Examination of Nonlocal Screening Effects on Protein Crystallization

by

Sawyer S. Hopkins

B.S., Michigan Technological University, 2014

A THESIS

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2017

Approved by:

Major Professor
Jeremy Schmit

Copyright

Abstract

Over twenty percent of amino acids are ionized under biological conditions, and the subsequent electrostatic interactions have substantial effect on protein crystallization, binding, catalyzation, and recognition. These electrostatics along with other intermolecular forces create a delicate balancing act of repulsive and attractive forces. This thesis explores the effects of electrostatics on the formation of dense ordered structures.

In dense protein aggregates the repulsive electrostatics are dominated by the entropic cost of compressing salt ions in the electrostatic screening layer. A non-local electrostatic interaction was derived to describe this behavior, and was used to examine the interplay of attractive energies and repulsive entropy on protein colloid stability and the crystallization process.

Using a simple analytical model it was predicted that the derived electrostatic effects describe a finite window in phase space in which crystallization can occur. This simple model was expanded upon via computational methods simulating hard spherical particles aggregating under short-ranged attractive interactions and the repulsive electrostatics.

From the computational simulations phase and dynamical data was extracted to confirmed the initial insight of the analytical model. The simulations also introduced new information not described by the simple model, most notably a metastable amorphous phase caused by the competition of energies and entropies.

Table of Contents

List of Figures vi

1 Introduction 1

2 Nonlocal Electrostatics 4

2.1 Poisson-Boltzmann 4

2.2 Free energy 5

2.3 Screening effects 7

2.4 Phase Diagrams 9

3 Computational Methods and Techniques 11

3.1 Equations of Motion 11

3.2 Parameter Space 12

3.3 Seeding and Placement 13

3.4 Linear Comparison 14

4 Numerical Results 16

4.1 Phase Definitions 16

4.2 Phase Diagrams 19

4.3 Structure and Dynamics Analysis 20

5 Conclusion 23

5.1 Summary 23

5.2 Future Work 24

Bibliography 25

A Computational Units 28

B Extended Figures And Plots 30

List of Figures

2.1 Illustration of a single protein monomer represented as a hard sphere surrounded by a spherical electrostatic screening layer. 6

2.2 The excluded volume v_e created via the overlap of electrostatic screening layers. 8

2.3 (left) Free energy from screening effects for varying protein charges, Q. (right) Relation of free energy and protein charge for varying concentrations charges c_s. Both are presented at a static volume of $v = 0.5$. Values represented are unitless. 8

2.4 Ideal phase diagrams for a three state system, showing the relationship between protein charge Q and concentration ρ with attractive potential energy ϵ. 10

3.1 Depiction of complex screening layer overlap. 13

3.2 The functional behavior of the electrostatic energy f when consider as a function of coordination for varying protein charges q in a one molar salt solution. 15

4.1 Typical distribution of coordination number for specified aggregate phases. . 17

4.2 The structure of the four characterized phases shown in both a per particle view (a-d) and a surface overlay (e-h). Solution (a,e). Amorphous (b,f). Gel (c,g). Crystal (d,h). Periodic boundary conditions apply to all images. . . . 17

4.3 Example of an amorphous structure showing no defined fcc ($q_6 = 0.57$) or hcp ($q_6 = 0.48$) indicators in local bond order, and a gel aggregate displaying both fcc and hcp indicator spikes. 19

4.4 Averaged local bond order parameters $\bar{q_6}$ and $\bar{q_4}$ for charge $Q = 15$. Ordered structures roughly correspond to the region of $0.3 < \bar{Q}_6 < 0.6$ and $0 < \bar{Q}_4 < 0.2$ 20

4.5 Phase diagram for the non-linear and Debye-Hückel interactions. Sampling of data point occurs at $5Q$ and 0.25ϵ increments. See appendix B. 20

4.6 The dynamics of the average number of particles in each cluster in time for charge $Q = 15$. 22

B.1 Phase diagram for the non-linear and Debye-Hückel interactions. Sampling of data point occurs at $5Q$ and 0.25ϵ increments. Sampled data marked as circles. 30

B.2 The time evolution of the average coordination number of the four identified phases associated with non-local electrostatics 31

Chapter 1

Introduction

A colloid is a two phase dispersion of particles, in which the particles being dispersed are referred to as the dispersed phase, and the medium in which they are being dispersed is known as the continuous phase. The dispersed phase is composed of particles with one dimension that is at least one nanometer and less than one micrometer, such as proteins and nanoparticles[1]. As proteins are commonly on the scale of several nanometers[2], globular in shape[3], and charged under biological conditions[4 5], this paper will consider proteins as charged spherical colloidal particles.

The effects of electrostatics on the stability of protein solutions, and the ability of salts to increase or decrease the solubility of a solution (salting in/out) has been recognized for over one hundred years.[6 7] Derjaguin, Landau, Verwey, Overbeek (DLVO) theory is the classic attempt to describe this effect for charged colloid systems. DLVO theory considers two interactions, an attractive Van der Walls attraction between particles and repulsive electrostatics.[8 1] In this model charged particles attract salt counter-ions and repel salt co-ions. This creates a charged shell around each particle which quickly decays in intensity to that of the neutral bulk solution. The thickness of this charged sphere is described as the screening length, and corresponds to an electrostatically screened volume around each colloidal particle referred to as the screening layer. The colloidal particles described by

DLVO theory are assumed to be much larger in size than their screening length.[9]

In this theory, the Van der Walls potential, created by dipole-dipole, induced-dipole, and dispersion forces, causes these particles to aggregate.[10] As two particles move closer to each other, there is an overlap, or rather compression, of their charged shells. This compression decreases the volume around the particles accessible to salt counter-ions; creating an entropic penalty towards aggregation. These two interactions create an opposing set of forces that dictate the colloid's stability.

For solutions with low salt concentrations, the screening volume around each particle is large enough that the interaction between particles is dominated by pair-wise Coulomb interactions. In these dilute solutions a linear approximation to the electrostatics known as the Debye-Hückel equation is often used, rather than an exact solution to the Poisson-Boltzmann equation. However it has been shown in Schmit, Whitelam, Dill 2011[11], that in dense aggregates that this assumption is not only incorrect, but that the repulsive forces are dominated by non-pairwise entropic forces.

For example consider building a dense crystalline structure around a stationary central particle, one particle at a time. As each particle is removed from the dispersed phase and added to the central structure, there must be an attractive interaction binding the particles together. This attractive interaction is necessary to prevent the immediate reversibility of this addition to the crystal due to random thermal motion and internal repulsive forces from electrostatics.

However, as each particle is added, the volume around the central particle which salt counter-ions are allowed to occupy is further decreased. This results in each subsequent particle facing a higher entropic cost than the prior. This higher entropic cost results in greater internal repulsive forces and a greater tendency toward the reversal of the addition process.[10] As further particles are added to the system, the non-pairwise electrostatics, which are again negligible in sparse systems, become the dominant repulsive interaction between particles.

Due to the decreasing stability, and increasing propensity for a particle to reenter the dispersed phase, there exists an intricate balance between attractive energies driving the system towards a crystalline like state, and entropy favoring a dispersed or open gel-like phase. Together the energy and entropy dictate the stability of colloids, and create a finite window in phase space in which crystallization can occur.[12]

Experimentally precipitating crystals is a tedious process of scanning a complex parameter space that includes, but is not limited to, temperature, salt and protein concentrations, and pH.[12] This thesis will approach the concepts of non-local electrostatics as described above via computational means. This will first be done by deriving a non-local interaction between protein molecules. Phase and dynamics data will then be collected from both Debye-Hückel and the derived interaction. Direct comparison will be made between these two models to identify the effects of non-local electrostatics on the crystallization process.

