. Hoskam, and L. H. v.d. Brandhof-Shaegen, *ibid.*, 1941, **44**, 1099.

there would be no tendency to separate. Since in fact both are highly and oppositely charged, in dilute solution the very high electrostatic attraction will predominate. The colloidal ions will therefore come together from high dilution, but, when concentrated, the shorter-range forces considered in the theory of solubility of neutral macromolecules will predominate, since they increase much more rapidly with decreasing distance of separation than do the Coulomb forces. Separation of two concentrated phases cannot occur, as this would violate electrical neutrality or require it to be balanced by the equivalent separation of simple counter ions with their much greater entropy. The only possible compromise occurs—the oppositely charged colloidal ions are brought much nearer together by Coulomb forces but kept apart at short range by the non-Coulomb forces. A mutual concentration, but falling short of mutual precipitation, results.

Coacervation of type (3) (above) does not seem to the Reviewer to involve any substantially different principle. Provided both ions be highly charged, it is not necessary for both to be very large. The gum ions, if neutral, would disperse in water. The short-range energy promotes dispersal. The Coulomb energy promotes concentration.

Whether, on addition of a non-electrolyte precipitant in the presence of salt, or of a precipitating concentration of salt, a coacervate results rather than a solid precipitate, is greatly influenced by geometrical factors in the colloidal substance.

If this has a regular molecular structure capable of crystal packing, the precipitate will always be solid or only very little swollen and only over a very narrow range of non-solvent concentration. The non-electrolyte, ordinary cellulose acetate (about 2.3 acetyl groups per glucose residue), being heterogeneous geometrically and non-crystalline, is precipitated as a gel by many non-solvents. Cellulose triacetate, geometrically regular and known to be crystalline, is precipitated in solid form. A coarse gelatinisation of the whole liquid may occur due to fine division of the precipitate, but the latter is itself but little swollen. Exactly comparably, gum arabic, having a complex geometrical structure and being non-crystalline,³⁵ can be precipitated as a coacervate, while calcium alginate, regular and crystalline,³⁶ precipitates sharply as a solid.

The extension to the more complex case of colloidal electrolytes of statistical theories of solubility of non-ionic macromolecular substances and of the thermodynamic properties of their solutions is a potential fruitful field of research awaiting development. We have discussed briefly above the nature of the chief complications which will be present in electrolytes. In conclusion, we may emphasise the necessity of remaining aware of a serious danger. In some circumstances the complications arising directly from the fact that our solutions are electrolytes can be eliminated—as, for example, in osmotic-pressure measurements, by working in excess of a simple salt. We have seen, for example, that, in the latter case, information can

³⁵ K. H. Meyer, "High Polymers" (Interscience, 1942), Vol. IV, p. 375.

³⁶ G. Lunde, E. Heen, and E. Oy, *Kolloid-Z.*, 1938, **83**, 196; H. Kringstad and G. Lunde, *ibid.*, p. 202.

be derived as precise as that obtainable about the molecular weight of non-ionic macromolecules in organic solvents—as precise, but no more precise. The deviations from ideal behaviour with non-ionic macromolecules are much greater than in simple solutions and important at far lower concentrations.³⁷ The same is true for colloidal electrolytes even when the major electrolyte effects have been eliminated. We must not let our preoccupation with electrolyte effects tempt us into neglect of these other factors.

Among the proteins, the importance for water solubility of the ionic groups, which are carried only on the side chains contributed by some of the amino-acids, is variable. Some proteins and related polypeptides carry a very large number of ionic groups, as, for example, the naturally-occurring polyamide of glutamic acid.³⁸ Others, for example silk fibroin, are relatively poor in ionisable groups, the chains being formed mainly of glycine and alanine. Nevertheless, silk fibroin, presumably because of the relative richness in amide groups due to the predominance of small amino-acids, is water soluble in the state in which it is produced by the silk glands. It has the remarkable property of becoming insoluble simply when the solution is subject to shear.³⁹ The insoluble body gives an X-ray crystal diagram and we can assume that the shearing process enables the mutual fitting of the complex molecules to occur. Conversion of soluble into insoluble forms is of very widespread occurrence in proteins although in no other case is the so-called denaturation achieved with such beautiful simplicity as in silk fibroin.

The importance of geometrical regularity, as a determining factor in crystallisability, is well emphasised by the insolubility of the necessarily regular polyglycine and the water solubility of the mixed polypeptide from glycine and alanine.⁴⁰ The addition of methyl groups, which must reduce the attraction for water, is more than offset by the irregularity produced by random distribution of different groups in the chain, which renders crystallisation impossible.

Electrochemical Properties

The fact that colloidal electrolytes are electrolytes, while responsible for many of the complications arising in the elucidation of the nature of their solutions, also provides us with an important new range of measurements.

Measurements of total electrical conductivity alone tell us disappointingly little about the size of the colloidal ion or aggregate, except under very favourable conditions. J. W. McBain⁴¹ was the first to point out that the aggregation of simple ions should increase their velocity in the electric field. If the total charge of the constituent ions of the aggregate remains unneutralised on aggregation, the force acting will be proportional to the weight of

³⁷ P. J. Flory, *J. Amer. Chem. Soc.*, 1943, **65**, 375; G. Gee, *Trans. Faraday Soc.*, 1940, **36**, 1171; G. Gee and L. R. G. Treloar, *ibid.*, 1942, **38**, 147.

³⁸ M. Bovarnick, *J. Biol. Chem.*, 1942, **145**, 415; see also ref. (71).

³⁹ W. Ramsden, *Nature*, 1938, **142**, 1120.

⁴⁰ E. J. Wilson and E. Pacsu, *J. Org. Chem.*, 1942, **7**, 117, 126.

⁴¹ *J.*, 1919, **115**, 1279; 1922, **121**, 2330; 1923, **123**, 2417.

the aggregate and therefore, approximately, to its volume, whereas the resistance to motion, by Stokes's law, is proportional to a linear dimension only.

