

Comments for "Phase Transitions of Associative Biomacromolecules"

Armaan Ahmed

March 2023

Note: Any outside sources are cited by "([Last Name], [Article Title], [Year])"

1 Introduction

Biomolecular condensates: Assemblies of usually hundreds of molecules concentrated into a small volume between the nano and micro scales (size of individual proteins to size of cells).

1. Not usually homogenous, and can have different stoichiometries of different macromolecules.
2. These condensates usually behave, in the long time scales, similar to viscous or elastic materials.

Phase separation and condensate formation occur hand-in-hand.

1.1

Physical gelation: A type of phase-separation instantiating the concept of *bond percolation*, in which a transition branching across a given network driven by inter-polymer cross-linking interactions, occurs. Physical vs chemical gelation is a matter of the nature of the cross-linking interactions. Physical gelation is reversible.

1. "Percolation theory deals with the numbers and properties of the clusters formed when sites are occupied with probability p " (Christensen, Percolation Theory, 2002)
2. For example, *in vitro* gelation can happen in a protein solution by cohesive sticker-sticker interactions.
3. Relevant amino-acid level stickers are Gly or Gly-Leu-Phe-Gly motifs.
4. Physical gelation occurs, in part, as a cause of exceeding a certain critical protein sticker concentration

5. Gels have properties of both liquids and solids, ie a liquid that has elastic-resistance.

In vitro reversible gelation can be formed based on:

1. Number of stickers
2. Temperature
3. Peptide structures (eg, alpha helices)
4. Intrinsically disordered regions of RNA binding proteins

Questions:

1. How does the interplay between the time scales of molecular motion and cross-linking interactions determine gel properties/whether something even is a gel?

1.2

Gel "strength" (ie, strong or weak) depends on the strength of the component cross-link interactions.

Gels are not necessarily equivalent to solids or glass, and analogously, gelation is not equivalent to solidification nor vitrification (solidification without forming structured crystal structures).

However, gelation may imply vitrification if the rate of molecular movement is faster than the rate of cross/uncross-linking. **Why is this so? What is the molecular difference between gelation and vitrification?** The article mentions that although liquid water is a network of hydrogen-bonded water molecules, it is not a gel, which is okay because the rate of bond making/breaking is faster than water-molecule movement (diffusion), **Why?**

The rheology (related to topological properties/deformation characteristics) of gels are determined by the rate of molecular transport versus that for making/breaking bonds.

Throughout this article, gels are defined only focusing on connectivity!

Macroscopic definitions of gels are very intuitive, liquid and solid parts.

1.3

Intrinsically disordered regions (IDRs; regions without any defined shape, usually allows for non-specific associations) are important regions for when discussing the polymer-polymer association in liquids and polymer-solvent associations.

1. "A major surprise was the discovery that water is a poor solvent for homopolypeptides such as polyglutamine and polyglycine" - Charged and/or

hydrophobic residues minimize interactions with polar water. So, polymers composed of these types of residues can be more easily phase-separated (ie molecules segregate from one another).

2. Importantly, specific amino acid composition of chains can tip the balance between chain-chain, solvent-chain, and solvent-solvent interactions. (ie, we can also think about having an insoluble salt in water, generally in this case the salt-water interaction is weaker than the water-water interaction, thus remains phase separated).

1.4

Phase separation in cells: First example is the formation P granules in *C. elegans*. Distinct granules formed or dissolved at specified protein and RNA concentrations. Interestingly, these granules obeyed laws of Newtonian fluids (droplets fused, flows in response to applied forces, and droplets have round shapes)

1.5

There is a deep connection between physical gelation and phase separation, at least for proteins with multiple binding sites (ie protein multivalency) having folded domains (stickers) connected by disordered linkers.

Linkers are important: If a linker can be well-solvated, the both percolation and phase-separation is weakened. If a linker is completely Gaussian (ie random walker, does not have solvent interactions, "purely entropic"), percolation is enhanced by being coupled with phase separation.

1.6

In a macromolecular solution, there are two terms whose sum correspond to its overall free energy: **the free energy of mixing, and the free energy of reversible associations among macromolecules**. Ie, sum of energies when macromolecules are being mixed and when re-separating.

