

The Phase Diagram of Globular Protein Solutions: The Role of the Range of Interaction

by

Neer Ruben Asherie

B.A., Natural Sciences, University of Cambridge, 1991

M.A., Natural Sciences, University of Cambridge, 1995

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the


MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1998

© Massachusetts Institute of Technology 1998. All rights reserved.

Author

Department of Physics
October 31, 1997

Certified by 

George B. Benedek
Alfred H. Caspary Professor of Physics and Biological Physics
Thesis Supervisor

Accepted by

George F. Koster
Chairman, Departmental Committee

FEB 10 1998

Science

The Phase Diagram of Globular Protein Solutions: The Role of the Range of Interaction

by

Neer Ruben Asherie

Submitted to the Department of Physics
on October 31, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The phase diagram of globular proteins is studied both theoretically and experimentally. Emphasis is placed on understanding how the microscopic interactions of the proteins lead to the phase diagrams observed.

The theoretical part of this work uses a combined analytic and computational approach. The proteins are assumed to be hard spheres of diameter σ . The interactions between the proteins are modeled by a square-well potential with range λ and depth ϵ . This model is used to show how the relative positions of the liquid-liquid and liquid-solid phase boundaries in globular protein solutions are related to the short-range nature of the protein interactions. The theory presented successfully describes the features of the phase diagrams observed in a wide variety of protein and colloidal systems. The theoretical study is applied to the phase diagram of aqueous solutions of γ -crystallins to gain insight into the microscopic interactions between these proteins. Analysis of the experimental data for the critical volume fractions permits the determination of the range of interaction appropriate for these proteins. A comparison of the experimental and calculated widths of the liquid-liquid coexistence curves suggests a significant contribution from anisotropy in the real interaction potential of the γ -crystallins.

The experimental part of this work focuses on the role of small aggregates in shaping the phase diagram. Oligomers of γ -crystallins are produced by crosslinking native γ -crystallin proteins. Experimental results for the liquid-liquid coexistence curves of the crosslinked dimers and trimers are presented. These results are analyzed within the context of the model used to describe protein monomers. It is found that the protein oligomers may be described as having longer effective ranges of interaction than the monomer protein. The experimental findings are used to illustrate the important connection between aggregation and phase separation in globular protein solutions.

Thesis Supervisor: George B. Benedek

Title: Alfred H. Caspary Professor of Physics and Biological Physics

Acknowledgements

First and foremost I wish to thank my advisor, George Benedek. His constant support allowed me to carry out the work presented here. He rekindled my love of physics when I thought it gone and for that I will be forever grateful.

I thank Profs. Mehran Kardar and Min Chen for being on my committee.

I have had the privilege to work with a stimulating set of people during my thesis research. I am deeply grateful to Aleksey Lomakin and Jayanti Pande for their help with my work. Their contradictory and bewildering advice has been a tightrope I have enjoyed walking. I am in debt to Olutayo Ogun for his patience in supplying me with protein. I appreciate the willingness of George Thurston, Yevgeniya Zastavker, and Yoko Kusumoto to listen to me when they had better things to do. I thank Ajay Pande for poolside chats. Finally, I wish to thank Canwen Liu, my teacher during my early days in the lab, for “monthès and monthès of inexhaustable enerch.”

I wish to thank my family for their support over the years. I am especially grateful to my brother Oded, with whom I lived as I completed this thesis. Whenever I was despondent about my research he was always there to make me do the shopping, wash the dishes and clean the house. I also want to thank my friend Ludwig Chincarini, who through phonecalls, letters, postcards, faxes and emails was constantly reminding me to “slide that thesis on to the stacks.”

To my parents
who taught me the value
of a good education

Always the beautiful answer who asks a more beautiful question

E.E. CUMMINGS

Contents

1	Introduction	21
1.1	The Phase Transitions of Globular Protein Solutions	21
1.2	Overview of the Thesis	26
2	The Liquid Phase	31
2.1	Introduction	31
2.2	Computer Simulation	35
2.3	Results and Discussion	43
2.3.1	Results of this Study	43
2.3.2	Other Monte Carlo Results	50
2.3.3	Connection with the Mean Field and Adhesive Sphere Models	54
2.3.4	Comparison with Experimental Data for the γ -crystallin Proteins	60
2.4	Summary and Conclusions	63
3	The Solid Phase	67
3.1	Introduction	67
3.2	The Cell Model	71
3.3	Results and Discussion	72
3.3.1	Construction of the Phase Diagram	72
3.3.2	Comparison with Experimental Results	77
3.4	Summary	80
4	Aggregation and Phase Separation	81
4.1	Introduction	81

4.2	Materials and Methods	83
4.2.1	Preparation of Pure γ_{IIB} -Crystallin Solutions	83
4.2.2	Production of Crosslinked Oligomers	84
4.2.3	Characterization of Crosslinked Oligomers	86
4.2.4	T_{ph} measurements	90
4.3	Results and Discussion	90
4.3.1	Coexistence Curves of Oligomers	90
4.3.2	Comparison with Theoretical Models	92
4.3.3	Effect of Oligomers on the Phase Behavior of Globular Proteins	101
4.4	Conclusions	106
5	Conclusions	107
A	The Incompressible Protein-Water Solution	111
B	Widom's Formula	115
C	Chemical Potential Extrapolation	119
D	Flowchart of the Monte Carlo Algorithm	121
E	The Number of Contacts	127
F	The Chemical Potential of the Solid	129

List of Figures

- 1-1 The phase diagram of γ_{II} -crystallin [2, 3, 4] . The circles are points on the liquid-liquid coexistence curve (CC). The squares are points on the liquidus line (L). The triangle is a point on the solidus line (S). The lines are guides to the eye. The critical temperature is $T_c = 278.4\text{K}$. The critical volume fraction is $\phi_c = 0.21$ 23
- 1-2 The phase diagram of argon [8, 9, 10] showing the coexistence curve (CC) and the liquidus (L) and solidus (S) lines. The critical temperature is $T_c = 150.86\text{K}$. The critical volume fraction is $\phi_c = 0.133$, assuming a hard core diameter $\sigma = 3.162\text{\AA}$ [12] . The triple point temperature is $T_t = 83.78\text{K} = 0.56T_c$ 24
- 2-1 Illustration of the temperature expansion method for $\lambda = 1.25$. The open symbols represent the Monte Carlo results for the chemical potential for three different values of the reduced energy $\hat{\epsilon}=1.318$ (triangles), 1.267 (circles), and 1.216 (squares). The solid line is a fit to the $\hat{\epsilon}=1.267$ Monte Carlo results using Eq. (2.17) with $n_0 = 4$. The dashed lines are the chemical potentials obtained by extrapolating the $\hat{\epsilon}=1.267$ chemical potential to $\hat{\epsilon}=1.318$ (coarse dashed line) and to $\hat{\epsilon}=1.216$ (fine dashed line). 44

2-2	Reconstruction of the spinodal and the coexistence curve. The open circles are the Monte Carlo results for the chemical potential with $\lambda = 1.25$ and $\hat{\epsilon} = 1.267$. The isotherms which result from the temperature extrapolation (from $\hat{\epsilon} = 1.267$ to $\hat{\epsilon} = 1.317$ in steps of 0.005) are shown as solid lines. The coexisting points and spinodal points at each temperature are shown as dashes and crosses respectively.	45
2-3	Coexistence curves. The curves with progressively larger widths represent the results obtained for the reduced ranges $\lambda=1.8, 1.5, 1.25$ and 1.1	46
2-4	Comparison of the coexistence curves. The coexistence curves from our simulations at $\lambda=1.25$ (coarse dashed line) and 1.5 (fine dashed line) are shown together with the coexisting points obtained by Vega <i>et al.</i> [47] for the same ranges (open circles and squares respectively). Note that $T^* = kT/\epsilon$	51
2-5	Variation of the critical volume fraction with the reduced range. Our results (solid circles) are presented together with those of Henderson <i>et al.</i> [45] (open squares), Vega <i>et al.</i> [47] (open triangles), and Lomba <i>et al.</i> [49] (open circles). The solid line is a linear extrapolation of our results to $\lambda = 1$. The dashed line is the mean field result.	52
2-6	Variation of the critical reduced energy with the reduced range λ . Our results (solid circles) are presented together with those of Henderson <i>et al.</i> [45] (open squares), Vega <i>et al.</i> [47] (open triangles), and Lomba <i>et al.</i> [49] (open circles). The solid line is Eq. (2.31) with $\tau_c = 0.13$. The dashed line is Eq. (2.30) with $a_c = 10.6$	53