Chapter 2

Nonlocal Electrostatics

2.1 Poisson-Boltzmann

In a protein salt solution, the total electrostatic energy, f_{es}, is given by the sum of the coulomb potential energy, $f_{coulomb}$, and the free energy of the salt solution, f_{salt}; in which ρ is the local charge density in the solution, Ψ the electrostatic potential, and S the entropy of the salt solution.[5]

$$\begin{aligned} f_{coulomb} &= \frac{1}{2}\int_V \rho\left(r\right)\Psi\left(r\right)\mathrm{d}V \\ f_{salt} &= -TS \\ f_{es} &= f_{coulomb} + f_{salt} \end{aligned} \tag{2.1}$$

The electrostatic potential is found as a solution to the Poisson equation given in equation 2.2.

$$\epsilon\nabla^2\Psi = -\rho \tag{2.2}$$

The Boltzmann distribution defines the local ion concentration of a solution by the probability of a salt ion existing at a given energy level within an electric field. This is shown in equation 2.3, in which c_0 is the bulk solution concentration, z is the valency of the salt ion,

and q is the fundamental electron charge.

$$\rho \propto c = \sum_i c_0 e^{-z\mathrm{q_i}\Psi/kT} \tag{2.3}$$

The combination of these two equations results in the non-linear partial differential equation known as the Poisson-Boltzmann equation. There are various approaches to approximating a solution to this equation, the likely most well known of which for spherical charges is the Debye-Hückel approximation given by equation 2.4[13], in which U_{dh} is the electrostatic potential energy.

$$U_{dh}(r) = \gamma \ell_d \frac{e^{r/\ell_d}}{r} \tag{2.4}$$

$$\ell_d^2 = \frac{\epsilon kT}{4\pi \sum_i c_{i_0} z_i q^2} \tag{2.5}$$

In this equation ℓ_d is referred to as the Debye screening length and is calculated in equation 2.5, where subscript i denotes ion species. The following sections will derive a non-local solution to the Poisson-Boltzmann equation, for which this solution will serve as a base point of reference and comparison.

2.2 Free energy

Figure 2.1 depicts a protein represented as a spherical particle of radius R surrounded by an electrostatic screening layer of thickness a. In a salt solution, aggregation of neutrally charged proteins occurs without penalty from the electrostatics in the solution. Aggregation becomes entropically unfavorable for charged proteins due to the local enrichment of counterions within the screening layer. This penalty is given by the chemical potential caused by the change in local ion concentration, c, from the fixed solution concentration, c_s, as shown in equation 2.6[11 14 15].

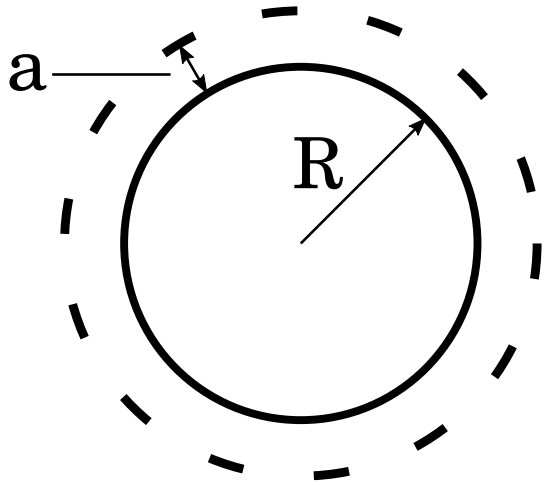


Figure 2.1: Illustration of a single protein monomer represented as a hard sphere (solid) surrounded by a spherical electrostatic screening layer (dashed).

$$\frac{\Delta S}{k_B} = ln\left(c\right) - ln\left(c_s\right) = ln\left(\frac{c}{c_s}\right) \tag{2.6}$$

The entropy for a monovalent solution is obtained by first integrating equation 2.6 from the bulk solution concentration to the local positive and negative ion concentrations, and then integrating over the solvent accessible volume in the screening layer.

$$\frac{S}{k_B} = \int_V \left(\mathrm{d}^3 r \sum_{i=(+,-)} \int_c^{c_i} ln\left(\frac{c'}{c_s}\right) \mathrm{d}c' \right) \tag{2.7}$$

As previously posited the Coulomb energy as defined in equation 2.1 will be dominated by the non-local interactions,[11] and as such will be neglected from this point forward. The total electrostatic free energy of f_{es} will consist only of contributions from the salt entropy of the solution, and will further be referenced simply as f.

$$\frac{f}{k_B} = -TS = T\int_V \left[c_+ ln\left(\frac{c_+}{c_s}\right) - c_+ + c_s\right] + \left[c_- ln\left(\frac{c_-}{c_s}\right) - c_- + c_s\right] \mathrm{d}^3 r \tag{2.8}$$

Equation 2.9 is then found by substituting the Boltzmann relation previously given in

equation 2.3 into equation 2.8.

$$\frac{f}{k_B T} = 2c_s \int_V \left[\Psi' \sinh(\Psi') - \cosh(\Psi') + 1\right] \mathrm{d}^3 r \tag{2.9}$$

$$\Psi' = \frac{e\Psi}{k_B T} \tag{2.10}$$

2.3 Screening effects

Using the assumption of charge neutrality, it is posited that a protein of charge Q must have an equal ionic counter charge in its screening layer. This charge neutrality relation is given by equation 2.11, in which v is the solvent accessible volume of the protein's screening layer.

$$Q = -v\left(c_+ - c_-\right) = 2vc_s \sinh(\bar{\Psi}') \tag{2.11}$$

The free energy given in equation 2.12 is derived under the approximation that at high protein concentrations the electrostatic potential can be considered constant, in which $\bar{\Psi}'$ is the average electrostatic potential in the screening region (Jellium approximation)[16]. Using this approximation the integral in equation 2.9 becomes trivial. Combining equations 2.11 and 2.9 the following solution is found.[17]

$$\frac{f(\zeta)}{k_B T} = Q\left(\sinh^{-1}\left(\frac{1}{\zeta}\right) - \sqrt{1+\zeta^2} + \zeta\right) \tag{2.12}$$

$$\zeta(v, Q, c_s) = \frac{2vc_s}{Q} \tag{2.13}$$

The solvent accessible volume, v, is given by equation 2.14; in which v_t is the volume of a sphere with radius $R_d = R + a$ (see Figure 2.1), v_p is the protein volume given by a sphere of radius R, and v_e the volume excluded from the solution due to the overlap of screening layers.

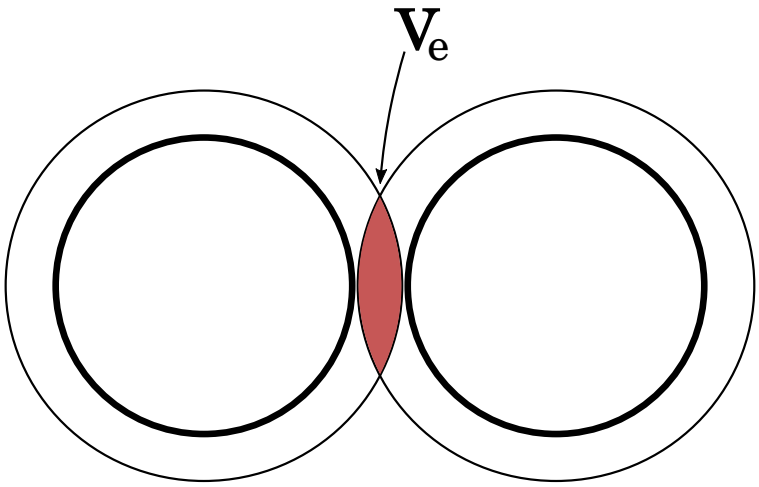


Figure 2.2: The excluded volume v_e created via the overlap of electrostatic screening layers.

$$v = v_t - v_p - \frac{1}{2}v_e \tag{2.14}$$

For a protein i with n neighbors with centers within a sphere of radius $2R_d$ from the center of i, the excluded volume is given by equation 2.15[18].

$$v_{e_i} = \sum_{j=1}^{n} \frac{4\pi}{3} R_d^3 \left[1 - \frac{3r_{ij}}{4R_d} + \frac{1}{16}\left(\frac{r_{ij}}{R_d}\right)^3 \right] \tag{2.15}$$