Phase separation: "a segregative transition that gives rise to two or more compositionally distinct phases that coexist with one another."

Reversible associations among macromolecules are driven by physical interactions between stickers (cohesive motifs). Interactions include hydrogen bonding, Coulombic interactions, and aromatic/hydrophobic group interactions.

A system of hard-sphere fluids (Figure 1) can percolate and phase-separate.

- Percolation is an associative transition: macromolecules associate/cross-link with one another. Associative transitions are driven by geometric/topological considerations. Conformation and/or self-assembly are drivers for this type of transition.

- Phase separation is segregative, ie system is separated into ≥ 2 phases.
- Associative/segregative transitions can happen hand in hand. COupled Associated and Segregative phase Transitions (COAST).

*In the Mintonian view, phase separation is binarized as just being associative or segregative depending on just the driving interactions. However, especially for a system of n -macromolecules, this is over-simplifying: we have to consider all the pairwise segregative and associative terms between these macromolecules in addition to pairwise solvent interactions.

Multivalent associative macromolecules: Ie, macromolecule containing stickers that allow for site-/chemistry-specific interactions more favorable than solvent-mediated ones.

Patchy colloids: Stickers are of defined sizes and orientations enabling site-specific interactions between the colloids (like intrinsically folded domains of proteins encoding binding specificity).

Linear associative polymers: Linear polymers having sticker motifs allowing for specific intra-/inter-polymer cross-linking.

Strengths of these types of reversible, specific cross-linkings spans orders of magnitudes, even just depending on available thermal energy (temperature).

Regions in between stickers are called **spacers**. Spacers influence solvent-polymer interactions (solubility), thus influencing percolation thresholds and also inter-polymer sticker-sticker interactions.

Generally, it can be non-trivial to distinguish between stickers and spacers. Ie, to distinguish requires precise knowledge about the context the polymer-of-question is in. However, this binirization is a useful abstraction to organize the driving factors for transitioning.

Sticker interactions allow for the formation of reversible cross-links. Relevant sticker-sticker interactions: salt bridges, hydrogen bonds, ionic interactions, and varying π system interactions.

The sticker-spacer formalism can be applied to nucleic acids as well: stickers are nucleotides driving the specificity of base-pairing, secondary structure formation, and inter-polynucleotide interactions, while spacers are the rest.

Questions:

- Dr. Zhang, a lot of your recent works relates to studying 2-body systems. Are there plans to expand to n -body systems, and how much more difficult is their study?

1.7

"An order parameter is a measure of the degree of order across the boundaries in a phase transition system" (Wikipedia).

1. Segregative transitions: the order parameter is the density vector of a system.
2. Associative transitions: order parameter related to the number of cross-links, topology, cluster density, etc.
- 3.

Order parameters can change quickly during a phase transition.

COAST examples include Phase separation coupled to percolation (obviously, abbreviated as PSCP) and complex coacervation (which is the associative liquid-liquid phase separation of a mixture of oppositely charged polymers; Sing and Perry, Recent progress in the science of complex coacervation, 2020). PSCP and complex coacervation are basically the same thing, except coacervation implies electrostatic associative interactions (so, PSCP is used for non-electrostatic associative interactions).

There are other COAST processes.

Is order parameter related to the entropy across the boundary?

2

Purely segregative phase separation gives rise to multiple coexisting phases, and there can be liquid-liquid, solid-liquid, liquid-liquid-crystalline, etc, phase separation.

For multipolymer systems, segregative phase separation occurs because of a mixture of pairwise favorable and unfavorable interactions between different types of polymers.

Density transitions: segregative phase transitions of a system with only one type of polymer.

2.1

Simple example: when there is a system of elastically colliding (but impenetrable) solid balls, all with radius σ Potential energy is:

$$U(r_i, r_j) = \begin{cases} 0 & |r_i - r_j| \geq \sigma \\ \infty & |r_i - r_j| < \sigma \end{cases}$$

Which is quite obvious... This is essentially saying that there is infinite repulsion between intersecting balls and otherwise no interaction.