- 2-7 The average number of contacts per particle. The average number of contacts per particle, $\bar{\eta}$, is shown as a function of ϕ for several ranges: (i) $\lambda=1.05$, $\hat{\epsilon}=2.650$ (triangles); (ii) $\lambda=1.25$, $\hat{\epsilon}=1.267$ (squares); (iii) $\lambda=1.65$, $\hat{\epsilon}=0.605$ (bow ties); (iv) $\lambda=2.20$, $\hat{\epsilon}=0.260$ (circles). The straight dashed lines represent the low ϕ behavior for the different ranges as given by Eq. (2.32). The solid line is the adhesive sphere result for $\tau = 0.13$ 59
- 2-8 Comparison with the experimental results for the γ -crystallins. The coexistence curve generated by the Monte Carlo simulation for $\lambda=1.25$ is shown as a coarse dashed line. The fine dashed line represents the coexistence curve obtained analytically in the mean field limit. The experimental results of Broide *et al.* [2] are presented for γ_{IIIa} (circles), γ_{IIIb} (squares), γ_{II} (triangles), and γ_{IVa} (bow ties). The solid line is the coexistence curve obtained for $\lambda=1.25$ and a temperature dependent interaction energy of the form $\epsilon = kT_c \hat{\epsilon}_c (1 + \frac{T-T_c}{T_c} \kappa)$, with $\kappa = -3$. . . 61
- 3-1 The phase diagram of γ_{II} -crystallin [2, 3, 4]. The circles are points on the liquid-liquid coexistence curve (CC). The squares are points on the liquidus line (L). The triangle is a point on the solidus line (S). The lines are guides to the eye. The critical temperature is $T_c = 278.4\text{K}$. The critical volume fraction is $\phi_c = 0.21$ 69
- 3-2 The phase diagram of argon [8, 9, 10] showing the coexistence curve (CC) and the liquidus (L) and solidus (S) lines. The critical temperature is $T_c = 150.86\text{K}$. The critical volume fraction is $\phi_c = 0.133$, assuming a hard core diameter $\sigma = 3.162\text{\AA}$ [12]. The triple point temperature is $T_t = 83.78\text{K} = 0.56T_c$ 70
- 3-3 The phase diagram for a square-well system for $\lambda = 1.25$. The liquidus lines for three values of n_s are shown: $n_s = 12.0$ (A); $n_s = 11.6$ (B); $n_s = 11.5$ (C). The vertical line (D) is the solidus for $n_s = 12.0$. Curve E is the coexistence curve taken from Fig. 2-3 74

3-4	The experimentally determined phase diagram for the γ -crystallins from Refs. [2, 3] . We present the liquid-liquid coexistence curve and the liquidus line for γ_{IIIa} (circles), γ_{IIIb} (squares), γ_{II} (triangles) and γ_{IV} (bow ties). The solid lines are guides to the eye.	78
4-1	Size exclusion chromatography results (on a Superdex 200HR column) for the crosslinked protein after a three hour reaction. The absorbance at 280 nm is shown as function of retention time in minutes.	85
4-2	The expected reaction of bismaleimido-hexane (BMH) with thiol groups on a protein, schematically shown as P. The final product is the expected structure of the dimer.	87
4-3	SDS/PAGE of the oligomers of γ_{IIIb} without DTT. Lanes (left to right): 1 and 7, molecular mass markers (from bottom to top) are 14.4, 21.5, 31.0, 45.0, 66.2 and 97.4 kDa; 2, native γ_{IIIb} ; 3, control γ_{IIIb} ; 4, monomer fraction after crosslinking; 5 and 6, dimer fraction; 8 and 9, trimer fraction; 10, n-mer mixture.	89
4-4	Coexistence curves for the oligomers of γ_{IIIb} -crystallin in 0.1 M sodium phosphate (pH 7.1). Native γ_{IIIb} monomers, from Ref. [2] (open squares); control γ_{IIIb} monomers, this report (open circles); γ_{IIIb} monomers after crosslinking (solid bowties); γ_{IIIb} dimers (solid circles); γ_{IIIb} trimers (solid triangles). The solid lines are fits to Eq. (4.1)	91
4-5	The width of the coexistence curve A and the critical volume fraction ϕ_c for various systems. The center of each word lies approximately at the values of A and ϕ_c corresponding to that system.	93
4-6	Coexistence curves for (1) monomers and (2) dimers according to various models. A: mean field (solid lines); B: square-well potential with $\lambda=1.5$ (coarse dashed lines); C: adhesive spheres (fine dashed lines).	99

4-7	Comparison of the coexistence curves for γ_{IIIb} and γ_{II} oligomers. Native γ_{IIIb} and γ_{II} monomers, from Ref. [2] (open squares and open triangles respectively); γ_{IIIb} dimers (solid circles); γ_{IIIb} trimers (solid triangles); γ_{IIIb} n-mers (solid squares); γ_{III} from Ref. [23] (crossed squares). . . .	102
4-8	QLS data for γ_{IIIb} - and γ_{II} -crystallins. Normalized scattering intensity is plotted against apparent R_h . (A) γ_{IIIb} monomer. (B) γ_{IIIb} dimer. (C) γ_{IIIb} trimer. (D) γ_{IIIb} n-mer mixture. (E) γ_{II} monomer. (F) γ_{III} .	103
D-1	A flowchart for the main routine of the Monte Carlo algorithm. . . .	122
D-2	A flowchart for the calculation of the chemical potential.	124
F-1	(A) The solid for the case where $\alpha \approx 1$ and $\lambda > 1$. The large solid circles represent the hard cores of the particles. The small solid circle encloses the region where the center of the central particle may move without overlapping the hard core of a neighbor (“sphere of excursion”). The dashed circle encloses the region where the central particle could move and still be in contact with all of its neighbors (“sphere of attraction”); (B) The sphere of excursion and sphere of attraction shown separately. The hatched region is excluded to the particle for it cannot move beyond the sphere of excursion.	132
F-2	(A) The solid for the case where $\alpha - 1 > \lambda - \alpha$. i.e., $\alpha > (\lambda + 1)/2$. The large solid circles represent the hard cores of the particles. The small solid circle represents the sphere of excursion. The dashed circle represents the sphere of attraction; (B) The sphere of excursion and sphere of attraction shown separately. In the gray region the particle does not feel the attraction of all of its neighbors.	134
F-3	A schematic illustration of the partition function Z as a function of cell size α in the $\hat{\epsilon} \rightarrow \infty$ limit.	136

List of Tables

2.1	Results and parameters from the Monte Carlo simulations at different reduced ranges λ ($\lambda = \infty$ is the mean field limit). The quantities presented are: (i) the critical volume fraction ϕ_c ; (ii) the reduced critical energy $\hat{\epsilon}_c$; (iii) the average number of contacts per particle at the critical point \bar{n}_c ; (iv) the number of successful attempts made K_{tot} in units of 10^6 ; (v) the reduced energy at which the simulation is performed $\hat{\epsilon}_1$; (vi) the diameter of the particles σ ; (vii) the maximum number of particles used in the simulation M_{max} ; and (viii) the maximum volume fraction ϕ_{max}	47
2.2	Variation of the simulation results with the order of the chemical potential fit. n_0	49
3.1	Metastability gap parameters at different reduced ranges λ (See Eq. (3.7)). The quantities presented are: (i) the average number of contacts per particle in the solid n_s^* at the metastability boundary ($T_L = T_c$); (ii) the change in the total number of contacts per particle in the liquid n_l upon the the addition of an extra particle; (iii) the critical volume fraction ϕ_c ; (iv) the reduced critical energy $\hat{\epsilon}_c$	76
4.1	Coexistence curve parameters (see Eq. (4.1) in text) for the monomer and dimer.	92
4.2	Coexistence curve parameters for the curves presented in Fig. 4-6 . . .	100