In their investigations of soap solutions, McBain and his collaborators established, among many interesting facts, that the equivalent conductivity increased with increase of concentration in the range from about 0.1 to 0.5N. This was unfortunately attributed to the direct effect of ion aggregation just referred to. Later researches have shown that the explanation of this peculiarity must be sought on different lines, and that ion aggregation in these solutions is already fully established at a much lower concentration, where it is associated with a sharp fall of equivalent conductivity. A rising value similar to that in soaps but less marked has been reported for gum arabic and casein salt solutions.^{42, 43} Before returning to this interesting peculiarity, we will consider other aspects of conductivity, particularly in the dilute solutions where interpretation of the results is less uncertain.

The mobility of a highly charged colloidal ion would, in general, be very high at infinite dilution, since it is the ratio of charge to linear dimension which determines it. The equivalent conductivity at infinite dilution might therefore be expected to exceed that of most simple electrolytes. In the case of some aggregating electrolytes this effect has been directly observed.⁴⁴ In opposition to this direct effect of high charge to radius ratio, however, there is that of the increased attractive forces between colloidal and counter ions which increases both the breaking effect of the free ions on one another (the relaxation and endosmotic effects of the Debye-Hückel-Onsager theory) and the tendency for counter ions actually to adhere to the surface of the colloidal ion.

At infinite dilution, therefore, the conductivity of colloidal electrolytes may be abnormally high, but the fall of conductivity with rise of concentration will be so much greater than in the case of simple salts that, in the range of concentration accessible to accurate conductivity measurement (or at least in the range where measurements have so far been made), the conductivity is more usually below that of corresponding simple salts.

While the osmotic coefficient is restored to unity in presence of a swamping excess of simple salt, the reduction of the mobility below the infinite-dilution value is maintained and increased. Of the two parts of the interionic effect on mobility, as analysed by L. Onsager,⁴⁵ it is, however, probable that the relaxation effect is small and the endosmotic effect still fairly simply calculable.^{46, 47} On this assumption, the mobility of the colloidal ion can be shown to be directly proportional to the surface potential, as is that of a large particle according to the Smoluchowski theory. We shall refer to this identity later.

⁴² D. R. Briggs, *J. Physical Chem.*, 1934, **38**, 867.

⁴³ G. Schmidt and E. C. Larsen, *Z. Elektrochem.*, 1938, **44**, 651.

⁴⁴ G. Robinson and H. E. Garrett, *Trans. Faraday Soc.*, 1939, **35**, 771.

⁴⁵ *Physikal. Z.*, 1926, **27**, 388; 1927, **28**, 277.

⁴⁶ D. C. Henry, *Proc. Roy. Soc.*, 1931, **A**, **133**, 106.

⁴⁷ G. S. Hartley, *Trans. Faraday Soc.*, 1935, **31**, 31.

The Debye-Hückel-Onsager theory for simple electrolytes predicts that, in general, the mobility of a multivalent ion in an unsymmetrical electrolyte is reduced more rapidly by increase of concentration than is that of univalent ions. The transport number of the multivalent-ion constituent will therefore decrease with increase of concentration. If, however, the Coulomb forces, with or without the assistance of short-range forces which could then come into play, cause some of the counter ions actually to adhere to the colloidal ion, the effect on transport number will be reversed. The initially z -valent ion will have its mobility reduced, by the adherence of one counter ion, by approximately $1/z$ th of its initial value. The average mobility of the z counter ions would be similarly reduced were this one ion to remain stationary. Since actually it will be carried in a direction opposite to its direction if free, the reduction of average mobility will be by *more than* $1/z$ th of the initial value.

Transport-number determinations can therefore give us some information on this important point. It must, however, be borne in mind that the different effect of an attached counter ion from that of a unit charge in the atmosphere is not likely to be arrived at by a sudden transition. Other entirely different types of measurement must be called into use, as has been done in the study of association of simple ions as previously mentioned.³ The Reviewer has elsewhere⁴⁸ drawn attention to this possibility in one case. Cetylpyridinium iodide, even when free iodine is completely eliminated, gives yellow solutions in non-ionising organic solvents, *e.g.*, benzene. In ionising organic solvents, *e.g.*, wet alcohol or acetone, the salt is colourless. In water it gives pale yellow solutions. In aqueous solutions, except in extreme dilution, the paraffin-chain ions are aggregated into very highly charged colloidal units, while in ionising organic solvents no aggregation occurs. The yellow colour in water can therefore be taken as evidence that some of the iodide ions are very closely attached to, or included in, the colloidal aggregate. The phenomenon would be worth quantitative investigation, and extension into the ultra-violet would open up a wider field of application.

From transport-number and conductivity determinations, the mobility of the colloidal ion can be determined. In the case of aqueous solutions of paraffin-chain salts⁴⁹ and of certain dyes,⁵⁰ the application of this method has greatly extended the meagre information which could be derived from conductivity measurements alone. Thus, it has been shown that, in the case of aqueous solutions of paraffin-chain salts, the mobility of the paraffin-chain ion constituent increases abruptly where the total equivalent conductivity falls abruptly. Since all normal concentration effects on mobility are depressant, the only possible explanation of a rise of mobility with increase of concentration, to values well above the infinite-dilution value,

⁴⁸ *Kolloid-Z.*, 1939, **88**, 33.

⁴⁹ B. Collie, C. S. Samis, and G. S. Hartley, *Trans. Faraday Soc.*, 1936, **32**, 795; C. S. Samis and G. S. Hartley, *ibid.*, 1938, **34**, 1288; O. R. Howell and H. Warne, *Proc. Roy. Soc.*, 1937, *A*, **160**, 440.

⁵⁰ C. Robinson and J. L. Moilliet, *ibid.*, 1933, *A*, **143**, 630.

is the decreased resistance to motion offered by a cluster as compared with that offered by the ions separately, *i.e.*, the effect predicted by J. W. McBain⁴¹ but erroneously applied at that time to the rise of the total equivalent conductivity occurring at much higher concentrations.

