A review of the early development of the thermodynamics of the complex coacervation phase separation

Arthur Veis *

Department of Cell and Molecular Biology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

Abstract

Coacervation was defined as the phenomenon in which a colloidal dispersion separated into colloid-rich (the coacervate), and colloid-poor phases, both with the same solvent. Complex coacervation covered the situation in which a mixture of two polymeric polyelectrolytes with opposite charge separated into liquid dilute and concentrated phases, in the same solvent, with both phases, at equilibrium, containing both polyelectrolytes. Voorn and Overbeek provided the first theoretical analysis of complex coacervation by applying Flory–Huggins polymer statistics to model the random mixing of the polyelectrolytes and their counter ions in solution, assuming completely random mixing of the polyelectrolytes in each phase, with the electrostatic free energy, \( \Delta G_{\text{elect}} \), providing the driving force. However, experimentally complete randomness does not apply: polyelectrolyte size, heterogeneity, chain stiffness and charge density (\( \sigma \)) all affect the equilibrium phase separation and phase concentrations. Moreover, in pauci-disperse systems multiple phases are often observed. As an alternative, Veis and Aranyi proposed the formation of charge paired Symmetrical Aggregates (SA) as an initial step, followed by phase separation driven by the interaction parameter, \( \chi_{23} \), combining both entropy and enthalpy factors other than the \( \Delta G_{\text{elect}} \) electrostatic term. This two stage path to equilibrium phase separation allows for understanding and quantifying and modeling the diverse aggregates produced by interactions between polyampholyte molecules of different charge density, \( \sigma \), and intrinsic polyelectrolyte structure.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

H.G. Bungenberg de Jong, one of the early leaders in colloid chemistry at Utrecht University in The Netherlands, defined and described the phenomenology of complex coacervation in several superb chapters in a book, Colloid Science II, edited by H.R. Kruyt [1], written over several years and finally published in 1949. The

* Biochemistry and Molecular Biology, Department of Cell and Molecular Biology, Feinberg School of Medicine, Northwestern University, 303 E. Chicago Avenue, Chicago, IL 60611, USA. Tel.: +1 312 502 1355; fax: +1 312 503 2544.
E-mail address: aveis@northwestern.edu.

0001-8686/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cis.2011.01.007
discussion in these chapters was rooted in the prevailing theories and language of colloid chemistry, and concerned the phenomenon of “coacervation”, in which the original dispersed colloid “state” or sol separated into a colloid-poor and a colloid-rich phase. The case of particular interest was that in which both phases retain liquid solution character, differing in “colloid component” concentration. The separated colloid-rich phase could remain as a turbid suspension of amorphous liquid drops or coalesce into a clear liquid phase. Most importantly, Bungenberg de Jong pointed out the key element in these colloidal systems. The phases were thermodynamically in equilibrium, and not governed entirely by particle surface charge considerations as in classical colloidal particle instability.

The coacervation equilibrium could be established in two ways. A macromolecular solute could be brought to coacervation phase separation by adjusting the solvent conditions, such as bringing mixtures of water and various alcohols to different mixing ratios. In this ternary solution of a water soluble macromolecule and two solvent components, demixing at a particular range of alcohol/water ratios would lead to two phases, with a high concentration of macromolecule in one phase in equilibrium with a more dilute phase of the macromolecule having a higher water concentration. This type of coacervation was called “simple” coacervation. The subject of this brief review, “complex coacervation”, involves the system in which two macropolymers in a single solvent demix to form two or more phases, each phase containing both polymers. In complex coacervation demixing the macropolymers are of opposite net charge and at electrostatic equivalence in the concentrated phase. This is a true equilibrium in which the polymer components are in solution. By themselves each of the macropolymers is completely miscible in the solvent, but the mixture is driven to coacervate demixing by the solute electrostatic interactions.

In that same era, along with studies of polymer synthesis, the theoretical understanding of polymer physical chemistry was being developed. However, experimental studies on polymeric polyelectrolytes were basically restricted to the use of polymers of biological origin: proteins, plant gums and ionic polysaccharides. These were difficult to isolate and characterize as to purity, homogeneity and molecular weight, and particularly, for charge density and backbone charge distributions. Nevertheless, Overbeek and Voorn [2] and Voorn [3,4] in a series of related papers brought the application of a liquid lattice model of polymer solutions via the Flory–Huggins [4–6] theories of the thermodynamics, and explicitly, the statistical evaluation of the entropy and free energy of linear chain polymers in solution, to bear on the coacervation equilibrium. Although the Flory–Huggins model is well known, it is important here to explore its consequences with regard to the phase equilibria in polymer systems in general where a mixture of polymers is considered before going directly to the Overbeek–Voorn models.