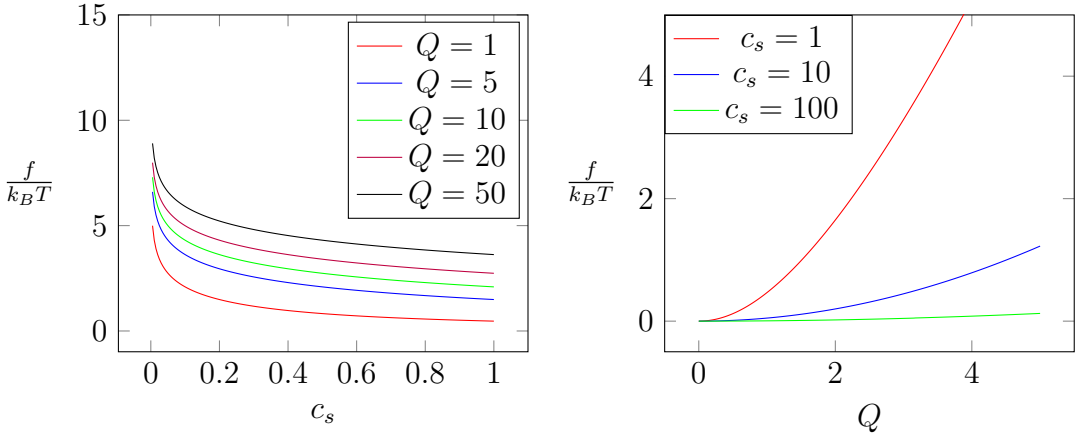


Figure 2.3: (left) Free energy from screening effects for varying protein charges, Q. (right) Relation of free energy and protein charge for varying concentrations charges c_s. Both are presented at a static volume of $v = 0.5$. Values represented are unitless.

Asymptotically this free energy vanishes as the protein becomes more neutral. This is

expected due to the lack of local ion enrichment around the neutrally charged proteins. This behavior is demonstrated in Figure 2.3 which also shows the initial non-linear response in the free energy in relation to small protein charges.

The width of the screening layer, a, is inversely proportional to the square root of concentration, c_s; creating an implicit relationship between c_s and v. For this reason the relationship between f and c_s and the corresponding limits should be considered with care.

2.4 Phase Diagrams

Consider three ideal states: a single particle monomer, a linear gel, and a dense FCC or HCP crystal. These states will serve as the basis for a simple model, however are unrealistic due to the constraints on growth and dimensionality. These states will be allowed to develop in a salt solution that provides some inter-particle attraction F_A and a repulsive interaction f.

In this system there exists a pseudo-equilibrium between the monomer and crystal states described by $f = \epsilon$ and a pseudo-equilibrium between the monomer and gel states described by $f << \epsilon$. The area between these two equilibria defines the region in parameter space in which crystallization can occur. These boundary conditions have been previously been derived in Schmit, Whitelam, Dill 2011[11] for F_A given as a Van der Waals potential, and is shown in figure 2.4.

These phase diagrams are idealized and the structure of the corresponding aggregates unrealistic. However, this model is useful in gaining initial intuition into the complex forces at work. The next chapter will discuss building a computational model in which to simulate crystalline growth from the derived free energy. This computational modeling will provide two key advantages over this ideal model. The first being that aggregates will be allowed to grow freely in three dimensions, rather than being dimensionally restricted. The second advantage being the inclusion of random motion via Brownian Dynamics.

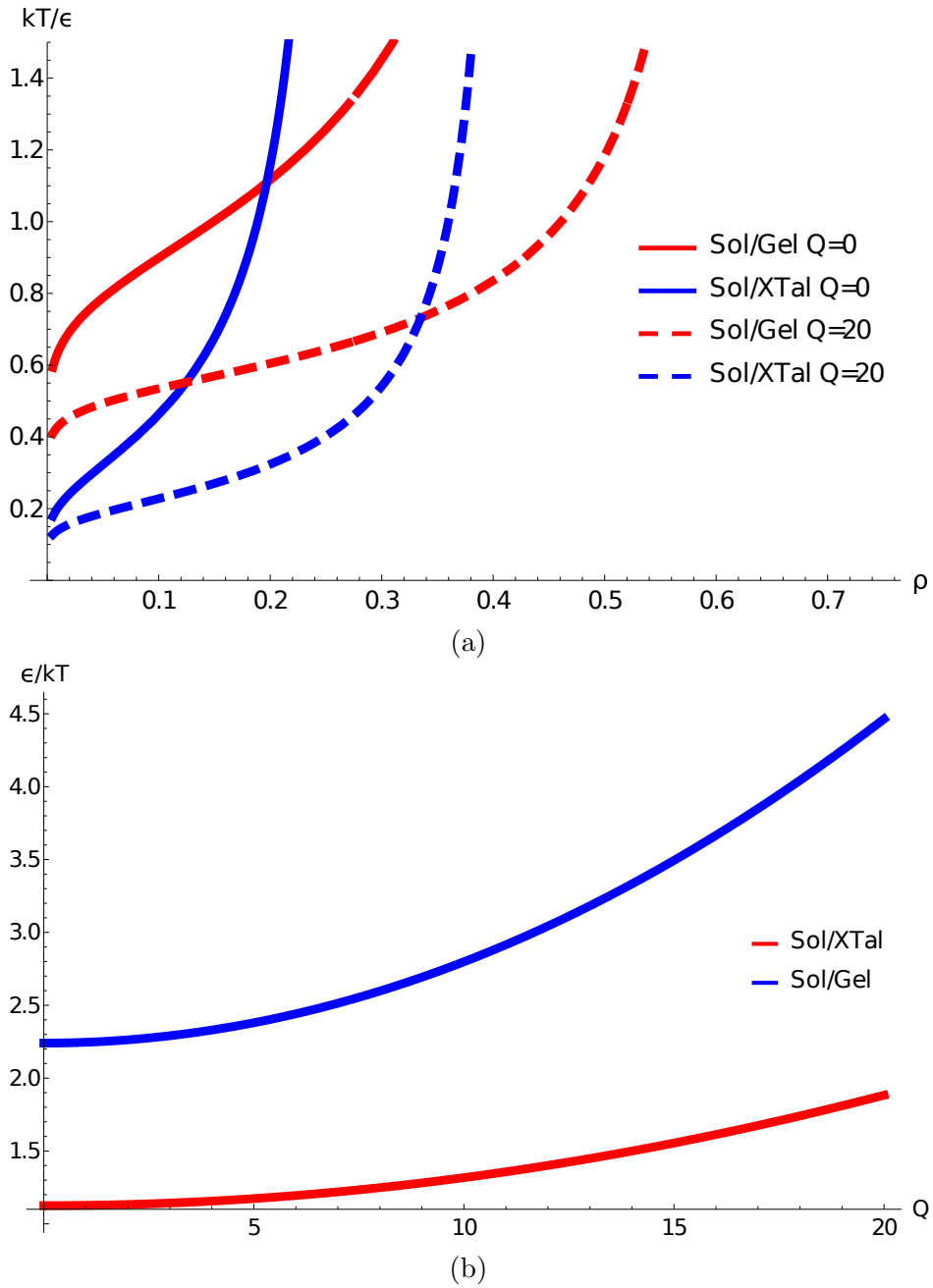


Figure 2.4: Ideal phase diagrams for a three state system, showing the relationship between protein charge Q and concentration ρ with attractive potential energy ϵ.

Chapter 3

Computational Methods and Techniques

3.1 Equations of Motion

A particle system consisting of spherical particles of uniform radius evolving under the influence of one or more systematic forces ($F_i(r)$), a stochastic force ($R_i(r)$), and viscous drag (Γ) was used to examine protein crystallization processes. This system was defined by the Langevin equation as shown in equation 3.1.[19]

$$m_i\ddot{r}_i = F_i(r) - m_i\Gamma\dot{r}_i + R_i(t) \tag{3.1}$$

The viscous drag term describes the linear drag experienced by particles due to motion through the solvent, in which Γ describes the strength of the drag.