This kind of system undergoes "freezing" transitions, where there is a solid phase and a dilute liquid phase after a saturation concentration.

Phase separation occurs because after a certain concentration, the system will obtain its lowest free energy (which here is equivalent to its higher entropy) state by forming a solid phase.

Radial distribution function: This function $g(r)$ describes for radially symmetric systems the relative probability of finding a particle at a distance r from a fixed particle. There is usually a short-range peak (for example, liquid argon has one in the single digit Angstroms range) and a long-range tail (ie, short range order and long range disorder for liquids).

As the density increases, the probability distribution begins "evening out", so an equilibrium freezing transition occurs to generate a more disordered dilute phase. (like if you increase the pressure of a gas by pumping in more, at some point it some will condensate out.)

Now we take the configuration integral:

$$Z_N = \int e^{\frac{\sum_{pairwise} U(r_i, r_j)}{k_b T}} d^{3N} r$$

I understand the concept of partition function as the discrete sum of all possible configurations of a system, but I am a little uncertain as to how this is related to the above integral, which is also known as the configurational partition function.

N is the number of hard spheres. When using the simple elastic potential, The above configuration integral becomes independent of temperature, and so the thermodynamics depends only on the density. Higher density means lower accessible volume (less number of possible states), meaning lower entropy, thus the system seeks to minimize its free energy by reducing its density (ie creating a dilute phase).

For a one-component system, this entropy driven liquid-to-solid transition is called a **density transition**.

The shape of the crystals is determined by the shape of each particle (ie rods, disks, etc).

2.2

Now, instead of completely elastic fluids, we can talk about more realistic **van der Waals fluids**. These are fluids where the component particles experience short-range gradually increasing repulsive forces and long range attractive forces (attractive because of London dispersion forces, which are attractive). The standard 12-6 Lennard-Jones potential is an example potential for van der Waals fluids:

$$U_{LJ}(r_i, r_j) = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{|r_i - r_j|} \right)^{12} - \left(\frac{\sigma_{ij}}{|r_i - r_j|} \right)^6 \right]$$

- ε_{ij} gives the depth of the potential trough
- σ_{ij} is the distance between the particles at which the potential is zero

The short-range forces are most important for determining the structure of the fluid at equilibrium! Changing the exponent of the repulsive term (power 12) can change when the liquid phase separates and forms a coexisting solid phase. The attractive term also gives rise to temperature sensitivity to the fluid (and hence the phase transition).

Phase transitions happen according to a phase diagram (gas, liquid, or solid phase).

2.3

In some biological systems, the model of self-avoiding polymers (high repulsive forces at short range) works, for example cDNA in certain rod-shaped bacteria. The self-avoiding polymers have an entropy reduction effect from being in close vicinity, so they form distinct chromosome-rich phases.

Even simply, entropically-modeled particles can explain some of the spatial organization of cells, even the simplest kinds.

Hard sphere transitions can also be catalyzed by active molecular steering.

3

3.1

In reality, cells are complex: made of many molecules, non-ideal solvents, etc.

Flory-Huggins Theory describes complex polymer solutions:

An important quantity is $\Delta\mu_{mix}$:

$$\Delta\mu_{mix} = k_B T \left(\sum_i \frac{\phi_i}{N_i} \ln \phi_i + \frac{1}{2} \sum_{i,j} \phi_i \phi_j \chi_{ij} \right)$$

For a binary system and my conservation of mass:

$$\Delta\mu_{mix} = k_B T \left(\frac{\phi_A}{N_A} \ln \phi_A + \frac{1 - \phi_A}{N_B} \ln(1 - \phi_A) + \phi_A(1 - \phi_A) \chi_{AB} \right)$$

This is the **free energy density of mixing**: This describes the free energy change per molecule of the system when transferring molecules from single-component, homogenous phases. This model assumes no higher order interactions (three body), no volume change upon mixing, and that the components are randomly dispersed in the system.

- ϕ_i is the volume fraction of component i
- N_i is the number of monomers in component i (around 1 for spherical macromolecules or small molecules)
- χ_{ij} is a pairwise interaction term between the components i and j compared to single-component interactions.