### 2. Thermodynamics of polymer mixing

For a binary solution of solvent (1) and solute (2) the free energy of mixing is given by Eq. (1) in which \( \varphi_1 \) and \( \varphi_2 \) are the volume fractions of the components in the mixture and \( r_1 \) and \( r_2 \) are the lattice elements occupied by each component molecule. By convention we can set \( r_1 = 1 \) so that \( r_2 \) represents the number of solvent volume elements filled by a polymer molecule.

\[
\Delta G_m = RT[(\varphi_1 / r_1) \ln \varphi_1 + (\varphi_2 / r_2) \ln \varphi_2] + \chi_{12} \rho_1 \rho_2
\]

(1)

Several points become clear immediately. The first two terms represent the entropy of mixing of solvent and solute, and since the volume fractions are \( < 1 \), these provide a negative contribution to \( \Delta G_m \), favorable to mixing. In the third term, \( \chi_{12} \rho_1 \rho_2 \) counts the number of solvent contacts per polymer molecule and the \( \chi_{12} \) coefficient in the term is a measure of the difference between 1–1 and 1–2 segment interaction energy. In an ideal solution the 1–1 and 1–2 interactions are identical so \( \chi_{12} = 0 \). Thus, \( \Delta G_m \) will be negative and the solution will remain as a single, homogeneous phase. In real solutions \( \chi_{12} \) may be either endothermic (+) and unfavorable, or exothermic (−) and favorable. If the endothermic interaction is sufficiently large there will be a flat value at which \( \Delta G_m \) will be positive, and the polymer will no longer be soluble at all concentrations, a two phase demixing will occur at a critical concentration, \( \varphi_{2,crit} \). Flory [7] determined that when \( r_1 << r_2 \), \( \chi_{12} \) can be estimated by

\[
\chi_{12,crit} = 1/2 + 1/\sqrt{(r_2)} - 1/2
\]

(2)

so that at large \( r_2 \), \( \chi_{12} = \chi_{12,crit} \geq 0 \), and an endothermic \( \chi_{12} \) value only trivially larger than 0.5 will drive polymer phase separation, with \( \varphi_{2,crit} = 1/r_2^{1/2} \). The situation is shown graphically in Fig. 1, a plot of \( \Delta G_m / RT \) vs \( \varphi_2 \). The plot labeled ideal refers to the ideal mixing law, in which \( r_1 = r_2 \) and \( \chi_{12} = 0 \). The plot labeled “regular–athermal” is for the case of a binary mixture with \( \chi_{12} = 0 \) but \( r_1 << r_2 \). In this situation the plot still has a single minimum, but the entropy of mixing of polymer segments with solvent is less. The single minimum determines that the solutions will still be a homogeneous single phase at all mixing concentrations. The upper plot of \( \Delta G \), in this case with a \( \chi_{12} = 1.5 \), shows that after \( \Delta G \) decreases according to Raults’ law \( \Delta G \) increases to a positive value putting an inflection point into the plot which ultimately shows a second minimum. The tie line between the two minima, one at very low \( \varphi_2 \) the other at a very high value of \( \varphi_2 \), denotes that the solution mixture will demix to an equilibrium solution with a solution of very low concentration in equilibrium with one at a much higher concentration and the concentrated solution is thus a “simple coacervate”. The conditions for equilibrium are that the chemical potentials \( \mu_i \) of each component will be the same in each phase:

\[
\langle \mu_1 \rangle = \langle \mu_2 \rangle = \langle \mu \rangle = \langle \mu \rangle
\]

(3)

This is an important set of relationships for determining the equilibrium concentrations after phase separation since the chemical potentials of each component can, from Eq. (1), be written as:

\[
\mu_1 - \mu_0 = RT\left[ \ln(1-\varphi_2) + (1 - 1/r_2) \varphi_2 + \chi_{12} \varphi_2^2 \right]
\]

(4)

\[
\mu_2 - \mu_0 = RT\left[ \ln(\varphi_2) -(r_2 -1) + r_2 \varphi_2 (1 - 1/r_2) + \chi_{12} r_2 (1-\varphi_2)^2 \right]
\]

(5)

![Fig. 1. The basic effects of molecular size differences in simple phase separation of random chain mixtures based on the Flory–Huggins equations with the inclusion of Van Laar interaction term.](image-url)
and at the critical concentration for phase separation, $\phi_{2,\text{crit}}$, the first and second derivatives of the chemical potential, at constant $T$ and $P$ with respect to $\phi_2$ are,

$$\frac{\partial \mu_1}{\partial \phi_2} = \frac{\partial^2 \mu_1}{\partial^2 \phi_2} = \frac{\partial \mu_2}{\partial \phi_2} = \frac{\partial^2 \mu_2}{\partial^2 \phi_2} = 0$$ \hspace{1cm} (6)

so that from Eqs. (4) and (5)

$$\phi_{2,\text{crit}} = 1 / \left(1 + \frac{1}{2} \phi_{1,\text{crit}}^1\right)$$ \hspace{1cm} (7)

It was exactly this approach that Overbeek used to determine the theoretical equilibrium phase concentrations in complex coacervation. However, before adapting Eq. (7) to the complex coacervation systems, we need to consider the thermodynamic approach to ternary systems of two polymer components in a single solvent in which each separately is soluble. First consider the case where the two polymers components 2 and 3 are liquids in the pure state, and where $T_2 = T_3$ (equal molecular weights) and $\chi_{23} = \chi_{32}$. A mixture of 2 and 3, with zero solvent and with the understanding that each polymer molecule would occupy a single lattice site ($r_2 = r_3 = 1$), then the solution is identical to a a regular binary mixture of liquids, or

$$\eta_2 - \eta_0 = RT \left[ \ln \phi_2 + \chi_{23} \phi_2^2 \right]$$ \hspace{1cm} (8)

$$\eta_3 - \eta_0 = RT \left[ \ln \phi_3 + \chi_{23} \phi_3^2 \right]$$ \hspace{1cm} (9)

in which the $\chi_0$ terms are $r$ times larger than the polymer segment interaction energy $\chi_{1r}$ terms in Eqs. (4) and (5). Thus, to quote directly from Flory [7] "only a minute, positive first neighbor interaction free energy is required to produce limited miscibility. The critical value for interaction free energy is so small for any pair of polymers of high molecular weight that it is permissible to state as a principal of broad generality that two high polymers are mutually compatible with one another only if their free energy of interaction is favorable, i.e., negative." Including the small molecule solvent 1 in the ternary mixture leads to Eqs. (10), (11), and (12) for the component chemical potentials [8], comparable to Eqs. (4) and (5).