4.3	Hydrodynamic radii in the limit of zero concentration for oligomers made up of spheres [117] . $R_h(m)$ is the hydrodynamic radius of the spherical monomer.	105
A.1	Corresponding variables for the incompressible protein-water solution (at zero external pressure) and the compressible one component fluid systems.	113

Chapter 1

Introduction

1.1 The Phase Transitions of Globular Protein Solutions

Globular proteins are folded polypeptide chains which have molecular weights ranging from approximately six thousand to one million daltons and radii of twenty-five to two hundred angstroms [1]. Globular proteins are given their name because they all have compact structures in contrast to the filamentous structures of fibrous proteins. Solutions of globular proteins undergo several types of changes of state. In this thesis I will examine those transitions of globular proteins in which the physics, rather than the biochemistry, of the globular proteins plays the dominant role, since such transitions can provide very general information about the interactions of these proteins. The phase transitions of globular proteins which are mostly governed by physical interactions are crystallization and liquid-liquid phase separation.

The more common phase transition that is observed in these protein solutions is crystallization. Upon a change of conditions (such as temperature, protein concentration or pH) the proteins condense from solution and form a crystal. In liquid-liquid phase separation, a change of conditions causes an initially homogenous protein solution to form two separate coexisting liquid phases, one protein-rich, the other protein-poor. Both of these transitions are general in nature and have counterparts in simple

molecular fluids. The crystallization of proteins is equivalent to the fluid-solid transition while liquid-liquid phase separation corresponds to gas-liquid coexistence. There is however one striking difference: when most simple fluids are cooled the order of phases observed is gas to liquid to solid. In protein solutions the protein “gas” usually transforms into a solid without passing through a liquid phase.

I illustrate this unusual order of phases by presenting in Fig. 1-1 the phase diagram of γ_{II} -crystallin protein [2, 3, 4]. This phase diagram is typical of the γ -crystallins (a family of monomeric eye lens proteins) and of other small globular proteins, such as lysozyme [5, 6, 7]. The circles are points which represent the volume fractions (ϕ) of coexisting protein-rich and protein-poor liquid phases (liquid-liquid coexistence curve). The squares (liquidus line) and the triangle (solidus line) represent respectively the volume fractions of protein in the liquid and solid phases in equilibrium with each other. We see that there is no triple point and the coexistence curve lies below the liquidus line. It is important to note that liquid-liquid phase separation may be observed in many globular protein solutions despite it being metastable with respect to solidification. For comparison the phase diagram of argon is shown in Fig. 1-2 [8, 9, 10]. We see that for argon the critical point lies above the triple point (i.e., $T_c > T_t$) indicating the presence of a stable liquid phase. In addition the coexistence curve is narrower and the value of the critical volume fraction is smaller than for globular protein solutions.

There is one additional difference between the phase behavior of globular protein solutions and that of simple fluids which is not apparent from a comparison of the phase diagrams. In globular protein solutions both liquid-liquid phase separation and crystallization may be obscured or accompanied by protein aggregation, in which the protein forms irregular solid structures. Simple fluids on the other hand do not aggregate. It is well known that many simple fluids form clusters [11]. These however are either alterations in the local order of the liquid phase or defects in the arrangement of the solid. They are not aggregates, which exist as independent structures and which form irreversibly. Although aggregation is governed by the same molecular interactions that lead to liquid-liquid phase separation and crystallization,

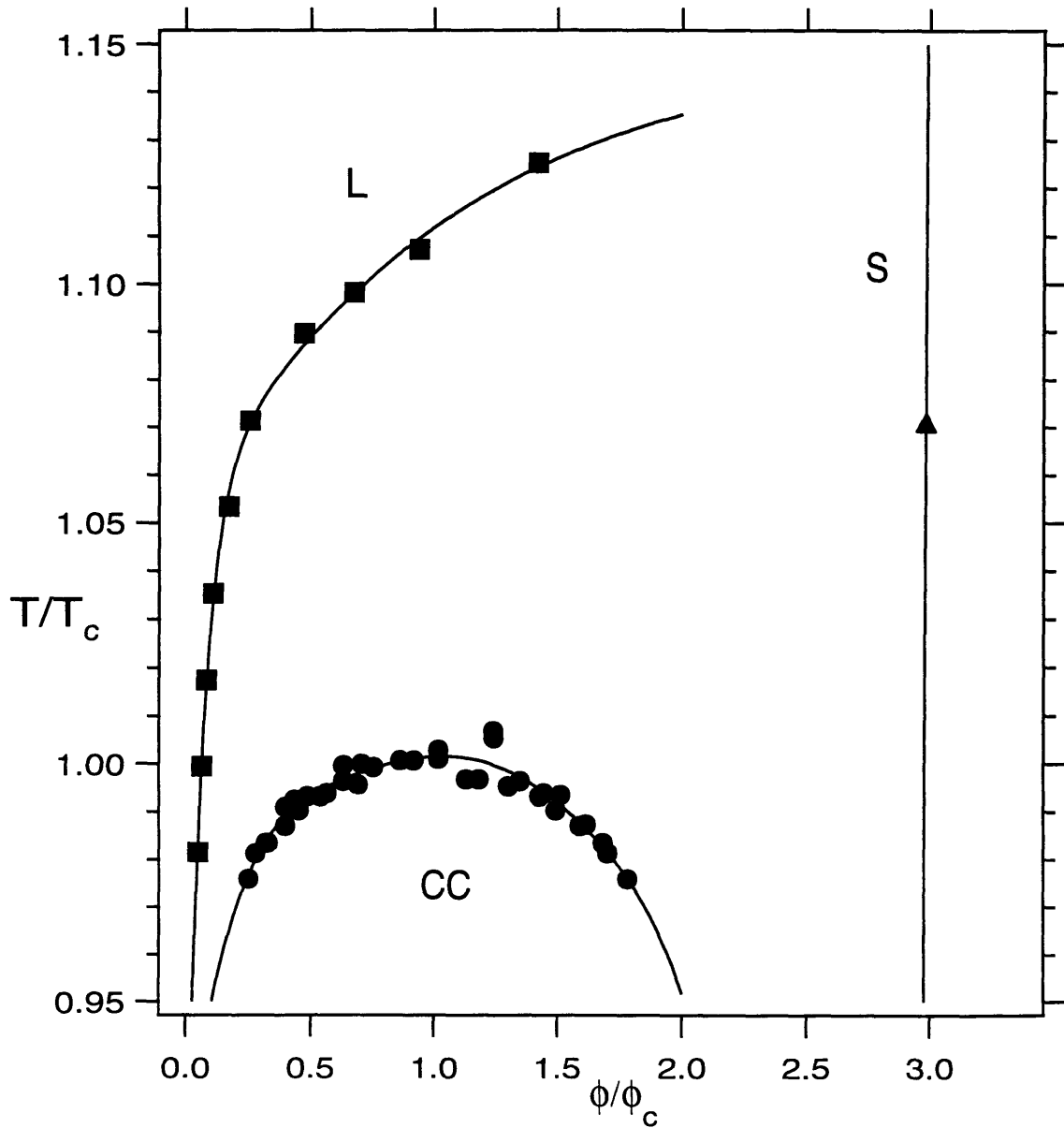


Figure 1-1: The phase diagram of γ_{II} -crystallin [2, 3, 4]. The circles are points on the liquid-liquid coexistence curve (CC). The squares are points on the liquidus line (L). The triangle is a point on the solidus line (S). The lines are guides to the eye. The critical temperature is $T_c = 278.4\text{K}$. The critical volume fraction is $\phi_c = 0.21$.