Transport-number measurements should always be made together with conductivity measurements whenever practicable. Logically the most accurate technique, that of the moving boundary, has been developed to be capable of very precise measurement by MacInnes, Longsworth, and their collaborators.⁵¹ The method has been extended to be applicable to cases where the mobility may be low by E. Drew, B. Collie, and G. S. Hartley.⁵² A much simpler form of apparatus for the normal case than that of MacInnes and Longsworth has been described by G. S. Hartley and G. W. Donaldson.⁵³ Application to macromolecular electrolytes, such as alginate, gums, and synthetic substances, such as polyacrylic acid and its salts, would be of great interest.

The simplification in the moving-boundary transport-number apparatus above referred to is made possible by the self-sharpening of the boundary between two well-chosen pure electrolytes under the influence of the applied potential gradient.⁵⁴ It is not therefore applicable where it is desired to make measurements in the presence of excess of simple electrolytes. The boundary in these cases is never so sharp and can only be maintained sufficiently sharp for quantitative observation if it is formed sharply by suitable mechanical means to start with. Formation of the boundary by shearing of two glass plates with a suitable disposition of holes, as is done in the apparatus of MacInnes and Longsworth, is here necessary. Application of an apparatus of this type, with refractometric or ultra-violet absorptiometric measurement of concentration-distance function, has been much used by A. Tiselius⁵⁵ for diagnosing the purity of proteins, analysing mixtures of proteins, and even for effecting fractionation on a small scale. The mobilities of different proteins are widely different, not only because of variations in the number of ionic side groups, but more especially because of the varying distribution of weak acid and basic groups. Mobility is therefore very sensitive to pH and zero mobility (isoelectric point) occurs at different pH values for different proteins.

In the case of proteins near the isoelectric point, we have the possibility of studying the mobility of large ions carrying a relatively small charge. If this charge can be determined, conditions are most favourable for the application of the Debye-Hückel theory. The atmosphere can be assumed to be built up only by deviations from bulk concentration of the uni- or bi-valent ions of the buffer solution, the equivalent concentration of colloid being relatively very small. A. Tiselius and H. Svensson⁵⁶ have in this

⁵¹ L. G. Longsworth, *J. Amer. Chem. Soc.*, 1935, **57**, 1185; D. A. MacInnes, I. A. Cowperthwaite, and K. C. Blanchard, *ibid.*, 1926, **48**, 1909.

⁵² *Trans. Faraday Soc.*, 1934, **30**, 648.

⁵³ *Ibid.*, 1937, **33**, 457.

⁵⁴ D. A. MacInnes and L. G. Longsworth, *Chem. Reviews*, 1932, **11**, 171; G. S. Hartley and J. L. Moilliet, *Proc. Roy. Soc.*, 1933, **A**, **140**, 145.

⁵⁵ *Trans. Faraday Soc.*, 1937, **33**, 524.

⁵⁶ *Ibid.*, 1940, **36**, 16.

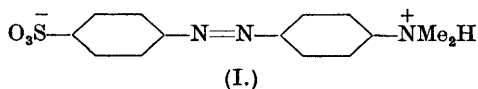
way found very satisfactory agreement between their results for egg albumin in acetate buffer and the theory of D. C. Henry.⁴⁶

The charge on the protein molecule in the presence of ions of a buffer solution can be measured by means of membrane experiments by the method worked out by G. S. Adair and M. E. Adair.⁵⁷ Determinations of the osmotic pressure, the concentrations of the various simple ions, and of the membrane potential enable the charge to be calculated with considerable precision if not too large. The membrane potential is the difference of potential between two liquids to one of which the protein is confined by the membrane. It arises from the tendency of the counter ions to diffuse through the membrane until arrested by the potential. Even where the charge is low, comparison of values from analytical data with those from membrane potential indicate that ions other than hydrogen may be attached to the protein molecule.

The colloidal ion is both highly charged and very large. The ionic groups on its surface are not necessarily very close together. To take gum arabic as an example, its molecular weight has been estimated as about 270,000^{58, 59} and its equivalent weight as *ca.* 1200.^{58, 42} Its partial specific volume in aqueous solution is 0.6.⁵⁹ If the colloid ion were spherical it would have therefore a radius of about 40 Å. and, if the ionic groups were situated entirely in the surface, their mean distance apart would be some 10 Å. In fact, the particle is considerably elongated and the mean distance of separation of charges will therefore be greater. In the paraffin-chain salts it will be rather less, but probably a greater fraction of the resident ions will be neutralised by attachment of counter ions.

Some comparative distances are of interest. In a 1N-solution of a uni-univalent electrolyte, the distance apart of the ions, assumed to be distributed in a body-centred cubic lattice pattern, will be *ca.* 12 Å. The radius of the ionic atmosphere according to the simple Debye-Hückel theory will be *ca.* 3 Å. In a 0.001N-solution the values will be *ca.* 120 Å. and *ca.* 100 Å., respectively.

Some interesting conclusions about spaced charges have been drawn by N. Bjerrum,⁶⁰ who showed that



a long "zwitterion" such as that of methyl-orange in acid solution (I), behaved as a neutral molecule in that its solubility in water was relatively unaffected by change in salt concentration, if small. At higher salt concentration the solubility was increased by further addition of salt. So long, therefore, as the ionic atmosphere radius is much greater than the spacing between the opposite charges, these charges cancel one another. At higher salt concentrations the atmosphere radius is so far reduced that the two ions tend to build up each their own atmosphere and a modified

⁵⁷ *Biochem. J.*, 1934, **28**, 199, 1230.

⁵⁸ H. B. Oakley, *Trans. Faraday Soc.*, 1935, **31**, 136.

⁵⁹ S. Saverborn, in "The Svedberg" (Uppsala, 1945), p. 508.

⁶⁰ *Z. physikal. Chem.*, 1923, **104**, 147.

salt effect therefore comes into play. Similarly, one might expect that a bivalent cation or anion with separated charges would be influenced as a bivalent ion by salts in low concentration and tend to be influenced more as a univalent ion at higher concentrations.