$$\mu_1 - \mu_0 = RT \left[ \ln \phi_1 + (1 - \phi_1) \chi_{12} \phi_2 \phi_3 / r_2 \right]$$ \hspace{1cm} (10)

$$\mu_2 - \mu_0 = RT \left[ \ln \phi_2 + (1 - \phi_2) \chi_{12} \phi_1 \phi_3 / r_2 \right]$$ \hspace{1cm} (11)

$$\mu_3 - \mu_0 = RT \left[ \ln \phi_3 + (1 - \phi_3) \chi_{12} \phi_1 \phi_2 / r_2 \right]$$ \hspace{1cm} (12)

3. Complex coacervation

The ternary system mixture of a single solvent and two macro-polymers of opposite charge introduces further complexities and differs markedly from the analysis provided by Eqs. (10), (11), and (12). The major difference is that both oppositely charged macropolymers remain in both phases after coacervation, strikingly different from the predicted separation of the polymers. The most general way to consider the phase separation which takes place when oppositely charged random polymers are mixed is to consider the free energy of mixing in its broadest possible formalism. That is:

$$\Delta G_m = \Delta G_{\text{Flory-Huggins}}(\text{Entropy}) + \Delta G_{\text{VanLaar Chain Segment Interactions}} + \Delta G_{\text{Electrostatic Interactions}} + \Delta G_{\text{Other factors}}.$$ \hspace{1cm} (13)

3.1. The Voorn–Overbeek model

Overbeek and Voorn took the approach that only the Flory–Huggins entropy of mixing and the electrostatic interactions were important. They estimated that Van Laar solvent–solvent interactions (1–2, 1–3) and solute–solvent non-electrostatic (2–2, 3–3, 2–3) interactions were all negligible and could be disregarded. They modeled the system by considering that the polymers were essentially random chains in both dilute and concentrated phases with the chain elements distributed randomly in the model lattice in both phases. Moreover, they treated the charges along the chain backbones as being distributed in the solution in both phases as if they were independent of their location on the polyp chain backbones, that is, the charges entirely distributed in the solution as predicted by the Debye–Hückel theory for monovalent backbone charges and monovalent counter ions and salts:

$$G_{\text{elect}} = e^2 / 3 \varepsilon \left(4 \pi e^2 N_e / e kT \right)^{1/2} N_e.$$ \hspace{1cm} (14)

Obviously $G_{\text{elect}}$ would need to be modified to accommodate counter ions of higher charge, but this needlessly makes Eq. (14) more complex since it is rare in the polypeptide solutions to have multiply charged side chains. In Eq. (14) $\varepsilon$ is the solvent dielectric constant, $\alpha$ the elementary charge, $k$ the Boltzmann constant, $T$ the absolute temperature, $V$ the volume of solution and $N_e = \sum_i n_i z_i$ the total number of + and − charges. The assumption is made that in volume $V$ containing $n_i$ molecules or ions, by setting the solvent molecular partial volume as $v$, each component $i$ has a molecular partial volume $r_i v$, and letting each microion also have a volume $v$, then each macronion will have a molecular volume of $r_i v$ and each volume element will then have a charge of $+, 0$ or $-$. Thus each chain segment for component $i$ will have an average charge density of $z_i / r_i = \sigma_i$. The volume fractions of each component will be $\phi_i = n_i r_i v / V$ with this notation, and with $N = V / v$, then Eq. (14) can be rewritten as:

$$G_{\text{elect}} / N kT = -a \left[ \sum_i \phi_i z_i \right]^{1/2}, \text{with } a = \left( e^2 / k e kT \right)^{1/2} \left(2 / 3 \sqrt{N / v} \right).$$ \hspace{1cm} (15)

With all of the above assumptions and limitations then Eq. (13) can be written as

$$\Delta G_m = k T \left[ \sum (\phi_i / r_i) \ln \phi_i - a \left[ \sum_i \phi_i z_i \right]^{3/2} \right].$$ \hspace{1cm} (16)

Obviously, the first term in Eq. (16) is the athermal Flory–Huggins entropy of mixing of the components. The second term provides the electrostatic interaction contribution to the free energy of mixing and, as stated earlier, is always favorable with increasing concentration of the polypeptide. The direct application of the electrostatic interaction...
term in Eq. (16) to assess a particular mixture of poliyons is quite difficult to generalize, so Voorn [3,4] devised an analysis of the “symmetrical” mixing situation in which poliyons $P^+, Q^-$, microcounter ions $A^-$ and $C^+$ in a common solvent, at initial concentrations $c_i$ and represented by: $[P^+ A^-]_i$ and $[Q^- C^+]_i$, establish the phase equilibrium:

$$[P^+ A^-]_i + [Q^- C^+]_i \rightarrow [P^+ A^- + Q^- C^+]_c^{1/2}$$

Further, if it is assumed that the chain lengths and sizes are the same $r_P = r_Q$ and $\phi_P = \phi_Q$ and have the identical linear charge densities, and that the microcounter ions have the same lattice size as the solvent and the individual chain segments, then Eq. (15) reduces to the simplified form