Here the stochastic term is used to simulate the random walk of particles through the solvent due to Brownian motion. It is described by a stationary Gaussian satisfying the following two conditions where bracketed terms denote an ensemble average.[20]

$$\langle R_i \rangle = 0 \tag{3.2}$$

$$\langle R_i(t) R_j(t) \rangle \propto \delta_{ij} \delta(t) \tag{3.3}$$

This stochastic force is also created such that it does not have any correlation between previous velocities or systematic forces, requiring two additional constraints on this force to be met for anytime greater than zero.[20]

$$\langle \dot{r} R_i(t) \rangle = 0 \tag{3.4}$$

$$\langle \ddot{r} R_i(t) \rangle = 0 \tag{3.5}$$

In this simulation the systematic force can be subdivided into two forces. The first being a short ranged attractive interaction with a hard core repulsion ($F_a(r)$), and the other being the repulsive non-local electrostatic term ($F_{es}(r)$). The short range attractive force encompasses H-Bond, Van der Waals, and hydrophobic effects, and was modeled using an extended Lennard-Jones potential of the form shown in equation 3.6, in which the particle diameter $\sigma = 1$, $K_b T = 1$, and the term ϵ is used as a control variable to tune the attractive strength of the force.

$$V_a(r) = -\int F_a(r)\, \mathrm{d}r = 4\epsilon K_b T \left[\left(\frac{\sigma}{r}\right)^{50} - \left(\frac{\sigma}{r}\right)^{25} \right] \tag{3.6}$$

3.2 Parameter Space

The effect of equation 2.12 was examined at a one molar salt concentration. This high salt concentration results in screening layers which are small enough to only produce binary overlap in highly ordered crystals. Here binary overlap describes two screening layers overlapping in a given region of space. More complex higher order overlaps would involve three or more screening layers. These complex overlaps begin to occur once the thickness of screening

layers are approximately 15% of the particle diameter. This overlapping of screening layers can be seen in the Figure 3.1

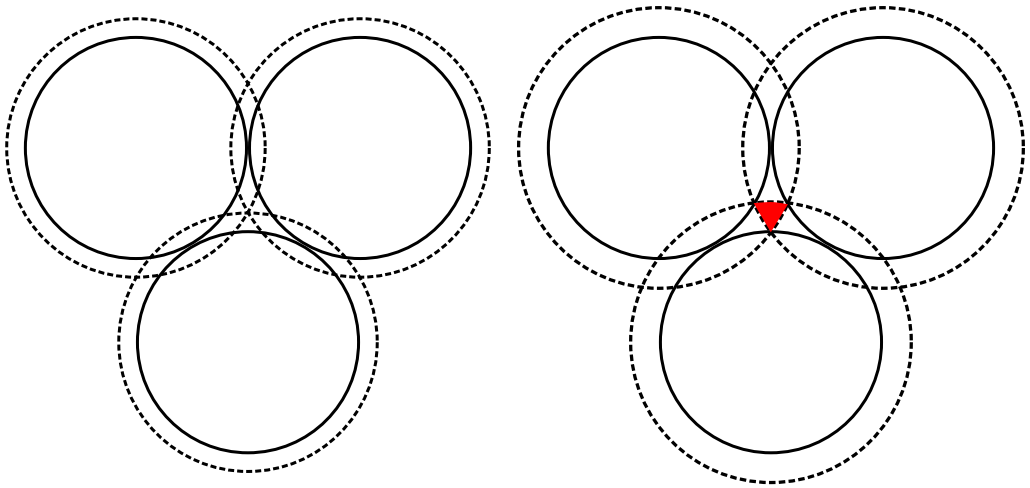


Figure 3.1: (Left) Excluded volume formed via binary interactions involving only two screening layers. (Right) A complex excluded volume (highlighted in red) is created from the overlap of three or more screening layers.

The system examined contained 2500 uniform spherical particles in a square box at a volume concentration of 9.5%. The particles evolved through the integration of the equations of motions provided in section 3.1 with a time-step of 0.001 (unit-less). During this integration periodic boundary conditions were implemented across all surfaces. Integration was halted for all trials after 10^8 integration cycles. Simulations were run adjusting the attractive potential strength, ϵ (Equation 3.6) in $0.25K_BT$ increments. The entire set of unit-less computational values can be found in appendix A.

3.3 Seeding and Placement

To assist in the initial nucleation of crystalline and gel structures a 4x4x4 primitive cubic seed crystal was placed in the center of each box at the beginning of each simulation. Moreover to ensure the possibility of valid comparison between simulations each box was seeded with an identical initial state of particle positions and velocities.

3.4 Linear Comparison

The derived electrostatics differs from the linear interaction described by Debye-Hückel in two important ways. The first being that the cost of adding a particle to an aggregate via Debye-Hückel is constant, where as it increases with each subsequent particle under non-local electrostatics. Second Debye-Hückel describes a screening layer which smoothly decays to the solution equilibrium, whereas the non-local electrostatics are derived under the assumption of a 'hard shell screening layer' which can be described by a step function. That is to say that the non-local interaction has a hard cutoff at the end of its screening layer.

A way was needed to compare the costs of these two interactions when building dense ordered structures. To do this consider a distance $R_c = 2R$, which is the distance at which two particles are directly in contact. Using this fixed distance the excluded volume in equation 2.15 can be rewritten as follows, in which a is the thickness of the screening layer as shown in Figure 2.1.

$$v_{e_i}(n) = n\pi a^2 \left[\frac{4a}{3} + R_c\right] \tag{3.7}$$

This allows the electrostatic free energy of a particle to be thought as a function of its charge, salt concentration, and the number of particles it is in contact with, n.

$$f(\zeta) \rightarrow f(n, Q, c_s) \tag{3.8}$$

The comparison point between Debye-Hückel and the non-local electrostatics can then be made at $n = 1$. In this special case for the non-local interaction, the cost of binding this one particle to another is only determined by the particle itself, and not those previously added. Combining equations 2.4 and 3.8 results in the following solution for the equivalent magnitude of the Debye-Hückel potential for a given non-local electrostatic energy.

$$f(1, Q, c_s) = U_{dh}(R_c, a) \tag{3.9}$$

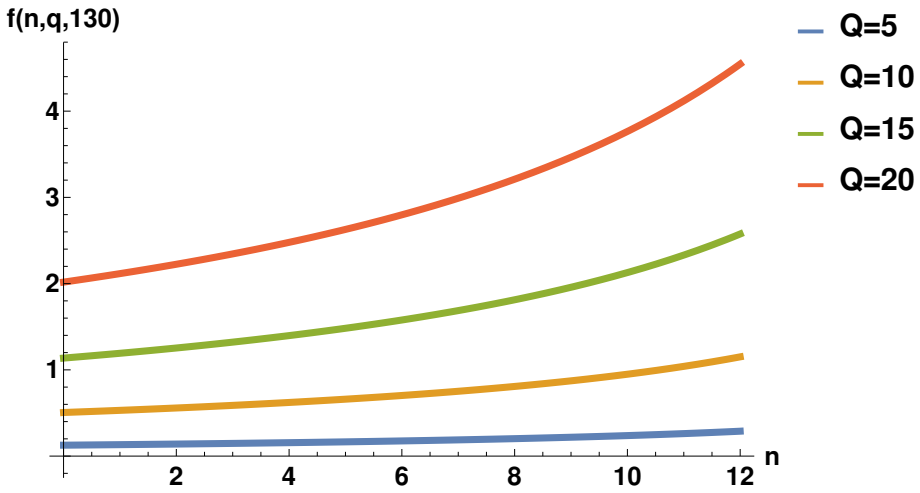


Figure 3.2: The functional behavior of the electrostatic energy f when consider as a function of coordination for varying protein charges q in a one molar salt solution.

$$\gamma = K_b T \frac{R_c}{a} e^{R_c/a} f\left(1, Q, c_s\right) \tag{3.10}$$

In this equation a is kept constant and Q and c_s are left as free variables for the electrostatics. Such that the Debye-Hückel potential can be simply thought as:

$$U_{dh}\left(r\right) = \gamma\left(Q, c_s\right) a \frac{e^{r/a}}{r} \tag{3.11}$$