If χ_{ij} is non-positive (0 for ideal mixture with no energetic effect, or negative), then there will always at least be an entropic favoring of mixing.

Question

1. When does the pairwise estimation fail to hold? When do we need to consider higher order interactions?
2. For one-component systems, density transitions can occur, when only having an entropic effect. Why does this not scale to n -component systems, intuitively? (article mentions that there is change in the degrees of freedom, but what is the fundamental reason?)
3. Is the entropic term in the free energy density of mixing related to Shannon entropy from information theory?

A simulation can be found here for binary monomer mixing

3.2

When χ is positive, this indicates that beyond some threshold, there will be a liquid-liquid separation (ie, coexisting “dense” ϕ_{dense} and “dilute” ϕ_{sat} [dilute phase at the saturation concentration] phases of a molecule in the system).

Interesting consequence of this phase separation: there will be a interface between the dense and dilute phases, and thus solvent (low-molecular weight molecules) will transport to the dense phase to minimize the chemical potential difference (ie normalized particle count) created by this phase separation. But, this transport will generate an opposing osmotic pressure.

So, at chemical and osmotic equilibrium, ϕ_{dense} and ϕ_{sat} must satisfy:

$$\mu_{m,T}(\phi_{sat}) = \mu_{m,T}(\phi_{dense})$$

$$\Pi_{m,T}(\phi_{sat}) = \Pi_{m,T}(\phi_{dense})$$

Equalizing the chemical potentials (μ) and the osmotic pressures (Π) of the two phases.

As a note, theorists use ϕ (volume fraction) and experimentals use c (concentration).

In theory, polymer “melts” is a case of an almost perfect dense phase ($\phi \approx 1$), but, in vitro, the volume fraction is almost always at least one order of magnitude lower.

3.3

In a mixture with linear polymer A (with n_p monomers) and solvent B , the excluded volume per monomer/effective solvation volume v_{ex} or v_{es} is given by averaging over the intermonomer distances in the solution volume using a specified potential of mean force W (this is a quantity that represents the projection of a free energy profile onto a distance axis):

$$v_{ex} = - \int_0^\infty \left(e^{-\frac{W(r)}{k_B T}} - 1 \right) r^2 dr$$

$e^{-\frac{W(r)}{k_B T}} = g(r)$, the radial distribution function, so v_{ex} encapsulates the probable volume occupied by any given monomer by taking an infinite sum over the distribution of position states.

For $v_{ex} > 0$ means that the polymer is well-solvated, so on average the monomers are more repulsive towards one another. If the monomer-monomer and monomer-solvent interactions are similar then, $v_{ex} \approx 0$. And if the solvent is poor, then $v_{ex} < 0$. This can be used to define the “osmotic second virial coefficient” (B_2), which is a measure of the strength of the monomer-monomer vs monomer-solvent interactions.

$$B'_2 = B_2 M_w^2$$

Where M_w is the molecular weight of the polymer and B'_2 is:

$$B'_2 = -2\pi \int_0^\infty \left(e^{-\frac{W(r)}{k_B T}} - 1 \right) r^2 dr$$

$$B'_2 = 2\pi v_{ex}$$

$$B_2 = \frac{2\pi v_{ex}}{M_w^2}$$

So, B_2 changes as the potential of mean force changes (ie if we change pH, addinh H^+ to the solvent) or temperature or concentration changes. B_2 also negatively linearly depends on the Florey-Higgins χ interaction parameter, but takes into account pressure!

To measure the second virial coefficient we just need to plot osmotic pressure normalized to solute concentration vs the solute concentration (Zimm plot)

3.4

Florey-Higgins theory scales well when consider multiple polymers dissolved in a solvent, we just use matrix form:

$$\Delta\mu_{mix} = k_B T (\phi'_N \ln \phi + \phi' X \phi)$$

Where $\phi_n = (\phi_1/N_1, \dots, \phi_n/N_n)$, $\phi = (\phi_1, \dots, \phi_n)$ and X is the n by n square matrix denoting the χ_{ij} pairwise interactions.

Obviously, find out what the elements of X are will help determine the driving factors of phase-separation and what interactions are the most important to study.