$$\Delta G_{\text{select}} / RT = (4\pi / 9\nu)^{1/2} \left( e^2 / \epsilon kT \right)^{3/2} \left( \phi_P + Q \right)^{3/2} \sigma^{3/2}$$

$$= y(\phi_P + Q)^{3/2} \sigma^{3/2}$$

With all of these simplifications, and taking the same approach as in Eq. (2) the critical values

$$2\chi_c = \left( 1 / (1-\phi_c)^2 \right) - (3 / 8) \left( \nu \phi_c^{1/2} \phi_c^{-1/2} \right)$$

$$1 / r_c = \phi_c^{2} / (1-\phi_c)^2 + (3 / 8) \nu \phi_c^{3/2} \phi_c^{1/2}$$

for $\chi_c$ and $\phi_c$ can be expressed as in Eq. (19), where the subscript $c$ is the critical value, and $\phi_c = (\phi_P + \phi_Q) = 2\phi_Q$ for this symmetrical system. Further, since all of the factors included in $\nu$ are fundamental constants, then at $T = 300^\circK$ and a dielectric constant at 80, and a reasonable value of $\nu = 3 \times 10^{-22}$ cm$^3$ molecule$^{-1}$ then

$$r_c \alpha_c^3 \approx 0.45$$

Assuming both $P$ and $Q$ are each polypeptides with $r = 1000$ residue chain length, each residue occupying 5.5 lattice sites, $r = 5.5 \times 10^3$ lattice sites then $\alpha_r$, the critical charge density for coacervation phase separation, would be $\alpha_r = 0.043$ or 240 charged residues per chain. Using the requirement that at equilibrium the chemical potential of each component must be the same in both phases, Voorn [4] found the equations too difficult to make any explicit relationships between $r_c$, $\phi_c^P$, and $\phi_c^Q$, so he resorted to numerical approximations to make predictions to compare with the extensive data in the literature. The result was that there was not reasonable agreement with experiment: the concentrations of the dilute phases were too low, and the charge densities too high, Fig. 2.

3.2. Experimental approaches to the symmetrical aggregate model

My own thinking about coacervation began immediately after I read the Bungenberg de Jong chapters mentioned above [1]. My dissertation research had concerned the evaluation of the electrostatic interactions along the backbone of linear chain polyelectrolytes and the differential interactions of microcounter ions ( gegen ions) of the opposite charge, namely Na$^+$ and K$^+$ ions [10] and their effects on the chain hydrodynamic properties [11]. These studies showed that even monovalent cations as close as Na$^+$ and K$^+$ had different effects on their binding as gegen ions to an anionic poliyon, and that the biological polyelectrolytes available for study in the early 1950s formed solutions far from ideality; their interactions with solvent (water) were different and chain stiffness was usually different depending on the poliyon source. In particular, many of the early complex coacervation experiments involved plant gums, polysaccharides with attached carboxyl (Gum Acacia, Gum Arabic) or sulfate (Agar) groups providing the charge, but with different charge densities and complex chain backbone elements. Gelatin was the usual low charge density positively charged poliyon, but gelatins could be prepared with different isoelectric point using different pretreatments and extraction and degradation procedures from collagen, the parent protein. Qualitatively, the Voorn–Overbeek assumptions were applicable, but one should have anticipated the deviations from ideality. Thus, I embarked on a series of experiments designed to use gelatins of essentially identical backbone sequence, but prepared to have different isoelectric points. Further, the gelatins were to be fractionated so that oppositely charged gelatin of nearly identical molecular mass and chain length could be mixed so that the conditions $r_2 = r_3$, $\chi_{12} = \chi_{13}$, $\chi_{23} = 0$ could be achieved.

3.3. Gelatin heterogeneity

While this plan was easy to state, it was difficult to achieve. Commericially available gelatins were of varying purity, and usually blended to achieve some particular goal, such as viscosity, gel strength, or adhesiveness. The actual molecular weight distributions within a gelatin preparation were neither predictable nor reproducible. Preliminary experiments with commercially available gelatin did show that gelatin–gelatin coacervates would form in the pH range between the gelatin chain isoelectric points, but not outside that range, confirming that this was indeed an electrostatically driven process. However, the inconsistencies drove us back to a different career changing path: the study of the parent protein, collagen, and the collagen–gelatin transition [12–20] which culminated in publication of “The Macromolecular Chemistry of Gelatin” in 1964 [24]. It was from that base that most of the following developed.

It is now well known that there are more than 29 types of collagen molecules with at least 40 distinct types of constituent polypeptide chains. We confine ourselves here to type I collagen, the species present in largest amount in the vertebrate body, and most commonly used in gelatin production. The type I collagen molecule is comprised of three polypeptide chains wound together in a compound triple helical array. The polypeptide chains of mature type I collagens are of three polypeptide chains wound together in a compound triple helical array. The polypeptide chains of mature type I collagens are of th