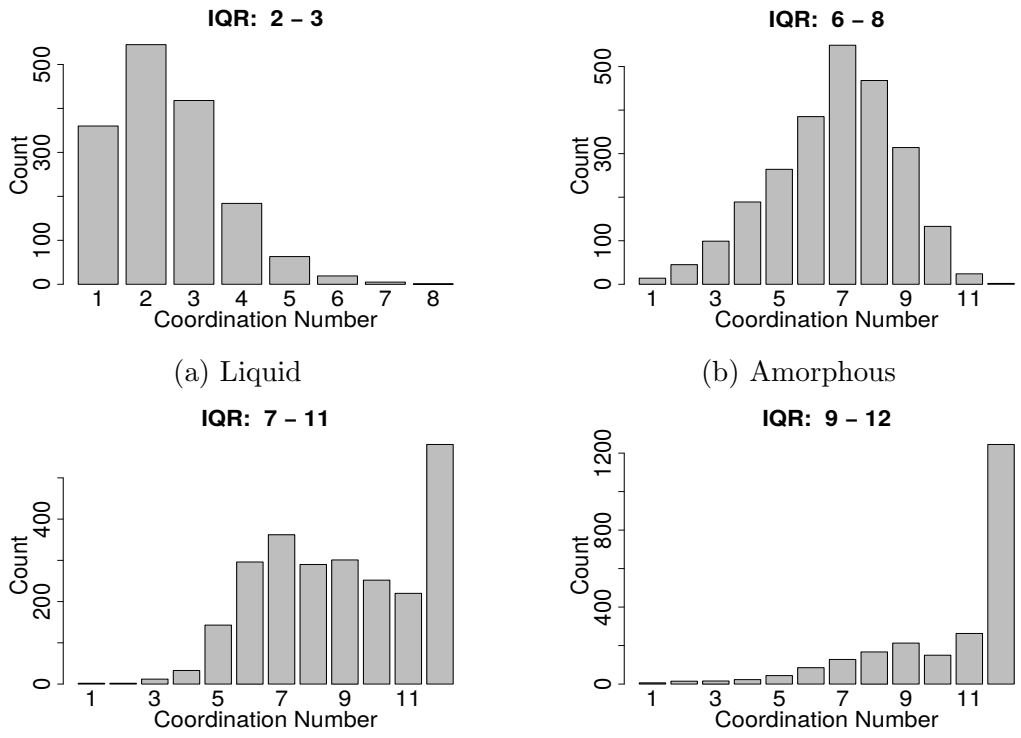
Chapter 4

Numerical Results

4.1 Phase Definitions

Three phases were uniquely identified via a histogram of their coordination number shown in figure 4.1. The solution phase was characterized by a peak coordination number of three or less. A gel was defined as an aggregate that is more structured than a solution, but the inner quartile range of the coordination numbers does not contain twelve[21]. Likewise a crystal would be a structure in which the IQR does contain twelve. Examples of these histograms are shown in figure 4.1.

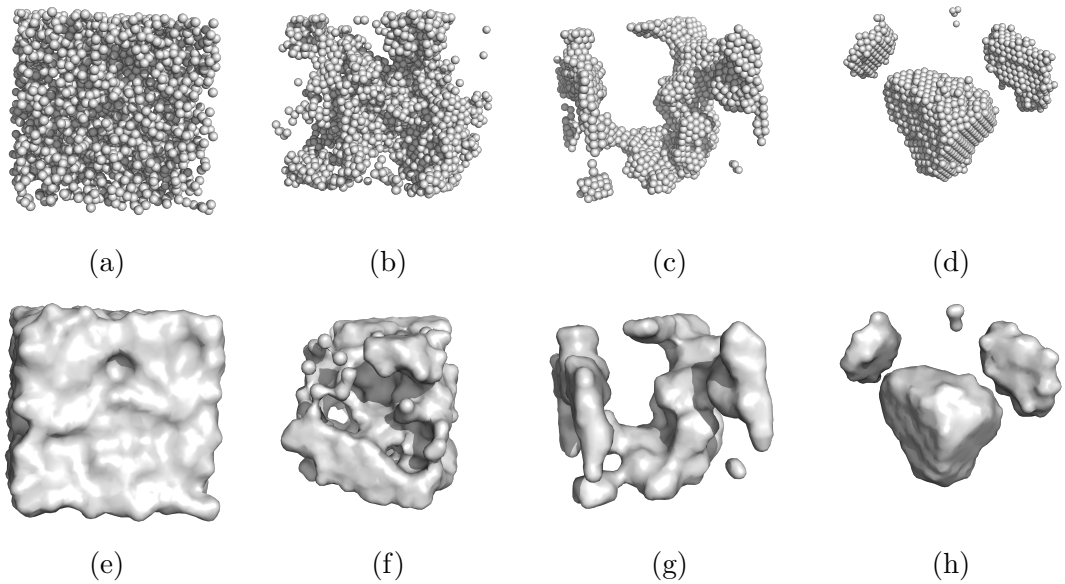
A fourth phase, not recognized in the idealized model described in chapter 2, was identified and from here on will be referred to as the amorphous phase. The distribution of coordination number throughout its structure would classify the phase as a gel, however the amorphous phase is qualitatively different than a gel. Examples of these histograms are shown in figure 4.1. The time evolution of the mean coordination number of these phases can be found in Appendix B Figure B.2.



(a) Liquid (b) Amorphous

(c) Gel (d) Crystal

Figure 4.1: Typical distribution of coordination number for specified aggregate phases.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 4.2: The structure of the four characterized phases shown in both a per particle view (a-d) and a surface overlay (e-h). Solution (a,e). Amorphous (b,f). Gel (c,g). Crystal (d,h). Periodic boundary conditions apply to all images.

As seen in figure 4.2, visually the amorphous phase is opaque, consuming most of the volume accessible to it, much like the liquid phase. Though the gel phase favors an open structure, maximizing surface area, and therefor consuming more volume than a crystal, it does this in a way such that particles are organized. In contrast the amorphous phase lacks any discernible structure. To quantitatively define the amorphous phase, describe its structure, or rather lack thereof, and distinguish it from the other three phases, local bond order parameters were calculated[22 23].

A bond order parameter for a particle i with n nearest neighbors is calculated by a weighted projection of the spherical harmonics of the n neighbors in respect to particle i.

$$q_{lm}(i) = \frac{1}{n(i)} \sum_{j=1}^{n(i)} Y_{lm}(r_{ij}) \tag{4.1}$$

$$q_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_{lm}(i)|^2} \tag{4.2}$$

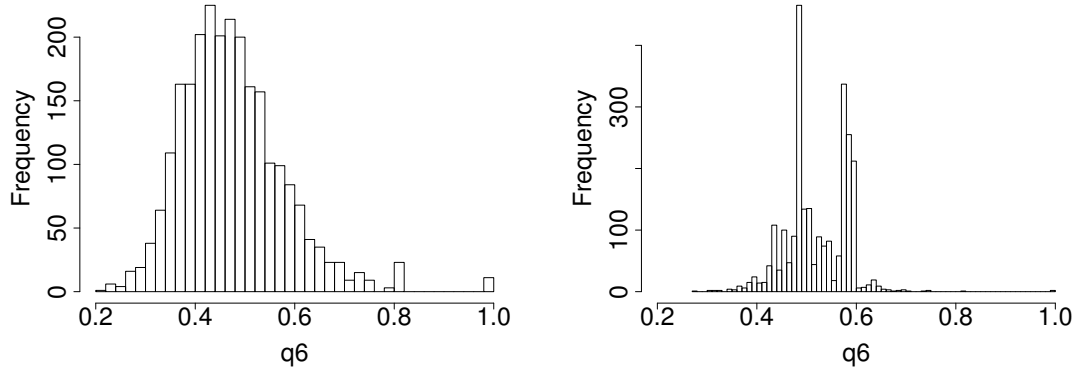
The averaged bond orders $\bar{q}_{lm}(i)$ is found by averaging the harmonic projections of particle i with those of the neighboring n particles. This is shown in equation 4.4.

$$\bar{q}_{lm}(i) = \frac{1}{n(i)} \sum_{k=0}^{n(i)} q_{lm}(k) \tag{4.3}$$

$$\bar{q}_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\bar{q}_{lm}(i)|^2} \tag{4.4}$$

This analysis was performed for harmonics $l = 6$, identified by q_6. It was found from this analysis that structures appearing to be amorphous showed no defined structure in their bond order parameters. This unordered structure is characterized the lack of well defined peaks at $q_6 = 0.57$ (fcc) and $q_6 = 0.48$ (hcp), which can be seen in figure 4.3. This lack of structure will be used as the primary metric for defining the amorphous region in the phase diagram.