Extension of this argument to a colloidal ion would lead us, however, to a quite false conclusion if we were to estimate the effect on any one ionic group of its immediate neighbours only. Let us consider the case of the colloidal aggregate present in dilute solutions of cetylpyridinium acetate, which we know to be approximately spherical, of radius 24 Å. and containing therefore about 80 cetylpyridinium ions. The charges can be regarded as uniformly distributed over the surface of the sphere. Their combined effect therefore at a point outside the sphere will be that of a point charge of the same total magnitude situated at the centre of the sphere. The distribution of free ions around any one ionic group in the surface we can best consider therefore to be due to combined influence of a charge of 79 electrons acting from the centre of the sphere and 1 electron acting at the location of the selected charge. If all 80 charges were free and the aggregate could exist at infinite dilution, the values of the electric potential ψ at distances r in Å. measured radially outwards from the selected ion would be, in units of ϵ/D where ϵ is the electronic charge and D the dielectric constant, those given in the table below. Corresponding values of $-d\psi/dr$ and $d^2\psi/dr^2$ are also given.

$r =$	0.		2.		5.		10.	
	Ion.	Aggregate.	Ion.	Aggregate.	Ion.	Aggregate.	Ion.	Aggregate.
ψ	∞	3.3	0.5	3.0	0.2	2.7	0.1	2.3
$-d\psi/dr$	∞	0.13	0.25	0.12	0.04	0.09	0.01	0.07
$d^2\psi/dr^2$	∞	0.01	0.25	0.009	0.008	0.006	0.002	0.004

When the aggregate exists in a finite concentration of ions, these values will all be reduced, and to a varying extent. The values of $d\psi/dr$ can in principle be calculated by describing spherical surfaces passing through the point considered with centres at the centres of the selected ion and the aggregate respectively and considering the contained charge of each sphere to act as from its centre. The contained charges will be the sum of the central charge and the charge (opposite in sign) in that part of the atmosphere within each sphere.

The fraction of the central charge left unneutralised by a layer of the atmosphere of certain thickness will, according to the simple Debye-Hückel theory, be independent of the central charge and dependent only on the charges and concentrations of the ions in the atmosphere. For an aggregate present in a swamping excess of simple uni-univalent salt, therefore, the contributions to $d\psi/dr$ will be reduced in the same proportion for the selected ion and the residual charge of the aggregate. Since $d\psi/dr$ falls off much more slowly with increase of r for the aggregate than for the ion,

the contribution of the aggregate to the integrated value for ψ at the ion surface will be much more reduced by the increase of salt concentration than will the contribution of the ion. In a N -solution of a uni-univalent electrolyte, where the effective atmosphere thickness according to the simple theory is *ca.* 3 Å., it can be shown⁴⁷ that $d\psi/dr$ will be reduced to half at 5 Å. distance. In this sense the effect of the aggregate charge on further changes of ψ due to increase of salt concentration will be small in high salt concentrations, but this is only true when the greater part of the ψ contribution of the aggregate charge at infinite dilution has been eliminated. At infinite dilution the contribution of the aggregate charge is much the larger. We may anticipate that further changes of potential in high salt concentrations will therefore be only of the same order as occur in simple salts. The total fall of potential from the infinite dilution value will, however, be very much greater.

Some other implications of the quantities given in the above table are of interest in getting a sense of proportion about the nature of a colloidal electrolyte solution. It will be noted that it is only in the values of the potential ψ itself that the contribution of the aggregate charge is much greater than that of the individual charges considered separately. It is this potential which determines the average prevailing concentration of counter ions in the immediate vicinity of the ions on the particle surface, in relation to the prevailing bulk concentration. Representing the surface concentration by c_s and that in bulk by c_b , the Boltzmann equation gives

$$c_s = c_b e^{e\psi/kT} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

At an assumed distance of closest approach of 2 Å., values of the exponential factor, for infinite dilution, will be 33 for the single ion alone and *ca.* 10^{10} at the surface of this ion due to the residual charge on the aggregate! A solution nominally $10^{-6}N$ in electrolyte would provide a solid layer of counter ions at the particle surface. This absurdity is resolved in two ways. First, it means that we cannot expect to approach infinite dilution as far as Debye-Hückel effects are concerned at any accessible concentration (even for a colloid with a stable particle) if the charge/radius ratio is as high as in the example taken (gum arabic, *if spherical*, would give an even greater ratio, despite its higher equivalent weight). Secondly, definite attachment of counter ions to a number of the resident ions is inevitable at finite concentrations. In the example quoted, probably about one-third of the charge is neutralised in this way, which will reduce the exponential factor in (5) to $10^{3.3}$ even in the absence of Debye-Hückel effects.

Despite this enormous direct effect of surface potential, the aggregate contributions to the values of $d\psi/dr$ and $d^2\psi/dr^2$ are much less at all near distances than the single ion contributions: $d\psi/dr$, the field strength, is the determining factor in the orientation of dipoles in the solvent; $d^2\psi/dr^2$, the divergence of the field, is the determining factor in attraction of oriented dipoles. As far, therefore, as hydration of the colloidal ion is caused by purely electrostatic forces, we can confidently predict that the high charge, as a whole, is quite unimportant. There is no reason at all to expect that

the colloidal ion as a whole will be heavily hydrated. Only the individual resident ions will be important in this respect, collecting no larger water sheaths than they would if existing separately in a simple salt solution. So much is loosely made in the literature of the very high degree of hydration of lyophilic colloids and its great importance that it is desirable to be emphatic on this point. A gum arabic ion with some 220 carboxyl groups upon it will certainly carry no more water, electrostatically attached, than would 220 separate acetate ions. It will probably carry less, because many of these carboxyl groups will be undissociated in the acid gum solution or associated closely with sodium ions in a sodium gum solution. Hydrogen-bond hydration of the hydroxyl groups in the gum will again be no greater because the molecule is a macromolecule. For the rest, the gum ion will be no more hydrated than will say polystyrene be solvated in a benzene solution. When gum is precipitated from water by alcohol there is no more need to speak of dehydration as an essential and peculiar part of the process than there is to speak of "debenzenation" when the polystyrene is precipitated from benzene by the addition of hexane. Water is a peculiar solvent but there is no need to consider it especially peculiar when it is behaving in a very ordinary way.