Fig. 2. The effect of the electrostatic free energy according to the Voorn–Overbeek free energy of mixing relationships. The Hory–Huggins mixing entropy, $\Delta G_{\text{select}} / RT$, is the upper curve. The concave downward dashed line if the $\Delta G_{\text{select}} / RT$ contribution, with the solid line being the sum of the two terms. The dashed line is the binodal tangent joining the total $\Delta G_{\text{select}} / RT$ minima at very low and very high $\phi_2$, not consistent with the data shown in Fig. 4B.
triple helix of the collagen molecule, with formula \([\alpha1\beta\gamma]\). The molecules self aggregate within the body into fibrils, which are then stabilized by an array of intermolecular (and some intramolecular) and interfibrillar covalent cross-linkages. Collagen fibrils are converted into gelatin by a combination of denaturation or melting of the triple helical portions and the cutting of the intermolecular cross-linkages. The conversion can lead to complete separation of the individual chains or, where some cross-linkages remain, networks of chains. In commercial gelatins the chains may also be degraded by hydrolysis or enzymatic cleavage of the peptide backbones. Obviously, this leads to the potential for tremendous heterogeneity in any gelatin preparation. The individual α chains of mature secreted type I collagen all have a central region of 1014 amino acids with sequence \((\text{Gly-X-Y})338\), of which ~120 residues are Proline (Pro) and ~100 are Hydroxyproline (Hyp). The X and Y in each triplet can be any amino acid, including Pro, but Hyp can only occupy position Y. Triplets Gly-Pro-Pro and Gly-Pro-Hyp are frequent. Thus, under the best of circumstances one cannot eliminate some heterogeneity from denatured and separated type I collagen α chains. Although the α chains are about \(10^5\) residues in length with molecular masses ~\(1 \times 10^5\) kDa, the intramolecular cross-linkages yield molecular sizes equivalent to \(3 \times 10^5\) kDa. Preparations from some tissues yield mixtures of intramolecular cross-linkages yielding molecular sizes equivalent to linked collagens can be even larger polymers. Commercial gelatins are mixtures of these three components. Gelatins from mature, cross-linked collagens can be even larger polymers. Commercial gelatins usually contain much lower Mr gelatins.

### 3.4. Gelatin charge density

Pro and Hyp constitute 200–250 of the 1000 residues of the α chain and dominate the temperature dependence of chain folding and gelation, but by holding the temperature to \(\geq 40^\circ\text{C}\) in aqueous solution the chain conformations are random except in the positions of possible interchain cross-linkages. Each α chain also carries about 25 Lys, 7 OH Lys, 48 Arg and 7 His residues for a potential maximum 87 positive side chain charges at \(pK_a\). The addition of a simple \(1 \times 10^{-3}\) salt at \(1.5 \times 10^{-3}\) M. Since each gelatin chain alone is soluble at its IEP and does not form either a simple coacervate or precipitate at any concentration, whereas a coacervate does form when gelatins of IEP9 and 5 are mixed under isoionic conditions, the process of coacervation is clearly driven by intermolecular electrostatic interactions and not by intrachain charge pairing. This behavior is very gegen ion concentration dependent.

#### 3.5. Gelatin concentration and mixing ratios

According to the Voorn–Overbeek analysis, and Eq. (16) in particular, at \(r\) above the critical size the Flory–Huggins entropy term indicates that the actual size of \(r\) is much less important than charge density and that molecular weight heterogeneity is also unimportant. However, the effects of polyion concentration, and molecular weight distributions of the gelatins on the coacervation were found to be profound [21–23,26]. When mixed in the isoionic, symmetrical fashion described, the amount of coacervate phase can be described by the volume \(V_c\) of coacervate phase, the concentration \(C_c\) of coacervate phase, the concentration \(C_e\) of the dilute equilibrium liquid phase, as functions of the initial mixing concentration \(C_m\). In a typical phase equilibrium in which the components did not change nature, then at each value of \(C_m\) in the range where \(C_e > C_{m}< C_c\), the concentrations \(C_e\) and \(C_c\) should remain constant while the volumes of each phase change: \(V_c\) increasing, \(V_e\) decreasing. In the gelatin–gelatin system studied, shown in Fig. 4A the volume of coacervate phase increased to a maximum at \(C_m = 6 \times 10^{-3}\) g/ml then returned to zero at \(C_m = 12 \times 10^{-3}\). Fig. 4B is a plot of the phase concentrations in the same experiment and shows that \(C_c > C_e\) while \(C_e < C_m\) but that neither \(C_e\) nor \(C_c\) was constant. The gelatins used were crudely fractionated by alcohol precipitation, and both IEP5 and IEP9 preparations selected for study had weight average molecular weights of \(3.3 \times 10^5\). The phenomenon seen here is what Bungenberg de Jong had called “self suppression” of coacervation and he hypothesized that this was due to the random mixing of the polyions in the

---

**Fig. 3.** The pH and salt concentration effects on coacervation of a mixture of acid and alkali processed gelatins with isoelectric points pH 5 and 9. [28], fractionated to obtain preparations of equal apparent molecular weight [29] and deionized to isionic form [30] Titration data are correlated with coacervation data represented as volume of coacervate phase, showing that coacervation takes place strictly between the two IEPs [21]. The addition of \(1.5 \times 10^{-3}\) M KCl sharpens the pH range at which phase separation is observed.
coacervate phase and the greater contribution of the electrostatic interaction as the concentrations increased. We have interpreted this as related to the excluded volumes of the polyions. In this case both gelatins have equivalent root-mean-square end-to-end chain dimensions and molecular domains that, at $C_m = 7 \times 10^{-3}$, would overlap — the excluded volume point at which the independent domains would either begin to be constricted as the concentration increased, or begin to intermix or overlap within the same space. In the Voorn–Overbeek model for random chains, this later effect is exactly what was proposed for the coacervate phase, leading to a charge distribution effectively independent of the polyion backbones. That would increase $\Delta G_{\text{elect}}$ and as in Fig. 2, essentially lead to precipitation of the solvated polyion mixture. This is obviously not what occurs.