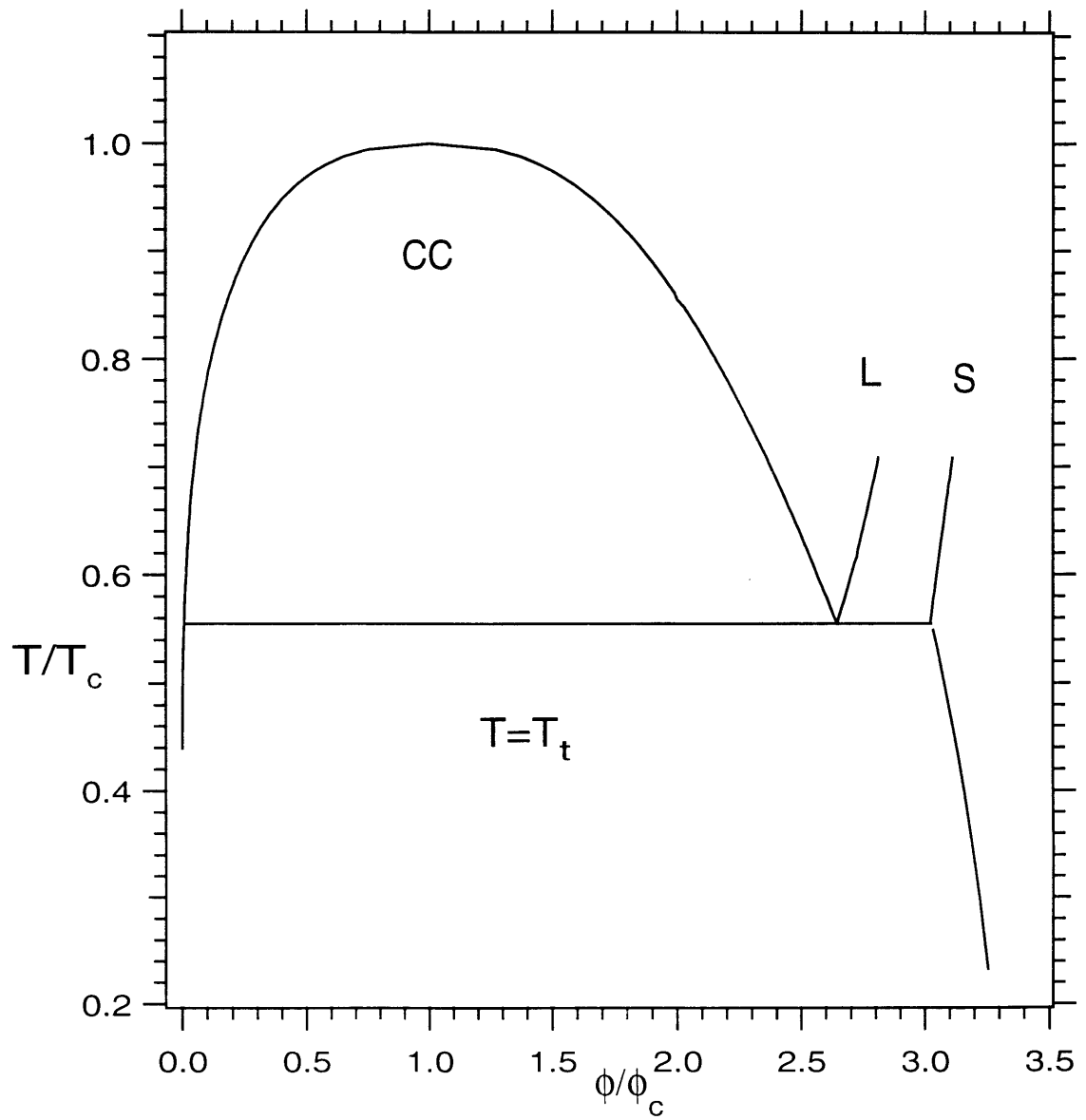


Figure 1-2: The phase diagram of argon [8, 9, 10] showing the coexistence curve (CC) and the liquidus (L) and solidus (S) lines. The critical temperature is $T_c = 150.86\text{K}$. The critical volume fraction is $\phi_c = 0.133$, assuming a hard core diameter $\sigma = 3.162\text{\AA}$ [12]. The triple point temperature is $T_t = 83.78\text{K} = 0.56T_c$.

it is not strictly a phase transition. Aggregates are not in thermodynamic equilibrium with another macroscopic phase; there is no aggregation phase boundary.

In this thesis I will address the following questions:

1. *Why is the shape of the coexistence curve for protein solutions different from that of simple fluids?* Specifically, I will examine why the critical volume fraction and the width of the coexistence curve are both larger for globular protein solutions than for simple fluids.
2. *Why is liquid-liquid coexistence metastable with respect to solidification in protein solutions?* I will also discuss why liquid-liquid phase separation can be observed despite it being metastable.
3. *What is the effect of aggregation on the phase diagram of globular protein solutions?* I will discuss how aggregation affects both liquid-liquid phase separation and solidification in these solutions.

I will argue that all of these questions can be tackled with a simple description of the microscopic interactions of the proteins. Such a description can help us understand the many biophysical phenomena in which phase separation processes are involved. I list a few examples of such phenomena below:

- The crystallization of proteins [13]: In order to carry out the x-ray analysis of protein structure high-quality protein crystals are needed. Often it is hard to find adequate conditions for crystal formation. It is therefore essential to understand how the microscopic protein interactions determine the location of the phase boundaries for the crystallization transition.
- Protein condensation diseases [14]: The phase transitions of proteins have been implicated in several diseases e.g., cataract, sickle cell disease and cryoimmunoglobulinemia. The understanding of the location of the phase boundaries and the strategies to shift them by modifying protein interactions are key elements in the search for the treatment of such diseases.

- The industrial purification of proteins [15]: a prominent method for purifying and concentrating protein solutions is “two-phase partitioning” or “membrane-less osmosis”. Here proteins are mixed with polymers, polysaccharides or other proteins. Upon phase separation, two liquid phases form which differ greatly in the concentrations of the macromolecular components. By an appropriate choice of phase-forming polymer, a high yield of the target protein in the concentrated protein-rich phase can be obtained. This method has been applied extensively to food colloids.
- Microcompartmentation of the cell cytoplasm [16]: It has been suggested that the macromolecular diversity and concentration in the fluid phase of the cell cytoplasm constitute conditions necessary and sufficient for aqueous phase separation. This phase separation would then be responsible for the partitioning or “microcompartmentation” of materials (e.g., ions, mitochondria and proteins) in the cytoplasm. It is still an unsolved question how much of the microcompartmentation of the cell cytoplasm is due to phase separation and how much depends on biochemical factors.

In the next section I will outline the work I present in this doctoral thesis: using the methods of physics, I explain how the microscopic interactions of globular proteins govern their phase transitions.

1.2 Overview of the Thesis

This thesis investigates how the microscopic interactions shape the phase diagram of globular proteins. The simplest picture of a globular protein is of a hard sphere of matter interacting with other hard spheres of matter. Though crude, this picture captures an essential structural feature of globular proteins, for the hard sphere model accounts for the repulsive interactions between proteins. However, to observe both liquid-liquid phase separation and crystallization in a systems of spheres, these spheres should also have attractive interactions. This way, at a sufficiently low temperature,

the energy of attraction will overcome the entropy of the spheres leading to a phase transition.

We know that globular proteins in solution have generally short ranges of interaction, because the proteins, which are at least fifty angstroms in diameter, can only interact with each other through a few layers of water, a distance of approximately ten angstroms [17]. In this respect globular proteins resemble colloids which also have short ranges of interaction. In the short-range regime the thermodynamic properties become universal, independent of the shape of the interaction potential. Therefore, we can use a simple model for the protein interactions to describe the more complex interactions real proteins exhibit. With such a simple model I will answer the questions I stated in the previous section.