The flattening-out of the potential-concentration curve in higher concentrations is reflected in the equivalent conductivity-concentration curve. We have seen that, in soaps⁴¹ and other paraffin-chain salts^{49,61} and other colloidal electrolytes,⁴³ the equivalent conductivity may actually increase with concentration over a limited range. That this is due to aggregation of the colloidal ion is rendered unlikely by the fact that, in the paraffin-chain salts, the main process of aggregation occurs at much lower concentrations, where it is marked by a sharp fall of total equivalent conductivity although the paraffin-chain mobility increases. Moreover, the explanation is hardly applicable to gum arabic. T. R. Bolam and A. K. M. Trivedi⁶² have described experiments on a stable Odén sulphur sol where there could be no question of change of aggregation during the experimental period and found that (on dilution) the specific conductivity decreased more rapidly than concentration. Transport-number measurements on the paraffin-chain salts⁴⁹ show that where the total equivalent conductivity rises, the mobility of the paraffin chains is still falling, although slowly. It is the average counter ion mobility which increases and more than compensates for this continued fall. Consistently with this, we find the rise more marked the more mobile (at infinite dilution) are the counter ions concerned.

This curious phenomenon has been discussed by the Reviewer elsewhere,⁴⁸ and the suggestion put forward that the most probable explanation was one proposed to him by C. W. Davies. In pure solution, the colloidal ions and their atmospheres, predominantly of counter ions, will tend to form separate locally neutralised systems. Within these systems conductivity will be higher than between them. Decreasing separation will establish

⁶¹ A. Lottermoser and F. Puschel, *Kolloid-Z.*, 1933, **63**, 175.

⁶² *Trans. Faraday Soc.*, 1943, **39**, 247.

better contact between them. On this view we should expect that the conductivities of a colloidal and a simple electrolyte will be more than additive over a limited range of composition. This has in fact been observed.⁶³

An alternative explanation is suggested by C. W. Davies's own estimation of the true degree of dissociation in simple electrolytes of unsymmetrical valence type.⁶⁴ He concludes that there will be a minimum degree of dissociation, the ordinary mass-law effect being more than offset by the decreased activity coefficient due to interionic forces. These explanations are really only distinct in so far as we can make a valid distinction between real association of ions and their very close proximity. Measurements by an optical method in a suitable system such as the solution of cetylpyridinium iodide (p. 169) would be the most definitive possible. The possibility of interionic effects outweighing the effect of increase of bulk concentration as far as ion association is concerned will become clearer in the discussion immediately below.

The potential at the surface of an ion is a determining factor in ionic equilibria, such as the dissociation of weak acids. Its variation with concentration of electrolyte is responsible for the variation of the apparent dissociation constant. The latter variation is usually formally eliminated by calculating a correcting factor—the activity coefficient—which varies with salt concentration. The Debye-Hückel theory analyses the effect of the non-random distribution of ions on the potentials at their surfaces and then calculates the values of activity coefficient from the predicted variation of potential. For the treatment of ionic equilibria this is a very roundabout procedure. The mass law cannot be expected to apply to the concentrations of ions in bulk because any chosen ion necessarily reacts only with the ions within its reach. It is the concentration of these ions prevailing at its surface which should be used in the mass-law equation. This is directly derived by the Boltzmann law from the bulk concentration and the surface potential [equation (5)]. G. S. Hartley and J. W. Roe⁶⁵ have shown that to treat the Debye-Hückel prediction of the surface potential in this way gives exactly the same result as the elaborate method of calculating the activity coefficient by a hypothetical discharging process.

The dissociation constant to be used in the surface-concentration equation is, however, different from that obtained from the normal calculation, extrapolated to infinite dilution, because, even at infinite dilution (indeed, more so at infinite dilution, since the potential is reduced in finite concentrations), the concentration of, say, hydrogen ions at the surface of weak acid anions is greater than that in bulk. We may apply the principle to a weak acid group in a colloidal ion, by an extension of the argument by which N. Bjerrum⁶⁶ applied it to the effect of the first dissociation of a dibasic acid on the second dissociation constant. In this case we start from the ordinary dissociation constant and need not be concerned with the relatively

⁶³ J. W. McBain and J. Searles, *J. Physical Chem.*, 1936, **40**, 493.

⁶⁴ *J.*, 1945, 460.

⁶⁵ *Trans. Faraday Soc.*, 1940, **36**, 101.

⁶⁶ *Z. physikal. Chem.*, 1923, **106**, 219.

small salt effect on this in a simple electrolyte solution. The potential at the surface of the colloidal ion, due to the whole assembly of charges and the distributed ions in the atmosphere, is directly calculable from the mobility of the ion in an electric field by the well-known Helmholtz formula. This potential would not be materially altered by removing any one ion and therefore determines the local concentration of ions from which the selected ion builds up its own atmosphere. If a weak acid anion is resident on the surface of a colloidal ion whose potential, calculated from mobility, is ξ , it will build up its local atmosphere from a hydrogen-ion concentration

$$[H^+]_s = [H^+]e^{-e\xi/kT} \quad . \quad . \quad . \quad . \quad (6)$$

where $[H^+]$ is the bulk hydrogen-ion concentration.

The equilibrium condition governing the dissociation of the acidic group resident on the colloidal ion is therefore

$$K[HA]/[A^-] = [H^+]e^{-e\xi/kT} \quad . \quad . \quad . \quad . \quad (7)$$

where K is the ordinary dissociation constant.

We may put this in a more convenient form by defining the apparent new pK of the acid as the actual bulk pH at which the acid in the colloidal form is half dissociated. If this value is $(pK)_s$ and the normal value to be expected for a corresponding acid in simple solution is pK , then, by (7)

$$(pK)_s = pK - \frac{e\xi}{kT} \cdot \log_{10} e \quad . \quad . \quad . \quad . \quad (8)$$

For water at 25° this becomes

$$(pK)_s = pK - 0.023l \quad . \quad . \quad . \quad . \quad (9)$$

or

$$= pK - \xi/60 \quad . \quad . \quad . \quad . \quad (10)$$

if the mobility, l , of the colloidal ion is measured in ordinary units (ohm^{-1} l.g.-equiv.⁻¹) or if its potential ξ is calculated in millivolts.