A series of investigations on the solution conformations of the two gelatins, both with comparable weight average molecular weights, by light scattering, viscosity, and sedimentation coefficients showed that the chain conformations of the IEP5 gelatin (B) conformed to the expected linear random chain model, whereas the acid processed IEP9 gelatins (A) were random in chain conformation, but better depicted as networked aggregates of chains, retaining interchain cross-linkages. Circular dichroism confirmed that both gelatins were completely in the random chain conformations at $T \geq 40^\circ C$. The intrinsic viscosity, $[\eta]_{A,B}$, of acid processed IEP9 gelatin was 0.47 dl/g, whereas the original mixture of A and B in NaCl and at $C_m = 2 \times 10^{-3}$ had $[\eta] = 0.58$, indicating that a molecular weight fractionation had taken place. The multiple plateaus further suggested that the gelatins might contain rather discrete sets of components, and could be considered as pauci-disperse, with a stepwise molecular weight distribution. Exploring that, the gelatins were fractionated by alcohol precipitation and equivalent populations were selected for mixing. As evident in Fig. 5B the plots of $C_c$ and $C_e$ vs $C_m$ showed aggregates formed and the phase intrinsic viscosities determined. It was found that a strong and specific fractionation had taken place, Fig. 5A such that at $C_m = 2 \times 10^{-3}$ g/ml, $[\eta]_A = 0.47$, $[\eta]_B = 1.18$ whereas the original mixture of A and B in NaCl and at $C_m = 2 \times 10^{-3}$ had $[\eta] = 0.58$, indicating that a molecular weight fractionation had taken place. The multiple plateaus further suggested that the gelatins might contain rather discrete sets of components, and could be considered as pauci-disperse, with a stepwise molecular weight distribution. Exploring that, the gelatins were fractionated by alcohol precipitation and equivalent populations were selected for mixing. As evident in Fig. 5B the plots of $C_c$ and $C_e$ vs $C_m$ showed...
plateau values suggesting that specific Mw components were preferentially collected in the coacervate phases. As Cm increased determination of the concentrations of the equilibrium liquid confirmed both the specificity of the fractionation and the stepwise nature of the coacervate complexes. Large scale coacervation was carried out and the coacervate was examined and found to have two distinct liquid layers, Fig. 6.

When the equilibrium liquid was collected from a centrifuged coacervated mixture at Cm between 2 and 7×10^{-3} g/ml, and maintained at 40 °C in the isionic condition, light scattering showed that high Mw aggregates were present. These were stable to dilution with water but disassociated upon addition of NaCl. This observation argues against the random mixing of the gelatin chains in the dilute phase, and thus led to the following reformulation of the basic equilibria. Eq. (17), the mechanism of the Voorn–Overbeek theory cannot be correct. Instead we proposed that the initial mixture immediately formed electrostatic-driven neutral aggregates which then can take either of two paths. In path 1, the dilute phase can remain as electrically neutral aggregates exhibiting a X > X_{cm} driving the phase separation as a simple coacervate. Upon formation, the coacervate phase at a concentration Cc, several fold exceeding the critical excluded volume. The overlapped chains at Cc may gain entropy by becoming indiscriminately mixed. The relevant overall reaction becomes

\[
(PQ)_{agg} ; C_m \rightarrow (\text{PQ})_{AGG} ; C_e \text{ hi} \text{ Dilute} ; \text{aggregate} + \left( (P^+ \text{OH}^-)_{Cc} + (Q^- \text{H}^+)_{Cc} \right)_{\text{Conc. Random Chains}}
\]

\[
(PQ)_{agg} ; C_m \rightarrow \left[ (PQ)_{AGG, C} \right]^{\text{I}} + \left[ (PQ)_{AGG, C} \right]^{\text{II}} \text{ Conc. Random Chains}
\]

A second path is the reaction in which aggregates remain in both phases:

\[
(PQ)_{agg} ; C_m \rightarrow \left[ (PQ)_{AGG, C} \right]^{\text{I}} + \left[ (PQ)_{AGG, C} \right]^{\text{II}}
\]

The data collected, particularly that from pauci-disperse systems and from analyses of mixtures of the polyions at non-equivalent concentrations (an excess of P^+ or Q^-) led clearly to the choice of the reaction path of Eq. (22), in which aggregates were present in both equilibrium liquid and coacervate phases: (1) polyion electrostatic equivalence was required in the concentrated phase, excess polyion remained in the dilute equilibrium liquid [25]; (2) in pauci-disperse
mixtures the coacervate phase was partitioned into coexisting separate phases; (3) molecular weight determinations showed that the coacervate phases showed selection by molecular weight or charge density [26]; and (4) in heterogeneous chain length polyion mixtures, the requirement for electrostatic neutrality leads to larger aggregates and at higher mixing concentrations reduces the intensity of coacervation so that the entire mixture remains as a single phase (Fig. 4A,B). A schematic illustration of the various options is provided in Fig. 7.