Why is the shape of the coexistence curve for protein solutions different from that of simple fluids? In Chapter 2, parts of which have already been published [18], we study the liquid-liquid phase separation of globular proteins by performing Monte Carlo simulations of spheres with isotropic attractive interactions. Since we understood that the interactions of the proteins in solution are short-ranged, we wanted to simulate short-range systems. Traditional Monte Carlo methods, however, become computationally prohibitive when applied to short-range interactions. Dr. Aleksey Lomakin has developed an innovative Monte Carlo scheme, which combines simulations with analytic techniques, to thoroughly explore short-range interactions between the proteins. Using this scheme we calculate, for each range of interaction, the chemical potential of a model protein solution as a function of protein volume fraction and temperature. From this we obtain the liquid-liquid coexistence curves by a method analogous to the Maxwell equal area construction. The coexistence curves so obtained allow us to make meaningful comparisons with the experimentally determined coexistence curves for solutions of globular proteins. Such a comparison was not possible with the results previously available in the literature. Our findings provide insight into the central role played by the range of interaction in determining the shape and the location of the phase boundaries. As expected, we find that there is much better agreement between our simulation results and the experimental results when we

describe the proteins as having short-range interactions.

Our Monte Carlo study helped us understand the importance of the range of interaction in the liquid-liquid phase separation of protein solutions. Since our model gave a reasonable representation for the chemical potential of the liquid, we were able to extend our microscopic model to address the other phase transition exhibited by globular proteins, namely crystallization.

Why is liquid-liquid coexistence metastable with respect to solidification in protein solutions? In Chapter 3, parts of which have also been published [19], we use a simple analytic model of a colloidal solid to study the crystallization transition. Specifically, we use the Lennard-Jones-Devonshire cell model to obtain an analytic expression for the chemical potential of a solid with short-range interactions. By using this chemical potential together with the chemical potential of the liquid phase we had obtained previously, we are able to describe the principal features of the phase diagram in a wide variety of colloidal systems, including globular proteins. In particular, we explain how the “metastability gap” of colloidal solutions (i.e., the difference between the temperature at which a stable solid phase appears and that at which coexisting liquid phases are first stable) is related to the range of interaction and to the number of contacts made by particles in the solid phase. We find good qualitative agreement between our results and the experimental observations for several colloidal solutions, not only for globular proteins but also systems such as colloid-polymer mixtures.

The two works discussed above use a very simplified model of protein interactions. Nevertheless, by identifying the most important elements required to describe the phase transitions of globular proteins, we have been able to relate the phase behavior of these proteins (both liquid-liquid phase separation and crystallization) to their microscopic interactions. In fact, the model used to characterize globular proteins (a repulsive hard core with a short-range attractive interaction) is equally applicable to various colloidal systems, such as silica spheres coated with stearyl alcohol [20] or mixtures of polystyrene spheres and polymers [21].

Although they are both short-ranged, there is an important difference between solutions of globular proteins and most colloidal systems. The interactions for simple

colloids are expected to be isotropic and therefore should be well described by the model presented in Chapters 2 and 3. In contrast, the surface of globular proteins is not uniform, since the proteins are made up of many amino acids residues, each of which interacts differently with the surrounding solution [22]. In particular, this anisotropy in the interaction between proteins leads to the formation of small aggregates which are held together by specific bonding between amino acid residues. Such aggregates have been found to strongly affect the phase diagram of globular protein solutions [23].

What is the effect of aggregation on the phase diagram of globular protein solutions? The role of small aggregates in shaping the phase diagram of globular protein solutions is discussed in Chapter 4. Little is known about the phase behavior of oligomers of proteins, where only a few monomers are connected to each other. It is important to understand how to describe the effect of oligomers on the phase diagram of globular protein solutions, for such oligomers often form naturally in protein solutions because of aggregation.

As we discuss in Chapters 2 and 3, aggregation is in competition with both liquid-liquid phase separation and crystallization in protein solutions. A major problem in the growth of protein crystals for x-ray structure analysis is the formation of amorphous aggregates [24]. In addition, there are several diseases, such as cataract, where the pathology is caused both by aggregation and phase separation [25]. It is therefore essential to understand what is the practical effect of small aggregates in a solutions of globular proteins which can phase separate.

To address the questions outlined above, I have produced oligomers of globular proteins by crosslinking native proteins. The protein used is a member of the γ -crystallin family. I present experimental results for the liquid-liquid coexistence curves of the dimers and trimers of the protein. These results are analyzed within the context of the model used to describe protein monomers. It is found that the protein oligomers may be described as having longer effective ranges of interaction than the monomer protein. The experimental findings are then used to illustrate the important connection between aggregation and phase separation in globular protein solutions.

1.2. OVERVIEW OF THE THESIS

In Chapter 5, I summarize the results presented in this thesis and make suggestions for further work.

Chapter 2

The Liquid Phase

Portions of this chapter have already been published in Ref. [18]: A. Lomakin, N. Asherie, and G. B. Benedek. “Monte Carlo study of phase separation in aqueous protein solutions.” *J. Chem. Phys.* **104**, 1646 (1996). The innovative Monte Carlo scheme which we use was designed by Dr. Aleksey Lomakin.

2.1 Introduction

The liquid-liquid phase separation of protein solutions is of great interest because the factors which govern the condensation of protein into coexisting protein-poor and protein-rich phases are believed to play a central role in several human diseases [26, 27, 28]. The understanding of the location of the phase boundaries and the strategies to shift them by modifying protein interactions are key elements in the search for disease treatment. An important example of such a disease is cataract [26], where opacification of the eye lens results from alterations in the spatial distribution of the lens proteins [29]. These alterations are known to be produced, in part, by the phase separation of the γ -crystallins, a family of monomeric lens proteins [30]. Several studies [2, 3, 23, 31] have investigated the phase separation in aqueous solutions of individual members of the calf lens γ -crystallin family. These experiments show that the γ -crystallins may be divided into two groups: “high- T_c ” proteins, such as γ_{IIIa} (γ_C)

and γ_{IVa} (γ_E), which exhibit high critical temperatures ($T_c \approx 38^\circ\text{C}$), and “low- T_c ” proteins, such as γ_{II} (γ_B) and γ_{IIIb} (γ_D), which exhibit low critical temperatures ($T_c \approx 5^\circ\text{C}$). The critical volume fractions of all the γ -crystallins are approximately the same $\phi_c = 0.21 \pm 0.02$ [2]. The coexistence curves are found to be upper consolute and very broad, as is observed in some colloidal dispersions [32, 33]. Mixtures of γ -crystallins have also been studied [34, 35].

From a theoretical point of view, the phase transition in the protein-water solution is analogous to that of the single component liquid-vapor system. Beginning with the van der Waals equation of state [36], there has been a quest for an analytical equation of state for simple liquids. The general approach is to assume a form for the intermolecular potential, almost always central and pairwise additive. One of the most common selections is the square-well potential, since it is the simplest model which includes both attractive and repulsive forces. At this stage, either one of two choices is made: (i) A “fundamental” statistical-mechanical equation, such as the Percus-Yevick formula [37], is invoked, from which a closed form solution for the equation of state is obtained [38, 39, 40]; (ii) A statistical-mechanical perturbation theory is used. Here, the main approach is to treat the attractive part of the potential being studied as a perturbation to the hard sphere model, which has only repulsive forces [41, 42].

Though these theories provide a recipe for how to calculate quantities of interest for phase separation, their complexity limits their utility when interpreting experimental results. One way to overcome this difficulty, is to begin with a phenomenological thermodynamic expression for the Gibbs free energy of the system. A simple analytic model, based on mean field theory, has been proposed [3, 43] to describe the phase separation phenomena of aqueous protein solutions. As we shall see in this chapter, this model corresponds to a long-range square-well intermolecular potential.