G. S. Hartley and J. W. Roe⁶⁵ tested this equation quantitatively for the dissociation of an adsorbed indicator in a paraffin-chain salt solution. Its significance for titration curves of colloids containing weak acid and basic groups is considerable. D. R. Briggs,⁴² for example, obtained dissociation "constants" for acid gum arabic ranging from $pK = 3$ between 0.05 and 0.10N-solution to an extrapolated value of 6.7 at infinite dilution. He attributed this to reversible aggregation of the colloid, although how aggregation could *strengthen* the acid he did not explain. The behaviour is exactly what we should expect on the basis of the above discussion, and the difference would be accounted for quantitatively by equation (9) if the infinite dilution mobility of the anion were some 160 units higher than its value in the higher concentration range quoted. This is not an improbably high value, since the mobility of a half-ionised molecule containing 220 carboxyl groups and having an effective radius of 42 Å. would, by Stokes's law, be about 230. This figure is quoted merely to indicate that the explanation is quantitatively adequate. The extrapolation used to get the infinite dilution value of the dissociation constant is subject to great error, and the calculation of the constant is necessarily arbitrary in some

degree because the potential of the colloid ion must change during dissociation and hence the apparent dissociation constant must have a large spread at any concentration.

The effect of the varying potential in spreading out the dissociation "constant" has a disturbing effect on potentiometric titration curves of weak acid or basic colloidal electrolytes and the amphoteric proteins. Not only will the various dissociation steps be much less sharply differentiated than in simple solutions but the estimated pK values will be different from those anticipated from the behaviour of corresponding simple acids or bases. Both difficulties can be overcome if the titrations are carried out in considerable excess of a simple electrolyte where this is not made impracticable by precipitation.^{66, 67} The pK values for the constituent groups must, of course, in this case be determined on the corresponding simple compounds in the same high salt concentration.

Aggregation in Colloidal Electrolytes

We have not so far been concerned explicitly with the question of how the large, multivalent ion of colloidal electrolytes is produced, but rather with its properties on the assumption that it is a stable unit. There is no doubt that in many cases the colloidal ion is a single chemical unit entirely knit together by covalent bonds—in short a macromolecule analogous to cellulose, the hydrocarbon constituent of rubber, or the synthetic polymers.

We have seen that a macromolecular substance will not dissolve in a poor solvent—*i.e.*, a solvent which would not show a high degree of thermodynamic compatibility with the equivalent monomer. The interaction between solvent and the constituent units of a macromolecule is therefore high. These units prefer dispersal in the company of solvent molecules to crowding in their own. It follows that aggregation of macromolecules in solution is improbable. Normally, if it occurs, it will occur completely, the macromolecular substance separating in a separate phase. There may, however, be exceptions, but they would require a special explanation.

This cannot be a residual microcrystallisation. Since a macromolecular substance and a liquid which would otherwise be a solvent do not form a solution if the macromolecules have largely crystallised we might envisage an intermediate state of affairs where the penetration of the solvent into the amorphous regions of the macromolecular matrix has proceeded so far as to dissolve these regions entirely. In solution, then, we might have microcrystalline aggregates of macromolecules brought into solution by the interaction of solvent with the free fringes of uncrystallised macromolecules.

If the microcrystals themselves were not soluble, larger crystals built on the same pattern must be even more insoluble. The fringes of uncrystallised macromolecules would be free to come into position to continue the crystal pattern. The microcrystals would thus grow at the expense of the free fringes, a process not possible in the solid macromolecular substance

⁶⁷ W. E. Hanby and H. N. Rydon, *Biochem. J.*, 1946, **40**, 302.

on account of the irreconcilable requirements of neighbouring crystallites. The action of solvent would thus be to make the substance still less soluble, because more highly crystalline. This does in fact occur in the extreme case of a crystallisable polymer obtained in the amorphous solid condition by "shock cooling". On being placed in a "solvent" vapour it becomes crystalline.⁶⁸

A macromolecular substance might, however, be brought into solution as microcrystalline aggregates were the free fringes of protruding uncrystallised molecules chemically modified by an irreversible (or effectively so) process. In a fresh preparation of viscose this state of affairs may exist temporarily, since the free fringes may consist of xanthated cellulose and the microcrystals be unmodified cellulose.

Limited aggregation of dissolved macromolecules will therefore only occur if these molecules are inhomogeneous. Moreover, the differently constituted parts of molecules must be of sufficient length and sufficiently widely separated to permit them to be involved in differently organised regions of the solution. If this condition is fulfilled it may not be necessary for any of the parts of the macromolecule to be crystallisable—*i.e.*, stereochemically homogeneous. For example, polyvinyl alcohol is highly swollen by water. If, during its preparation by hydrolysis of the acetate, a large fraction of the acetyl groups are left in the molecule, a product less swellable in water is obtained, the swelling being less the higher the acetyl content. If the distribution of hydroxyl and acetyl groups could be other than random (or nearly so), there being regions of many groups of one kind in succession, the necessary condition for aggregation of the molecules might arise. Amorphous aggregates, essentially of vinyl acetate, would be held in solution in water by the free water-attracting fringes of vinyl alcohol structure. Equally possible, depending on the distribution, would be a resulting complete insolubility, neither water nor a vinyl acetate solvent being able to break up both types of cohesion which would exist in the solid state.

Whether any known macromolecular substance is dissolved, but incompletely so, by this process, is very doubtful. Diagnosis would be far from simple. The elaborate structure of some natural macromolecular substances may make use of this principle, but the incidental complexities may obscure it. Aggregation may occur in certain cases on addition of a non-solvent to a solution, as a reversible or transient phenomenon, just before the stage of precipitation is reached. The solvent-solute interaction will here be the feeblest consistent with maintained solution, and the influence even of random heterogeneity may be effective.