3.6. The symmetrical aggregate and standard state change from random mixing

Early on, it was recognized [21–23] that if aggregates formed, the basic Voorn–Overbeek thermodynamic analysis could not be correct. That is, while the Flory–Huggins entropy might apply to the concentrated phase of randomly mixed chains, as in Eq. (21), that treatment did not describe the aggregates in the dilute phase. Once the first step in either Eq. (21) or (22) forms the aggregate the nature of the aggregate becomes important: conceptually this is the same as asking if the aggregate is a new component. If that is the case then one has to think of that formation reaction as representing a standard state change. The initial state of each molecule is a random chain as distributed in the molecular domain of the single molecule and these can be described by a radial distribution function with a particular average end-to-end distance $r$, or more properly $(\bar{r}^2)^{1/2}$. The basic question is how the polion segments are distributed within the domain of the individual unperturbed polion molecule as compared to within the aggregate of 2 oppositely charged polions within a common domain. Any distortion of chain element distributions affects the conformational entropy and the distribution of charged groups affects the electrostatic free energy. Wall [31] determined that the distortion factor, $\alpha = l_\text{dist}/l_0$, where $l_0$ is the freely jointed random chain length (or average root-mean square of the end-to-end distance)) and $l_\text{dist}$ is the equivalent length in the overlapped chain domains of an aggregate, is related to the entropy of distortion by

$$\Delta S_\text{dist} / k = N_2 \{ \ln \alpha^3 - 3/2 (\alpha^2 - 1) \}.$$  \hspace{1cm} (23)

At equilibrium, $N_2$ is the number of freely jointed chain elements in the aggregate. $\Delta S$ has a maximum at $\alpha = 1$ and is negative for all other values of $\alpha$: it makes no difference if the chains are crowded or stretched. Using this approach, the deviations from ideal behavior can be approximated by a virial expansion dominated by the second virial coefficient $B_2$. Flory [7] determined that the best approximation to $B_2$ is

$$B_2 = \left( 2^{5/2} \pi (\bar{r})^{3/2} / 27 \right) \ln \left[ 1 + \pi^{1/2} (\alpha^2 - 1) \right].$$  \hspace{1cm} (24)

This interaction term needs to be included in the $\Delta G_\text{m}$ equations, as suggested in Eq. (13). Thus the estimation of $\alpha$ and the distribution of chain elements within the aggregates is of major importance.

Gates [32] and Veis [27,33] attempted to determine an expression for $\Delta G_\text{m}$ for the case of aggregates in both phases by developing analytical expressions for the chemical potentials of the solvent in each phase based on an aggregate model for components 2 + 3. Then at equilibrium and with the use of the Gibbs–Duhem equation we equated the chemical potentials of the solvent in each phase. This expression contained terms for the phase compositions, the molecular weights of the solutes, the solute excess charge densities on each chain, and the constants such as temperature and dielectric constant, all of which could be measured and evaluated experimentally without any particular model for the aggregate structure. The Gibbs–Duhem equilibrium equation then yielded:

$$\ln \left( 1 - \phi_0^i \right) + \phi_2^i - 1 / M \left( \phi_0^i - \phi_2^i \right) + \chi \left( \phi_0^2 \right)^2 - \left( \phi_2 \right)^2 \right]$$  \hspace{1cm} (25)

$$+ \left[ 0.5 - 0.46Y \phi_0 \left( \phi_0^3 / \phi_2 \right)^{1/2} + 0.5 \Omega \right] \left( \phi_2 \right)^2$$  \hspace{1cm} (26)

Putting in the experimental values of $\phi^i$, $\phi_0^i$, $\phi_2^i$ and $M$ and $\alpha$ appropriate to the plateau regions in the experiment described in Fig. 5B, one can calculate that $\chi_{1-agg}$ the interaction coefficient between the solvent molecules and the new component, the electrostatically neutral 2 + 3 aggregate, had the values 0.63, 0.56 and 0.62 for $C_m = 0.3$, 0.9 and 1.2 x $10^{-2}$ g/ml, respectively, a remarkable near constant interaction parameter in both phases, and consistent with the observed phase separation. This also indicated that there was very little change in electrostatic free energy on transfer of a polion aggregate from the dilute to the concentrated phase, while the aggregates had a $\chi >$ that required for coacervation: in other words, clearly supporting the symmetrical aggregate model of Eq. (22).

3.7. Implications of the symmetrical aggregate model for polion interactions and complex formation

The work in my laboratory ended at this point as the NIH support for the study was not renewed. The main criticisms were that a): the theory was too difficult to develop further, and b) there was no clear biological significance. Fortunately others proceeded to study the problem from the theoretical perspective [34–36] as well as the many practical applications to polymer systems, with modeling of the structure of the complexes or aggregates. Tainaka [35,36] used the Veis–Gates second virial coefficient as applied to the dilute phase, but applied a new model for the distribution of charge within the coacervate phase, including the total charges within the aggregate rather than the net charge. He developed an interaction potential function $U(R)$ to describe the potential acting at a volume element $i$ from symmetrical aggregate elements at positions $l$ and $k$ separated by a distance $R$. He used this to describe the chemical potential in a fashion similar to the Veis–Gates model setting

$U(R) = \sum_{i} \left[ \Delta G(\phi_2 + \phi_0^i) - \Delta G(\phi_2^i) \right]$  \hspace{1cm} (26)

where $U(R)$ could be composed of 2 terms: $U_1(R) = X \exp(-0.75R^2/S^2)$ in which $S$ is the radius of gyration of the aggregate, and $U_2(R) = -X \exp(-0.5625R^2/S^2)$, $X_1$ includes the sizes of the polyion chains, and the $X_2$ interaction factor, and $X_2$ accounts for the electrostatic interactions within the aggregate, related to the charge density $\sigma$. Other factors such as ionic strength effects could be included simply by adding more terms: $U(R) = U_1(R) + U_2(R) + U_3(R) + \ldots$