Many Monte Carlo simulations have been made of systems which undergo phase separation [44, 45, 46]. The most recent of these [47, 48, 49] have employed the so-called Gibbs ensemble Monte Carlo technique [50, 51]. The focus of many of these studies is to examine the theory of critical phenomena for a variety of intermolecular

potentials, including the square-well potential. In the recent study of Vega *et al.* [47], the vapor-liquid phase equilibria of square-well systems with hard-sphere cores were studied for the reduced ranges $\lambda=1.25, 1.375, 1.5, 1.75$ and 2 . The critical points and the shapes of the coexistence curves (in terms of a critical exponent) were calculated. This information indicates that the interactions between the γ -crystallins are short-ranged as expected. The results, however, are not detailed enough in the short-range regime to interpret the phase diagram of these proteins.

Therefore, in order to gain insight into the microscopic interactions of the γ -crystallins, Dr. Aleksey Lomakin has developed a Monte Carlo method which we use to analyze the experimental observations of Broide *et al.* [2]. We also explore the applicability of mean field models, such as that proposed by Berland *et al.* [3] and Taratuta *et al.* [43], to aqueous γ -crystallin solutions. To simplify the analysis, as well as to save computational time, our Monte Carlo procedure uses theoretical extrapolation techniques, in addition to simulation, to calculate the quantities of interest, most importantly the chemical potential. To reconstruct the phase diagram of our model aqueous protein solution, we fit the Monte Carlo results for the chemical potential with an analytic expression. We then obtain the coexistence curve by a method analogous to the Maxwell equal areas construction for the van der Waals equation of state [52]. Since the use of an analytic form for the chemical potential neglects the contributions of critical fluctuations to the free energy of the system, our approach is unable to describe accurately the critical exponents. However, we are interested in aspects of phase separation which are not strongly affected by the fluctuations: the critical temperature T_c , the critical volume fraction ϕ_c , and the shape of the coexistence curve in regions relatively far from the critical point. To check the accuracy of our method, especially near the critical point, where the reliability of the method cannot be justified *a priori*, in Sec. 2.3.2 we compare our results with those available from other Monte Carlo simulations.

To begin our analysis, let us consider a system containing N_p protein molecules and N_w water molecules. We may write the microscopic free energy, E , of the protein-

2.1. INTRODUCTION

water solution as

$$E = E_{int} + N_w E_w^{0,w} + N_p E_p^{0,w}. \quad (2.1)$$

This form of the microscopic free energy represents a thermodynamic average of the energy of the system over all the positions of the water molecules and over the internal degrees of freedom of the proteins. Thus, E depends solely on the relative positions of the proteins. Here $E_w^{0,w}$ is the average free energy per water molecule in a solution of pure water (i.e., the chemical potential of one water molecule) and $E_p^{0,w}$ is the average free energy of one protein molecule, fixed in space, in solution of water. The interaction energy E_{int} results from the direct and indirect (i.e., through water) interactions between the proteins. The contribution of the water-water and protein-water interactions is independent of the relative positions of the proteins. However, E_{int} depends on the relative positions of the proteins and we will assume that it can be represented by a central and pairwise interaction. For convenience, we will refer to the microscopic free energy E as simply the energy of the proteins.

Now we proceed to make the model more specific. We consider the proteins to be spheres, of diameter σ , while the water is taken to be a continuous background (that is, the size of the water molecules is taken to be small as compared to σ). We assume that the effective potential energy $u(r)$ for a pair of proteins whose centers are separated by a distance r , is of the form of a square-well plus a hard core as given by Eq. (2.2) below.

$$u(r) = \begin{cases} +\infty, & \text{for } r < \sigma \\ -\epsilon, & \text{for } \sigma \leq r < \lambda\sigma \\ 0, & \text{for } r \geq \lambda\sigma \end{cases} \quad (2.2)$$

Here λ is the reduced range of the potential well and ϵ is its depth. With this potential, we can calculate the interaction energy E_{int} , and hence the total energy E , as a function of volume fraction and temperature. Note that for our particular choice of potential, we can define the number of protein-protein contacts N_{con} as the number of protein pairs whose centers are in the range $\sigma \leq r < \lambda\sigma$ from each other.

Thus we may write the interaction energy as

$$E_{int} = -N_{con}\epsilon \quad (2.3)$$

Of course Eq. (2.3) assumes that there are no overlapping hard cores in the configuration. If there are overlaps then $E_{int} = \infty$.

Once we have chosen an explicit form for the intermolecular potential of the system, we may use the Monte Carlo simulation procedure described in Sec. 2.2 to obtain the thermodynamic properties of our system. Thus we can reconstruct the phase diagram of our model aqueous protein solution and compare it with the experimental results of Broide *et al.* [2]. This comparison is made in Sec. 2.3.4.

2.2 Computer Simulation

The conditions for phase equilibrium in the protein-water solution are

$$\mu_p(I) = \mu_p(II) \quad (2.4)$$

$$\mu_w(I) = \mu_w(II), \quad (2.5)$$

where μ_p and μ_w are the chemical potentials of the protein and water respectively. Here I and II denote the two coexisting phases. We also write as a shorthand $\mu_p(I) \equiv \mu_p(\phi^I, T)$ and $\mu_p(II) \equiv \mu_p(\phi^{II}, T)$ and similarly for μ_w . We let ϕ^I and ϕ^{II} be the protein volume fractions in the two phases and we take $\phi^I \leq \phi^{II}$ without loss of generality. The temperature of the system is denoted by T . We may write Eqs. (2.4) and (2.5) in an alternative form, namely

$$\mu(I) = \mu(II) \quad (2.6)$$

$$\int_{\phi^I}^{\phi^{II}} \mu(\phi, T) d\phi = \left(\frac{\phi^{II} - \phi^I}{2} \right) [\mu(I) + \mu(II)]. \quad (2.7)$$

Here $\mu \equiv \mu_p - \gamma\mu_w$, where $\gamma \equiv \Omega_p/\Omega_w$, is the ratio of the volume of one protein

2.2. COMPUTER SIMULATION

molecule (Ω_p) to the volume of one water molecule (Ω_w). The volume fraction is defined as $\phi \equiv N_p \Omega_p / V$, with V the total volume of the system. The quantity μ represents the change in free energy (at constant volume) due to the replacement of γ water molecules by one protein molecule. We will work with μ for it is the analog of the chemical potential in a one component system. This equivalence is shown in Appendix A. Equation (2.6) follows directly from Eqs. (2.4) and (2.5). Equation (2.7) is equivalent to the equal areas rule proposed by Maxwell [52] for a pure fluid and can be derived by integrating by parts the Gibbs-Duhem relation, $\phi(\partial\mu_p/\partial\phi) + \gamma(1 - \phi)(\partial\mu_w/\partial\phi) = 0$, from ϕ^I to ϕ^{II} and using Eqs. (2.5) and (2.6), together with the definition of μ . We will call μ “the effective chemical potential” to distinguish it from the protein or water chemical potentials.

We can see that in order to use Eqs. (2.6) and (2.7) to study the phase separation of the system, we need to know the effective chemical potential as an analytic function of volume fraction and temperature. We should note that below the critical temperature T_c the effective chemical potential, as an analytic function of the volume fraction, has a region of negative slope. In this region, the system is unstable to microscopic fluctuations. At the critical temperature, this region reduces to a point with critical volume fraction ϕ_c . At the critical point, both the first and second derivatives of the effective chemical potential with respect to volume fraction are zero [53]. Thus, the two equations

$$\left. \frac{\partial\mu}{\partial\phi} \right|_{\phi_c, T_c} = 0 \quad (2.8)$$

$$\left. \frac{\partial^2\mu}{\partial\phi^2} \right|_{\phi_c, T_c} = 0 \quad (2.9)$$

determine the values of ϕ_c and T_c . The spinodal, which marks the boundary between areas of the phase diagram where the system is stable and unstable, is given by

$$\frac{\partial\mu}{\partial\phi} = 0. \quad (2.10)$$

According to Eqs. (2.6)-(2.10), if μ is known as a function of ϕ and T , then the whole phase diagram may be constructed. By focusing our attention on μ , we may

simplify the study of the phase separation phenomena as follows:

(i) We obtain the effective chemical potential, $\mu(\phi, T_0)$, as a function of volume fraction at a temperature T_0 above T_c , by using Widom's formula [54] (see also Appendix B):

$$\mu = \mu_0 + kT \ln \phi - kT \ln \langle \exp(-\Delta E^{test}/kT) \rangle. \quad (2.11)$$

Here, $\langle \rangle$ denotes a canonical ensemble average for the system at constant volume and temperature and ΔE^{test} is the change in the microscopic free energy of the system due to the addition of a test particle. As is shown in Appendix B, the standard part of the chemical potential is given by $\mu_0 \equiv -kT \ln(\Omega_p/V_F)$ (where V_F is the Fermi volume) and it is independent of volume fraction.