In the case of colloidal electrolytes, the possibility of a limited aggregation of the colloidal ions seems even more remote, because the long-range Coulomb repulsion will be added to the dispersive effects of short-range forces. In the present state of our knowledge it is reasonable to assume that, in solutions of the macromolecular electrolytes—gums, pectin, alginates, proteins, cellulose modified by ionogenic substitution—the state of dispersion is ultimate. The units are the covalently bonded molecules. Certainly no

⁶⁸ W. O. Baker, C. S. Fuller, and N. R. Pape, *J. Amer. Chem. Soc.*, 1942, **64**, 778.

clear evidence of definite grouped aggregation, as distinct from felt-like interlocking, has ever been produced.

In those electrolytes which are colloiddally dispersed as a result of aggregation, we are dealing with very much smaller ultimate molecules. No abnormal behaviour (except in many cases interfacial) would result in the absence of aggregation. The most important compounds of this class are the soaps and the now very numerous synthetic compounds of similar properties. Here, there is no doubt that the primary cause of aggregation is the inhomogeneity of structure envisaged above as a possible mechanism for the aggregation in macromolecular solutions. Its operation in much smaller molecules is effective because it occurs in a very extreme form—the combination in one and the same molecule of an ionic “ head ” having, with water, the strongest known solute-solvent interaction, and a paraffinic “ tail ” having the least possible interaction with the solvent (relative to the mutual attraction of the solvent molecules). This type of structure, sometimes referred to as polar-non-polar, is so fundamental to the behaviour of these compounds that it is convenient, as the Reviewer has elsewhere suggested,⁶⁹ to give to it a special name—amphipathy. If the greater part of the molecule has a chemical composition to which water is in effect antipathetic, but contains a group for which water has a very strong sympathy, we may call the resulting combination “ amphipathic ” towards water.

The behaviour of an amphipathic molecule towards water may result from an averaging of the opposing interactions. If, on the other hand, the molecule is large enough and the hydrophilic portion is situated at an extremity, a special organisation of the solution may arise, satisfying as far as possible the opposing tendencies separately and simultaneously. The paraffin tails may be huddled together away from the water, leaving the ionic heads in contact with the water on the outside of the resulting aggregate.

Our present knowledge of the nature of an aggregated solution of a paraffin-chain salt of the simplest type, where the paraffin chain is unbranched and saturated and the ionic group situated at the end, has been discussed by the Reviewer elsewhere^{48, 69} and by N. K. Adam.⁷⁰

It was established that, at least in dilute solutions of a salt which does not give an abnormally viscous solution, (1) the aggregate has a well-marked optimum size (corresponding to a content of about 80 paraffin-chain ions in cetylpyridinium chloride solution); (2) it is formed from the single ions with very low concentration of intermediate aggregates; (3) it confers on the solution a very general solvent power for non-polar substances, and (4) its constituent ions are packed more loosely than in the crystalline state.

These facts are explained by the assumption that the aggregate is approximately spherical and essentially liquid, the interior containing paraffin chains in as random a distribution as is consistent with the ionic

⁶⁹ “ Aqueous Solutions of Paraffin Chain Salts ” (Hermann, Paris, 1936).

⁷⁰ *Ann. Reports*, 1936, **33**, 103.

heads lying entirely on its surface, thus securing the maximum separation of paraffin part from the water while leaving the ions free to interact with the water. That this aggregate is preferred to smaller ones is due to the latter leaving more paraffin-water contact (energetically less satisfactory) or having a higher degree of organisation (an *a priori* less probable arrangement than reduction of size alone would suggest). That this aggregate is preferred to larger ones is due to the latter having a not greatly reduced area of paraffin-water contact (energetically not much more satisfactory) and requiring a more elaborate organisation in addition to increase of size, since the ionic heads could not remain in the surface unless the aggregate became anisometric (an *a priori* less probable arrangement on account both of increased size and of organisation).

No other picture has given a satisfactory explanation of all the established facts. In particular, that popularised by X-ray investigators, of a series of platelet aggregates with the paraffin chains parallel and separated by ionic and water layers,⁷¹ does not explain the density of the aggregate being less than that of the crystal (the increase of volume is about that which takes place when a solid paraffin melts) or the general solvent properties of the aggregate. Most important of all, it does not provide any mechanism for limitation of the size of the aggregates. This is known to be limited and in strict equilibrium.⁷²

This latter point is one of paramount importance in considering any kind of reversible association. When, from simple molecules or ions, aggregates of colloidal dimensions are created and maintained as a stable state of the system persisting over a wide range of concentration, we must discover not only a cause for the aggregation but also a cause for its limitation.

Recent work in the field has been concerned mainly with the properties of structurally more complex salts, with the solvent power of the aqueous solutions for materials but little soluble in water, and with X-ray examination of concentrated solutions. Developments, particularly in the latter field, have popularised a more elaborate picture of the structure of the concentrated solutions than was proposed by the Reviewer for the dilute solutions. The validity of the conclusions about the latter has consequently become suspect. The Reviewer is of the opinion that the X-ray measurements have been interpreted on too conventional lines and that an alternative explanation entirely consistent with the other substantial body of evidence is probable. The question will be discussed fully in a forthcoming paper.

The paraffin-chain salt aggregated in solution (we will take cetylpyridinium chloride as typical) is equivalent to a macromolecular colloidal electrolyte of molecular volume about 30,000. This is comparable to many of the simpler proteins, having molecular weight of 34,000 or near. It is rather less than one-fifth of the molecular volume of gum arabic. The

⁷¹ See references in review by J. W. McBain in "Advances in Colloid Science" (Interscience, New York, 1942), Vol. I, p. 99.