Further examination of the local bond order parameters shown in figure 4.4 shows that the amorphous phase exists in a region of $\bar{q}_6\bar{q}_4$ space that typically describes a liquid phase. The gel and crystalline phases exist in the region $\bar{Q}_6\bar{Q}_4$ space corresponding to FCC, HCP, and BCC structures. Interestingly the amorphous phase and liquid phase have significant overlap within this space.



(a) Amorphous

(b) Gel

Figure 4.3: Example of an amorphous structure showing no defined fcc ($q_6 = 0.57$) or hcp ($q_6 = 0.48$) indicators in local bond order, and a gel aggregate displaying both fcc and hcp indicator spikes.

4.2 Phase Diagrams

The phase diagram of the solution shows that the non-linear interaction has the same crystal and gel boundaries as the pair-wise model, satisfying previously stated expectations. The width of the crystallization window is 1ϵ which corresponds to the prediction of the idealized model shown in figure 2.4. The width of this window also shrinks with respect to Q. This behavior is also predicted in the idealized model as described in Schmit, Whitelam, Dill 2011[11]. In this diagram the crystalline phase vanishes past $Q = 15$, suggesting that the width of the crystallization window is less than the scanning increment of $0.25kT$ or zero.

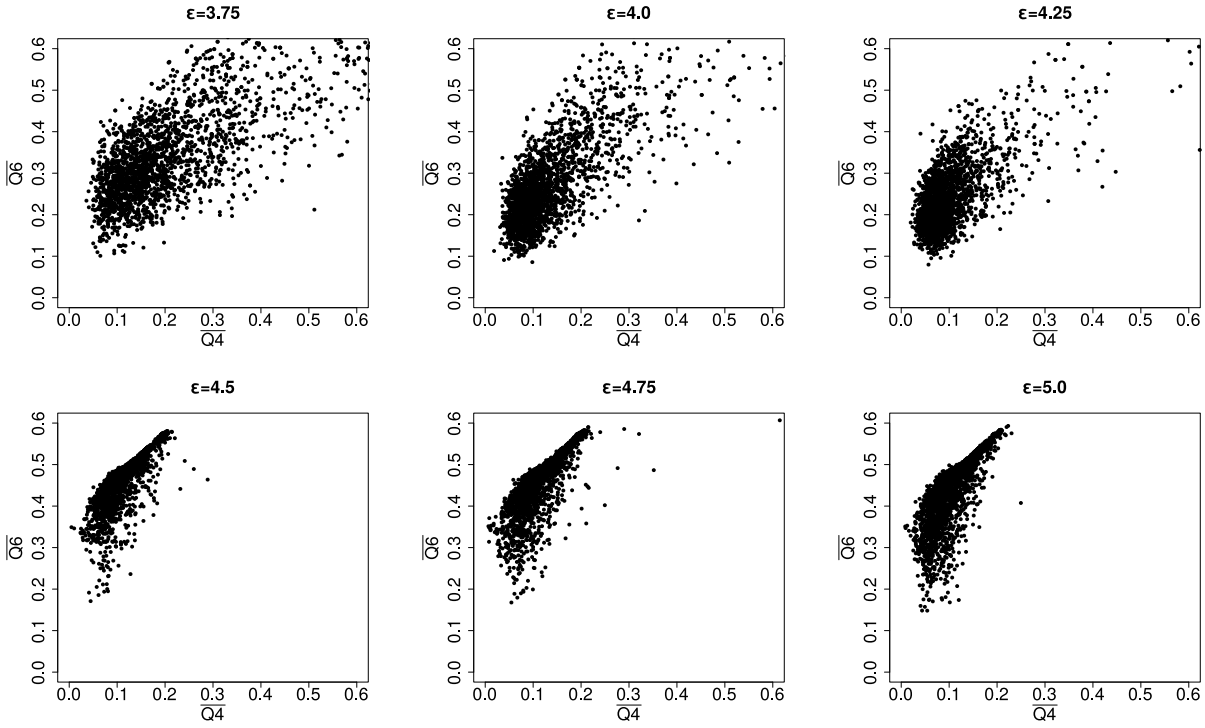
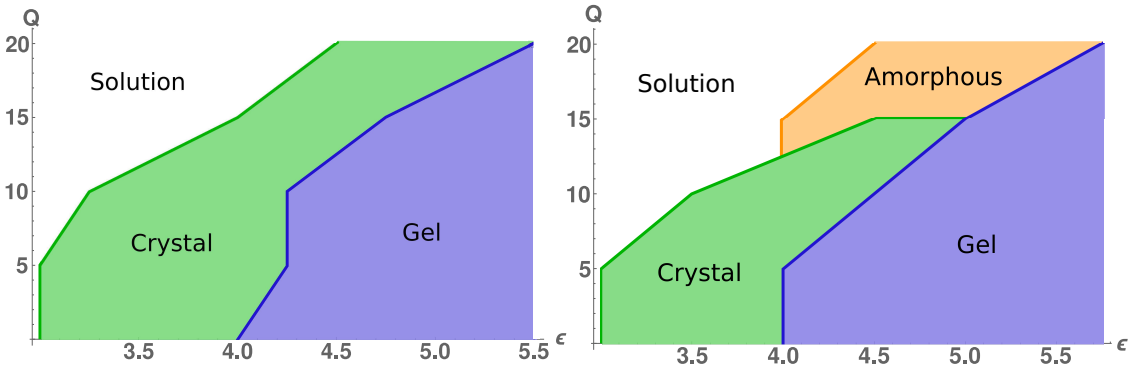


Figure 4.4: Averaged local bond order parameters $\bar{q}_6$ and $\bar{q}_4$ for charge $Q = 15$. Ordered structures roughly correspond to the region of $0.3 < \bar{Q}_6 < 0.6$ and $0 < \bar{Q}_4 < 0.2$



(a) Debye-Hückel (b) Non-linear

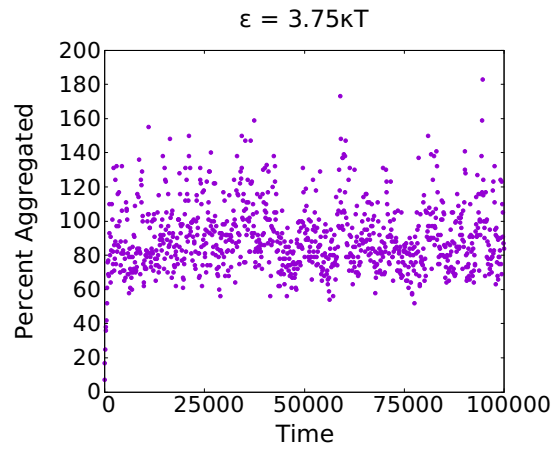
Figure 4.5: Phase diagram for the non-linear and Debye-Hückel interactions. Sampling of data point occurs at $5Q$ and 0.25ϵ increments. See appendix B.

4.3 Structure and Dynamics Analysis

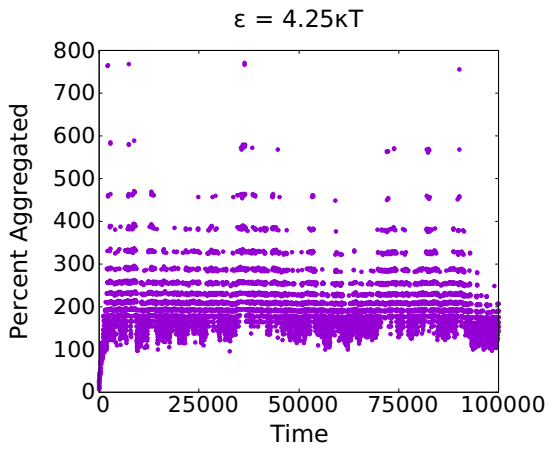
The set of aggregate structures generated for the systems corresponding to charge $Q = 15$ was used to examine the growth dynamics and behavior. These systems were chosen as they express all four identified phases. Figure 4.6 show the growth of the average number of

particles in each cluster $< N >$. The number of unaggregated monomers is then calculated as $< N_m >= N - < N >$.