For our particular system

$$\Delta E^{test} \equiv E(N_p + 1, N_w - \gamma) - E(N_p, N_w). \quad (2.12)$$

Using Eqs. (2.1) and (2.3) we obtain

$$\Delta E^{test} = -\nu\epsilon + E_p^{0,w} - \gamma E_w^{0,w}, \quad (2.13)$$

where ν , the number of new contacts made by the test protein, is given by

$$\nu \equiv N_{con}(N_p + 1) - N_{con}(N_p). \quad (2.14)$$

We note that Eq. (2.13) presupposes that the hard core of the test protein does not overlap with any other hard core. In the case of an overlap $\Delta E^{test} = \infty$. Substituting Eq. (2.13) into Eq. (2.11), we obtain the following form for the chemical potential

$$\mu = \mu_0 + E_p^{0,w} - \gamma E_w^{0,w} + kT[\ln \phi - \ln \langle \exp(\nu\hat{\epsilon}) \rangle], \quad (2.15)$$

where $\hat{\epsilon} \equiv \epsilon/kT$ is the reduced energy.

In Eq. (2.15) the ensemble average, represented by $\langle \rangle$, is to be taken over *all*

2.2. COMPUTER SIMULATION

attempts to add the test particle: both those for which there are no hard core overlaps and those for which hard core overlaps do occur. In the latter case $\Delta E^{test} = \infty$ and the corresponding exponential in Eq. (2.15) should be set to zero. For example, for $\hat{\epsilon} = 0$, which is the hard sphere limit, the quantity $\langle \exp(\nu\hat{\epsilon}) \rangle$ reduces to the ratio of the number of successful attempts to add a test particle to the total number of attempts.

We now introduce the reduced chemical potential $\hat{\mu}(\phi, T)$, defined as

$$\hat{\mu} \equiv \ln o - \ln \langle \exp(\nu\hat{\epsilon}) \rangle. \quad (2.16)$$

We can see from Eq. (2.15) that $\mu = kT\hat{\mu} + \mu_0 + E_p^{0,w} - \gamma E_w^{0,w}$. Since the last three terms of this expression will cancel in Eqs. (2.6) and (2.7), the phase diagram is determined entirely by $\hat{\mu}$. In fact, we can replace μ with $\hat{\mu}$ in Eqs. (2.6) and (2.7). As a shorthand, we will refer to $\hat{\mu}$ as the chemical potential. We use Monte Carlo simulations to calculate the quantity $\langle \exp(\nu\hat{\epsilon}) \rangle$.

(ii) We assume that the chemical potential may be represented by an analytic form, which we use to explicitly carry out the integration in Eq. (2.7). We will see in the next section that the error introduced by this approach in the reconstruction of the phase diagram is small, whereas the savings in computational time are great. We fit the Monte Carlo results to the following expression for $\hat{\mu}(\phi, T)$

$$\hat{\mu}(o, T) = \hat{\mu}_{CS}(\phi) + \sum_{n=1}^{n_0} A_n(T)\phi^n. \quad (2.17)$$

Here

$$\hat{\mu}_{CS} = \ln \phi - 3 + \frac{3 - \phi}{(1 - \phi)^3}. \quad (2.18)$$

In Eq. (2.18), $\hat{\mu}_{CS}$ is the Carnahan-Starling [55] approximation for the chemical potential of an assembly of hard spheres. The $A_n(T)$ of Eq. (2.17) are temperature-dependent coefficients to be determined. The parameter n_0 is chosen so as to obtain a smooth representation of the chemical potential. If n_0 is too large, the fit tends to follow in detail the statistical errors of the Monte Carlo simulation. On the other

hand, for small n_0 the systematic deviation of the fit from the Monte Carlo results becomes large. We typically choose $n_0 = 4$. Note that for $\hat{\mu}(\phi, T)$ to have the correct high temperature behavior (i.e., to reduce to the hard sphere limit), $A_n(T)$ must, to within the accuracy of the Carnahan-Starling approximation (Eq. (2.18)), tend to zero as $\epsilon/kT \rightarrow 0$. The form of the chemical potential, as given in Eq. (2.17), was chosen not only because it has the correct high temperature limit, but also because it properly reproduces the low ϕ behavior and it conveniently reduces to the mean field theory result if we set $n_0 = 1$. This is discussed further in Sec. 2.3.3.

We estimate T_c by extrapolating the chemical potential, as explained below, downward in temperature until we find a point where both Eqs. (2.8) and (2.9) are satisfied. We perform accurate Monte Carlo simulations at a temperature T_1 , where T_1 is within one percent above our estimated T_c , to find $\hat{\mu}(\phi, T_1)$. Note that the simulations are all performed in the single phase regime where the system is thermodynamically stable. We then use the extrapolation procedure described in the next paragraph to obtain a series of chemical potential isotherms for temperatures below T_c . Using these isotherms we are able to find the locations of the phase boundaries without any further time-consuming simulations.

To perform the temperature extrapolation, we expand the chemical potential at a temperature $T_2 < T_1$ in powers of $\Delta\hat{\epsilon} \equiv \hat{\epsilon}_2 - \hat{\epsilon}_1$ with $\hat{\epsilon}_1 \equiv \epsilon/kT_1$ and $\hat{\epsilon}_2 \equiv \epsilon/kT_2$. Here we take advantage of the fact that the chemical potential is a function of temperature only through the reduced energy $\hat{\epsilon}$. To first order we have

$$\hat{\mu}(\phi, T_2) = \hat{\mu}(\phi, T_1) + \Delta\hat{\epsilon} \left. \frac{\partial \hat{\mu}}{\partial \hat{\epsilon}} \right|_{\hat{\epsilon}_1}. \quad (2.19)$$

As is shown in Appendix C, the derivative $\partial \hat{\mu} / \partial \hat{\epsilon}$ may be written as

$$\frac{\partial \hat{\mu}}{\partial \hat{\epsilon}} = \frac{1}{\epsilon} \left(\frac{\partial \bar{E}_{int}}{\partial N_p} \right)_{V, T}. \quad (2.20)$$

2.2. COMPUTER SIMULATION

where \overline{E}_{int} is the average interaction energy of the system. Using Eq. (2.3) we have

$$\overline{E}_{int} = -\overline{N}_{con}\epsilon = -\frac{1}{2}N_p\overline{\eta}\epsilon. \quad (2.21)$$

Here \overline{N}_{con} is the average number of protein-protein contacts and $\overline{\eta}$ is the average number of contacts which each particle makes. Substituting Eq. (2.21) into Eq. (2.20), we reexpress Eq. (2.19) as

$$\hat{\mu}(\phi, T_2) = \hat{\mu}(\phi, T_1) - \frac{\Delta\hat{\epsilon}}{2} \frac{\partial}{\partial\phi} [\phi\overline{\eta}(\phi, T_1)]. \quad (2.22)$$

The quantity $\overline{\eta}(\phi, T_1)$ is also calculated during the Monte Carlo simulation at temperature T_1 using a cell list [58] to keep track of the number of neighbors surrounding each particle. Note that $\overline{\eta}$ is not equal to $\langle \nu \rangle$, the ensemble average of the number of contacts made by the test particle, for the test particle is not in thermodynamic equilibrium with the other particles in the system.