⁷² G. S. Hartley and D. F. Runnicles, *Proc. Roy. Soc.*, 1938, *A*, **168**, 420.

maximum charge is relatively greater, being about one-third of that of gum arabic. Unquestionably, a large fraction of this charge is not effective at accessible concentrations, and it must be remembered that for the pure salt, the critical concentration for breakdown of the aggregate into simple ions sets a lower limit to concentration much above that set by purely practical considerations. Comparison of the electrochemical behaviour of a paraffin-chain salt solution with that of an approximately globular carbohydrate type of electrolyte of similar particle volume and charge would be of great interest in determining the effect on superficial ionic association of the very low dielectric constant of the interior of the aggregate.

The most important applications of paraffin-chain salts are as emulsifying, wetting, washing, and foaming agents. These properties, in aqueous solution, are essentially those of the solution of ultimately dispersed ions. The aggregate is not *per se* surface active. The amphipathic ion is. If, in contact with the water, another phase exists which is less reluctant to accommodate the paraffin chain, the latter will seek this other phase but leave the ionic head in the water. The amphipathic ion acts as a bridge between the phases and tends to stabilise their surface of contact. The paraffin-chain salts are in fact surface active and have, as such, important auxiliary properties in colloid systems, not because they are colloidal electrolytes, but rather owing to the same primary cause which, in sufficiently concentrated solutions, makes them colloidal electrolytes.

From the enormous number of amphipathic compounds made available in recent years by industrial development in this field, there stand out as of particular importance for the study of the constitution of solutions the compounds where the ionic group has been displaced from the end towards the centre of the chain or where two short paraffin chains have replaced one larger one in attachment to a more or less complex ionic group. Compounds of this type may have, as the Reviewer has shown elsewhere,⁷³ greatly enhanced activity at oil-water interfaces. This is not due to a greater activity of the single ions at a given concentration, but to a higher concentration of single ions being attainable. Two short chains have almost as much tendency to seek a separate oil phase in contact with the solution as has a single one of the same total length. The tendency to form an aggregate is, however, much reduced because, as we have seen, the size of the aggregate is determined primarily by the length of the longest chain. Thus the compounds are more active because less colloidal. The advantage gained in this respect may, however, be largely lost through a much increased solubility in organic liquids. The very low interfacial tensions between *cyclohexane* and water recorded for the potassium mono-sulphonate of 1 : 3-diethyloxybenzene were not realised with benzene-water because the salt dissolves much more freely in benzene than in water. Similarly, numerous salts of Aerosol OT (sulphonate of dioctyl succinate) form only hard glasses in the solid state which are freely soluble in a wide variety of organic solvents.

Many of the commercially important direct dyes are aggregated to some

⁷³ *Trans. Faraday Soc.*, 1941, **37**, 130.

extent in aqueous solution. The amphipathic principle is probably of importance in these cases also. One or two benzene rings and perhaps a naphthalene ring constitute a fairly bulky load of hydrophobic material to be carried into solution by one ion. The water solubility of the organic part of the dye ion is not, however, so extremely low as that of an equivalent amount of paraffin, and the disposition of ionic groups is not usually favourable to the very simple type of aggregation occurring in the straight single paraffin-chain salts. Aggregation in dyes would not therefore be expected to be so great, so universal, or so simple. This is indeed found to be the case. Several investigations of dye solutions have been made by diffusion,⁷⁴ mobility,⁵⁰ and osmotic⁷⁵ methods. More recently, optical methods also⁷⁶ have been employed, depending on a change of absorption spectrum on aggregation, but the method does not appear to be of much value except with mixtures of dyes.

The aggregation of dyes in solution is of importance in dyeing practice as it is considered that only the single ions are in general capable of penetrating swollen fibres.⁷⁷ Since dyeing is normally carried out in concentrated salt solutions, it is the degree of aggregation in such solutions which is significant, and this, from diffusion measurements, is more accurately estimable than that in pure solution.

Estimated aggregation numbers are usually less than 10, are considerably dependent on salt concentration, and may vary widely between similar compounds differing only stereochemically.

Aggregation might be expected to be more pronounced in dyes where the ionic group or groups are less symmetrically disposed and where the equivalent weight is greater. A series of azo-dyes with a single ionic group and varying numbers of coupled constituents has been investigated by C. Robinson and H. E. Garrett,⁴⁴ and it was indeed found that aggregation number increased up this series. Aggregation, though slight, is frequently sufficiently persistent in dyes to great dilution to cause the equivalent conductivity to exceed the infinite dilution value, going through a maximum about 10^{-3} to $10^{-4}N$. In the case of the highest compound examined in the above series, having four benzene rings, the maximum was still not reached at less than $10^{-5}N$ dye concentration.

As with the paraffin-chain salts, an indefinite and much coarser dispersion can exist under suitable conditions. The clear solution of sodium palmitate at 60° becomes a silky suspension at room temperature. Benzopurpurin-4B to which an excessive amount of salt has been added becomes strongly birefringent on streaming. In both cases there is no reason to believe that the suspension of fine micro-crystals is in a true equilibrium state. In the

⁷⁴ Ref. (11); C. Robinson, *Proc. Roy. Soc.*, 1935, **A**, **148**, 680; E. Valko, *J. Soc. Dyers Col.*, 1939, **55**, 173; S. Lenher and J. E. Smith, *J. Physical Chem.*, 1936, **40**, 1005; F. H. Holmes and A. H. Standing, *Trans. Faraday Soc.*, 1941, **37**, 542.

⁷⁵ C. Robinson and H. A. T. Mills, *Proc. Roy. Soc.*, 1931, **A**, **131**, 596; C. Robinson and F. H. Selby, *Trans. Faraday Soc.*, 1939, **35**, 780; H. O. Dickinson, *ibid.*, 1947, **43**, 486.

⁷⁶ D. R. Lemin and T. Vickerstaff, *ibid.*, p. 491.

⁷⁷ J. Boulton and T. H. Morton, *J. Soc. Dyers Col.*, 1940, **56**, 145.

aggregation referred to in the above brief discussion, we seem to be concerned with an indefinitely stable arrangement, although there is evidence of some hysteresis. The mechanism of limitation of aggregation in dyes is not clear. It is possible that some amicrocrystallites are involved, built on a pattern which cannot be indefinitely extended.