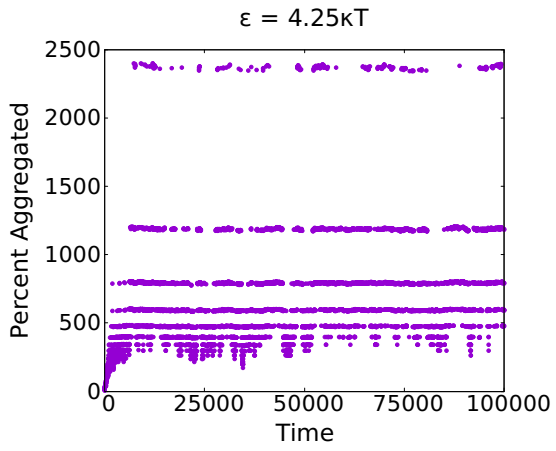
In figure 4.6 it can be seen that the stable crystalline and gel states iterate between several energy minima (represented by horizontal lines), before reaching a final equilibrium. The liquid state rapidly fluctuates as many small clusters form and fall apart in the dense solution (9.5% volume concentration). The amorphous phase shares behavior of both of these phases. Rapid fluctuations as in the liquid phases is seen, however the fluctuations transition between metastable states.



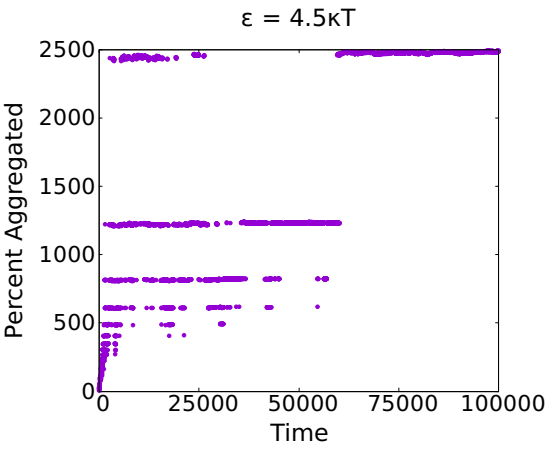
(a) Liquid: $\epsilon = 3.75$



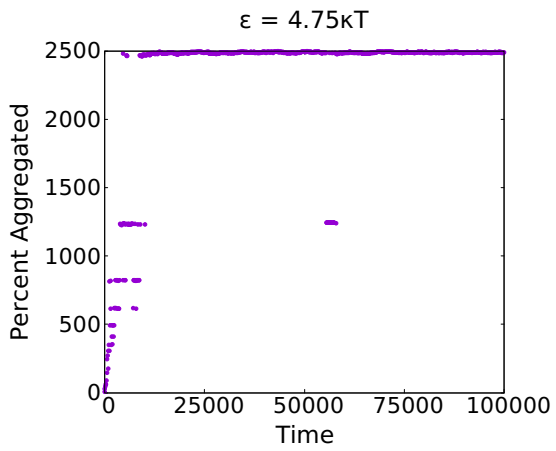
(b) Amorphous: $\epsilon = 4.0$



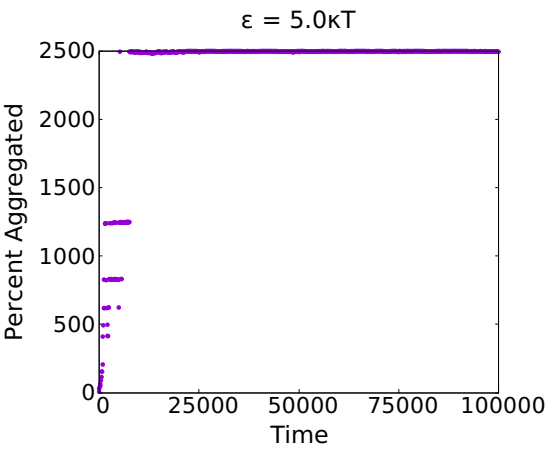
(c) Amorphous: $\epsilon = 4.25$



(d) Crystal: $\epsilon = 4.5$



(e) Crystal: $\epsilon = 4.75$



(f) Gel: $\epsilon = 5.0$

Figure 4.6: The dynamics of the average number of particles in each cluster in time for charge $Q = 15$.

Chapter 5

Conclusion

5.1 Summary

The observed behavior of the electrostatics is consistent with intuition gained from the idealized model of chapter 3, while exposing new structure and dynamics that this simple model cannot describe. The observed metrics also provide evidence for the unique and characteristically different behavior of non-local electrostatics in comparison to Debye-Hückel. Specifically the electrostatics have predicted a finite region in $Q - \epsilon$ space in which crystallization can occur.

The divergence in behavior of the electrostatics examined in this paper, to those of Debye-Hückel are repercussions of the non-pairwise interactions of the non-local electrostatics. The increasing cost to aggregation makes dense structures entropically unfavorable; creating a finite window for crystal precipitation, and a preference toward open gel-like structures.

The electrostatics also introduces an interesting non-equilibrium amorphous phase. Aggregates of this phase possess no defined structure and embody the underlaying concept of competition between attractive energies and repulsive entropies. In these amorphous structures, attractive potentials draw particles closer together, reducing the volume accessible to salt counter-ions, until the entropic cost of reducing this volume causes the aggregate to fall

apart.

The analysis of structure, phase, and dynamics of the electrostatics, along with the fundamental work laid out in Schmit, Whitelam, Dill 2011[11], indicates that when forming dense structures the entropic effects of the screening layer cannot be ignored; and that colloid stability is driven by the competing nature of energies and entropies.

5.2 Future Work

This thesis has provided a brief inspection of the presented non-local electrostatic interaction and its effect on protein crystallization. Improving the computational algorithms to handle complex overlaps of screening regions would allow for these electrostatics to be examined at lower salt concentrations. The simulated proteins were spherically symmetric, which could be removed by adding rotational binding sites, such that it is energetically favorable for particles to bind in certain rotational orientations with respect to each other.

Bibliography

[1] *Everything you want to know about coagulation and Flocculation*. Zeta-Meter, 4 edition, 1993.

[2] Harold P Erickson. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol Proced Online*, 2009.

[3] et. al. Antonina Andreeva, Dave Howorth. Scop2 prototype: a new approach to protein structure mining. *Nucleic Acids Res.*, 42, 2013.

[4] Jacob White Lian Hing Tan, Kian Meng Lim. Numerical study of the poisson-boltzmann equation of biomolecular electrostatics. 2006. URL http://hdl.handle.net/1721.1/30372.

[5] A Brigo F. Fogolari and H. Molinari. The poisson-boltzmann equation for biomolecular electrostatics: a tool for structural biology. *Journal of Molecular Recognition*, 15, 2002.

[6] John Arthur Wilson. Theory of colloids. *J. Am. Chem. Soc*, 38, 1916.

[7] F. Hofmeister. *Arc. Exp. Pathol. Pharmacol*, 24, 1888.

[8] Jifeng Zhang. *Protein-Protein Interactions in Salt Solutions, Protein-Protein Interactions -Computational and Experimental Tools*. InTech, 2012. ISBN 978-953-51-0397-4.

[9] E. J. W. Verwey. Theory of the stability of lyophobic colloids. *J. Phys. Chem*, 51, 1947.

[10] *General Principles of Colloid Stability and the Role of Surface Forces*. Colloids and Interface Science Series. WILEY-VCH, 2007. ISBN 978-3-527-31462-1.

[11] Ken Dill Jeremy D. Schmit, Stephen Whitelam. Electrostatics and aggregation: How charge can turn a crystal into a gel. *The Journal of Chemical Physics*, 135, 2011.

[12] Alexander McPherson and Jose A. Gavira. Introduction to protein crystallization. *Acta Crystallographica Section F*, 70, 2014. ISSN 2053-230X.

[13] E. Zaccarelli F. Sciortino, P. Tartaglia. One-dimensional cluster growth and branching gels in colloidal systems with short-range depletion attraction and screened electrostatic repulsion. *The Journal of Chemical Physics*, 109, 2005.

[14] BO LI. Minimization of electrostatic free energy and the poissonboltzmann equation for molecular solvation with implicit solvent. *SIAM J. Math Anal.*, 40, 2009.

[15] M. J. Holst. *The Poisson-Boltzmann Equation: Analysis and Multilevel Numerical Solution*. PhD thesis, University of California San Diego, 1994.