As mentioned, there is a proportionality between χ_{ij} and B_{ij} , thus finding either X or B is equivalent.

We can populate the B matrix by keeping all concentrations fixed except for the i and j th interaction of interest and generating a Zimm plot.

Definition: scaffold: a macromolecular driver of phase separation in n -component mixtures.

People use the B matrix and identify the normalized top negative interactions to obtain the scaffolds.

3.5

Obviously, $\chi_{ij} = \chi_{ji}$, thus there are $\frac{n(n-1)}{2}$ free χ parameters.

The number of components n scales with the number of degrees of freedom and the number of coexisting phases p reversibly scales with the number of degrees of freedom:

$$F = n - p + 2$$

Why is the number of degrees of freedom defined like this?

$$p_{max} = n + 2$$

So, for $n = 1$, there must be a maximum of three possible phases coexisting (ie triple point).

Regarding the X or B matrix, it can be interesting to look at the variance σ_2 (pairwise) of the elements.

Jacobs and Frenkel made certain predictions based on a mean-field model they developed:

For σ_2 small and n large, then there will be an overall dilute and overall dense phase (with respect to any polymer). For σ_2 large and n small, then the system will be solvent balanced but rich in different polymers in different regions.

If the diagonal elements dominate, then there will be n distinct phases enriched in the corresponding polymer and a phase dilute in each in each polymer.

The order parameter ϕ only captures the global concentration of each component in a given phase and thus cannot be used to probe local effects/inhomogeneities.

3.6

At least in aqueous solutions, temperature is a important parameter to determine phase diagrams (when phase transitions occur)

thermoreponsive phase behavior: The phase behavior in response to temperature changes

Polymer solutions can have an upper or lower critical solution temperature (UCST/LCST), which posits the existance of a maximum/minimum temperature when there is no longer any phase separation.

For UCST systems, upon approaching this critical temperature χ becomes less positive by an enthalpically driven process where polymer-polymer attractions become weaker.

For LCST systems, there is an entropic penalty to having the solvent “well-solvate” the polymer (ie interact with it strongly, as they become stuck), and this penalty increases with higher temperatures, so after a certain critical temperature we will observe phase separation.

Systems can be a combination of both UCST and LCST (elliptical or hyperbolic type phase diagrams).

4

4.1

A simple associative molecule: a hard sphere with an attractive patch, where three properties are important:

- size
- interaction strength
- types of organizational structures possible

In a gas, the time scale of inter-particle cross-links will be determined by the strength of the interaction.

In a liquid, this is true but also density will play a role.

Regarding organization, for example, if we have a system of two-patch particles and if the patches are an order of magnitude smaller than the particle, then the patches will be able to form linear network.

Intuitively, why does polymer organization depend on the size of the patch?

For a system of n -patch particles, there will be a characteristic distribution of cluster sizes, which shifts upward as particle concentration increases, until it networks the whole system (starting point at c_{perc}). As the concentration grows above c_{perc} , extra particles begin incorporating in the largest cluster.

Below c_{perc} there can be a wider degree of concentration inhomogeneity, especially comparing the bulk concentration to the small cluster concentrations, but this is tampered after c_{perc} (more homogenous), as clusters merge into one system spanning one.

The relation between connectivity and concentration constitutes a relevant order parameter (ie characterization) for a percolation.

Chemically-crosslinked gels can swell or shrink to include or exclude solvent, whereas physically-crosslinked gels forms/breaks apart according to solvent changes.

Percolation threshold is a function of valence (ie number of patches per particle) and the strength of the of each cross-link.

4.2

For the type of network described before, the number of molecules within a cluster is the **molecularity** of the cluster

microscopic binding reaction: x number of molecule A come together to form an associative cluster or x number of A and y number of B form a cluster according to their valency-determined stoichiometry.

There is an opposite limit - **isodesmic associations:** These don't involve exact stoichiometries and this process is characterized by the average molecularity of species in solution always increasing with concentration and there being an elementary association constant.

Now, consider a system of either patchy particles or linear polymers with n_s stickers. There is a mean field theory, assuming a "gas" of stickers (like spherical particles with patches or intermolecularly interacting linear polymers)