Tainaka took the data from Nakajima and Sato [34], a system of polyvinyl alcohol fractionated to the same size and derivatized to the same extent to make a symmetrical aggregate of equal charge densities and near equal backbone properties, and from Veis [8] on the gelatin mixtures described above and compared the experimental data on phase compositions and the theoretical calculated phase compositions and found, in both sets of data, reasonable agreement between theory and experiment. However, still further refinement of the potentials was necessary.
3.8. Molecular shape and size in the symmetrical aggregate

The Flory–Huggins, Voorn–Overbeek and Veis–Gates models are all predicated on using the random chain model as the basis for estimating entropy of mixing. The symmetrical aggregate model, as depicted schematically in Fig. 7D, brings the shape and internal structure of the aggregate into more focus. One of the more recent projects in my laboratory [37] involved the interaction between rod-like soluble type I collagen monomers (C) and a phosphate-containing protein, phosphophoryn, which has many phosphate groups along a backbone also rich in aspartic acid. [38,39]. The phosphophoryn (PP) appears, in calcium ion-free dilute solution at neutral pH, to have a random chain structure. Fig. 8 presents some direct rotary shadowing transmission electron microscopy evidence of C–PP complexes at differing C/PP mixing ratios. The semi-rigid rod-like structure of the collagen monomers (M = 300 kDa) with high axial ratios and the small coil-like PP (M = 90 kDa) are clearly seen before mixing. At very high C/PP (Arrow 1) there is a single specific PP binding site or pair of such sites that use PP as an internal cross-linking site forming loop structures, and no free PP can be seen. As the C/PP molar ratio decreases, PP–C interactions draw many individual collagen chains into a multi-valent, multi-chain aggregate (Arrow 2), and no free PP can be seen. As the ratio of PP to C becomes greater (Arrow 3), then the multitude of collagen chains joining in the aggregates increases and forms huge looped complexes and some free PP is present. At the highest PP/C mixing ratios (Arrow 4) no free collagen chains can be seen but the multi-collagen-multi PP aggregates coalesce into asymmetric complexes, with the shape clearly dominated by the high axial ratio collagen molecules. This is highly ionic strength dependent, with the aggregates easily disassociated. The Manning theories [40] on territorial binding of ions to polymeric ion backbones suggest that micro-counterions behave as if they are condensed with the polion backbone, although free to move along the chains and not specifically site-bound. Hence the Debye–Hückel treatment is not applicable within the polion local territorial domain. The same apparently is the case for the random chain PP–collagen rod-like complexes. The studies have not been made but it is likely that both the collagen conformation and the PP random-coil conformations are affected. The approach of Zhang and Shklovskii [41] considers all of these local and conformational effects on the phase diagrams. It will be interesting to see further examination of the phase diagrams along the lines proposed, but as indicated in [41] there are many caveats to be considered.

The work described above has focused on the work of the past, mainly attending to the early theoretical developments. More recently, an important paper was presented by D.J. Burgess [42] who summarized the prior work on gelatin coacervates and compared the various theoretical models. He extended the field considerably by using these theoretical works as a basis for understanding a major practical use of gelatin–gelatin, gelatin–albumin, and gelatin–acacia coacervates in microencapsulation. He very clearly analyzed the differences between various models and pointed out new directions for study. Bohidar’s laboratory [43] took a very interesting approach in considering if a random chain polyampholyte such a single type of gelatin could self-interact to form a coacervate by, in essence, dehydrating the molecules by using ethanol to alter the dielectric constant. It was found that indeed coacervate-like droplets did form. One problem with that is the extreme molecular weight dependence of precipitation of gelatin with ethanol [44] is well known. The fractionation by molecular weight or gelatin type needs to be...
considered. Perhaps the ethanol induced “coacervation” is mainly of the simple coacervate type. Tiwari et al. [45] made an interesting study of the behavior of mixtures of pl 5 and pl 9 gelatins, but used them without fractionation so that the maximum complex formation was at a mixing ratio of 3.2 for pl 9/pl 5. This indicates that the coacervate complexes must have contained different numbers of gelatin molecules of the two types, far from the symmetrical aggregate condition which leads to multiple phases in pauci-disperse systems. Much work remains to be done to clarify the details of these systems.

The key to moving further experimentally is to use means formerly unattainable but now feasible to investigate the structures of the complexes. Foremost would be to use fluorescent probes that can be differentially bound covalently to the polyions and then directly determine the molecular distributions within the coacervate phase and dilute aggregate phases. Coupling that with dynamic light scattering and chromatographic examination of the molecular aggregates would go a long way to clearing up the nature of the complexes. However, it does appear that the random mixing of oppositely charged random chain polyions under conditions of phase separation leads to the formation of non-random complexes which then form “simple” coacervate phases which are anything but simple.

Acknowledgements

I would like to thank Drs. Dubin and Kayitmazer for inviting me to review my early work in this still challenging field. I can see so many ways to use modern technologies to look at the nature of the coacervate complexes that were not available 30 years ago that I would go back to this study if it were at all feasible. A highlight of my career was a John Simon Guggenheim Foundation Fellowship year with Paul Flory at Stanford University, who listened to my approach and supported my continuing studies. I also would like to acknowledge Dr. James Harwood, Director of Basic Research, at Armour and Company, Chicago, IL, who allowed me to pursue this work just because I found it interesting.

The work from the author’s laboratory discussed in this review has been supported at various times by: Armour and Company, Central Research Laboratory, Chicago, IL (1952–1960), NIH Division of General Medical Sciences Grant GM 08114 (1960–1970), and John Simon Guggenheim Fellowship (1967–1968).

References