[16] Lyderic Bocquet Yan Levin, Emmanuel Trizac. On the fluidfluid phase separation in charged-stabilized colloidal suspensions. *Journal of Physics: Condensed Matter*, 15, 2003.

[17] Jeremy D. Schmit Shradha Mishra. Electrostatic interactions in concentrated protein solutions. 2013.

[18] Remco Tuinier Henk N. W. Lekkerkerker. *Colloids and the Depletion Interaction*. Springer Science+Business Media B.V, 2011.

[19] Albert S. Kim Jim C. Chen. Brownian dynamics, molecular dynamics, and monte carlo modeling of colloidal systems. *Advances in Colloid And Interface Science*, 112, 2004.

[20] W. F. van GUNSTEREN and H. J. C. BERENDSEN. Algorithms for brownian dynamics. *Molecular Physics*, 45:637–647, 1982.

[21] Stefan Auer and Daan Frenkel. Numerical simulation of crystal nucleation in colloids. *Adv. Polymer Science*, 2005.

[22] David R. Nelson Paul J. Steinhardt and Marco Ronchetti. Bond-orientational order in liquids and glasses. *Phys. Rev. B*, 28, 1983.

[23] Wolfgang Lechner and Christoph Dellago. Accurate determination of crystal structures based on averaged local bond order parameters. *The Journal of Chemical Physics*, 129, 2008.

Appendix A

Computational Units

The computational units are based on the protein diameter, σ, the protein mass, m, and K_BT all being equal to one.

$$\begin{aligned} K_BT &= 1 \\ m &= 1 \\ \sigma &= 1 \end{aligned} \tag{A.1}$$

with this the unit of time for the system is then given by

$$t = \sigma \left(\frac{m}{K_BT} \right)^{\frac{1}{2}} \tag{A.2}$$

Using the following physical values for the system of interest, in which η_h is the viscosity of water, and c the salt concentration of the solution.

$$\begin{aligned} \eta_h &= 0.001 \frac{\text{kg}}{\text{m} \cdot \text{s}} \\ m &= 5 \cdot 10^{-23} \text{kg} \\ \sigma &= 6 \cdot 10^{-9} \text{m} \\ T &= 300 \text{K} \\ c &= 1 \text{M} \end{aligned} \tag{A.3}$$

The viscosity Γ defined in equation 3.1 is then solved via the Stoke-Einstein equation.

$$\Gamma = 3\pi\eta\sigma \tag{A.4}$$

Arising to the computation value of

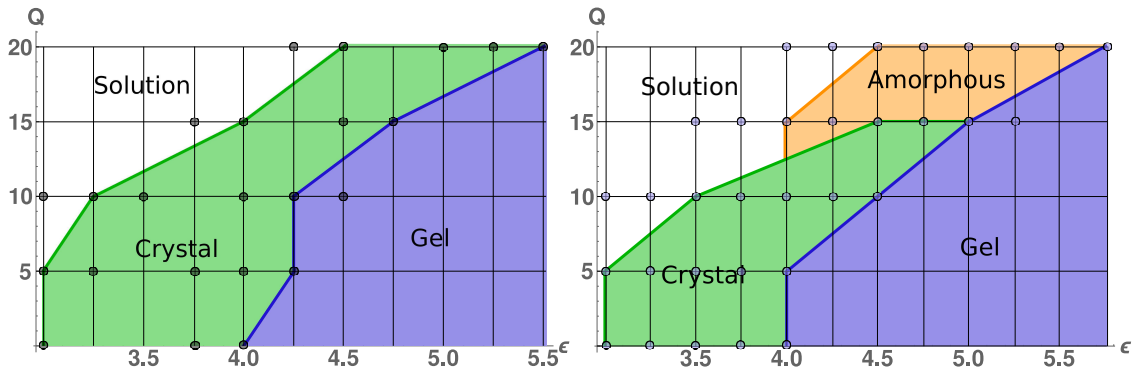
$$\Gamma = 735 \tag{A.5}$$

Likewise the computational unit of the salt concentration defined in equation 2.13 is found be to

$$c = 130 \tag{A.6}$$

Appendix B

Extended Figures And Plots



(a) Debye-Hückel (b) Non-linear

Figure B.1: Phase diagram for the non-linear and Debye-Hückel interactions. Sampling of data point occurs at $5Q$ and 0.25ϵ increments. Sampled data marked as circles.

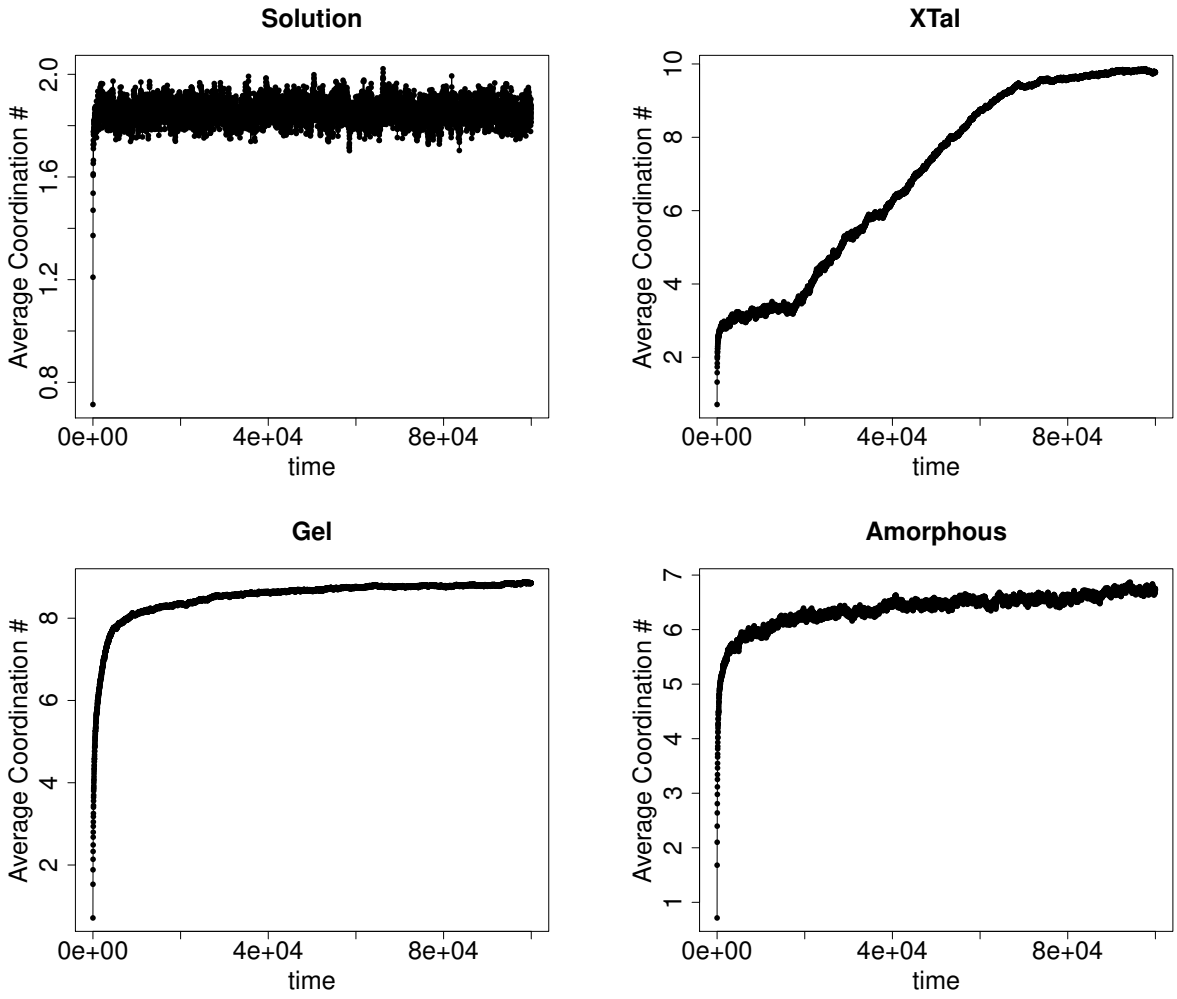


Figure B.2: The time evolution of the average coordination number of the four identified phases associated with non-local electrostatics