Therefore, once we have performed the Monte Carlo simulation at temperature T_1 , we construct $\hat{\mu}(\phi, T_2)$ by using Eq. (2.22). We have found empirically that we may reliably employ our temperature expansion provided that $\Delta\hat{\epsilon}/\hat{\epsilon}_1$ is less than ten percent. At each temperature we fit the Monte Carlo results for the chemical potential (both those obtained by direct simulation and those obtained from our extrapolation procedure) by Eq. (2.17) with the appropriate values of the coefficients $A_n(T)$. Once we have an analytic representation for the chemical potential we use Eqs. (2.6) and (2.7) to calculate the coexisting phases at each temperature and hence obtain the coexistence curve. In this way, we have calculated the critical volume fraction ϕ_c , the reduced critical energy $\hat{\epsilon}_c$, the spinodal and the coexistence curve, for a large number of square-well reduced ranges λ between 1.05 and 2.40.

For the Monte Carlo simulation we randomly place our particles inside a cube of unit volume with the usual periodic boundary conditions [56]. The hard core diameter, σ , of the particles was chosen to be in the range 0.14–0.18, so that we have $N = 100$ – 250 particles at the highest volume fractions $\phi = 0.3$ – 0.4 for which we perform the simulation. Note that here $\phi \equiv \frac{1}{6}\pi\sigma^3N$. To generate a statistical ensemble of

configurations, the particles are displaced using a time-saving modification of the well-established NVT Metropolis scheme [57]. In our scheme, as in the Metropolis scheme, the displacement of a particle is accepted unconditionally if the change in the total energy of the system, ΔE , due to the displacement, is negative and with probability $\exp(-\Delta E/kT)$ if ΔE is positive. From the ensemble so generated, we may calculate the quantities $\bar{\eta}$ and $\langle \exp(\nu\hat{\epsilon}) \rangle$ which are needed in Eqs. (2.16) and (2.22) to obtain chemical potential isotherms. The quantity $\langle \exp(\nu\hat{\epsilon}) \rangle$ is found through the addition of test particles to the system. To accumulate statistically significant information on the average value of $\exp(\nu\hat{\epsilon})$, we must continue testing each configuration of the system until, on average, at least one successful attempt to add a test particle is made. A successful attempt is one for which the core of the test particle does not overlap with that of any other particle. Thus, the addition of test particles not only enables us to calculate the quantity $\langle \exp(\nu\hat{\epsilon}) \rangle$, but also provides information on acceptable new positions for the particles of the system. Using this information, in our scheme we generate new members of the ensemble (as described in the following paragraph) by moving particles to any acceptable position inside the simulation volume. This should be contrasted with the standard choice for particle displacement. Usually, the step size for particle displacement is chosen in such a way that approximately half of the trial configurations are accepted [59]. This is the rule of thumb to optimize the speed of evolution of the system. However, since we are only interested in the chemical potential, we may use the same information for chemical potential tests and particle repositioning. In this way the system evolves several times faster than in the standard (small step size) algorithm. Naturally, the results of the two methods are the same. The use of the same information for chemical potential tests and particle repositioning in no way biases the results: if we refrain from calculating the chemical potential we simply have a Metropolis equilibration algorithm.

We describe below our Monte Carlo algorithm. A more detailed flowchart representation is given in Appendix D.

The fundamental cycle in our Monte Carlo simulation consists of the following

2.2. COMPUTER SIMULATION

sequence of steps : (i) A particle is selected at random. (ii) An attempt is made to add a new test particle at a randomly chosen position. (iii) If the attempt is successful, the number of contacts made by the test particle is calculated. (iv) For a successful attempt, or an unsuccessful one where the test particle *only* overlaps with the single particle selected in step (i), the next configuration is created by moving the particle chosen in step (i) to the position of the test particle. This last move is accepted in the standard way i.e., it is accepted unconditionally if the change in energy due to the move, ΔE , is negative and with probability $\exp(-\Delta E/kT)$ if ΔE is positive.

We see that the test particle of step (ii) can be thought of as simply a label for the position to which we are trying to move the particle chosen in step (i). By steps (i), (ii), and (iv), we generate the members of the canonical ensemble. Furthermore, during step (iii), the test particle is also used to calculate the chemical potential of the system by means of Widom's formula (see Eqs. (2.11) and (2.16)). The algorithm we have outlined above is significantly faster than one in which the evolution of the system, through small steps, is carried out independently of the calculation of the chemical potential.

For a given reduced range λ , we performed our main Monte Carlo simulation at a reduced energy, $\hat{\epsilon}_1 \equiv \epsilon/kT_1$, within one percent below the reduced critical energy, $\hat{\epsilon}_c \equiv \epsilon/kT_c$, as shown in Table 2.1. The reduced critical energy was estimated from auxiliary simulations by using the temperature extrapolation method. Our main simulation was continued until the statistical errors in $\hat{\mu}$ were no greater than the uncertainties associated with the analytic fit of $\hat{\mu}$ (Eq. (2.17)). As usual, the system was allowed to equilibrate before testing for the chemical potential [60]. Extrapolation and fitting techniques were used, as explained previously, to obtain the phase diagram. It should be noted that the final determination of ϕ_c and $\hat{\epsilon}_c$ (through Eqs. (2.8) and (2.9)) is made by a small extrapolation of the results from the thorough simulation carried out at $\hat{\epsilon}_1$. Thus, systematic errors in these quantities are very small. The whole procedure was repeated for a large number of reduced ranges $1.05 \leq \lambda \leq 2.40$.

In the next section we present the results of our Monte Carlo study.

2.3 Results and Discussion

2.3.1 Results of this Study

We begin our discussion by illustrating our temperature extrapolation method. In Fig. 2-1 we compare the direct Monte Carlo results for the chemical potential with those obtained by extrapolation. The open symbols represent the simulation results of the chemical potential for $\lambda = 1.25$ at three different values of the reduced energy $\hat{\epsilon} = 1.318$ (triangles), 1.267 (circles), and 1.216 (squares). The dashed lines are the chemical potentials obtained by extrapolating the chemical potential at $\hat{\epsilon} = 1.267$ to $\hat{\epsilon} = 1.318$ (coarse dashed line) and to $\hat{\epsilon} = 1.216$ (fine dashed line) using Eq. (2.22). The solid line is the analytic fit of Eq. (2.17) with $n_0 = 4$ to the $\hat{\epsilon} = 1.267$ Monte Carlo results. We see that the chemical potentials obtained by extrapolating from $\hat{\epsilon} = 1.267$ to either $\hat{\epsilon} = 1.318$ or $\hat{\epsilon} = 1.216$ (i.e., $\pm 4\%$ of the original temperature) are in satisfactory agreement with those calculated directly at $\hat{\epsilon} = 1.318$ and $\hat{\epsilon} = 1.216$ by Monte Carlo simulation. We find similar agreement between the simulation results and the extrapolation method over the whole range of λ studied: $1.05 \leq \lambda \leq 2.40$. This gives us confidence to use the extrapolation procedure in place of the many time-consuming simulations that would otherwise be required.

The coexisting volume fractions may then be determined from the chemical potential isotherms by applying Eqs. (2.6) and (2.7). The points which lie on the spinodal are given by Eq. (2.10). An example of the construction of the coexistence curve and spinodal is shown in Fig. 2-2 for the case $\lambda = 1.25$. The open circles are the Monte Carlo results for the chemical potential with $\lambda = 1.25$ and $\hat{\epsilon} = 1.267$. The isotherms which result from the temperature extrapolation (from $\hat{\epsilon} = 1.267$ to $\hat{\epsilon} = 1.317$ in steps of 0.005) are shown as solid lines. The coexisting points and spinodal points at each temperature are shown as dashes and crosses respectively. The coexistence curves so constructed are shown in Fig. 2-3 for the reduced ranges $\lambda = 1.8, 1.5, 1.25$, and 1.1. The coexistence curves become broader as the range of the interaction decreases, and the corresponding critical volume fraction increases.

In Table 2.1 we list the results for a group of representative Monte Carlo simula-

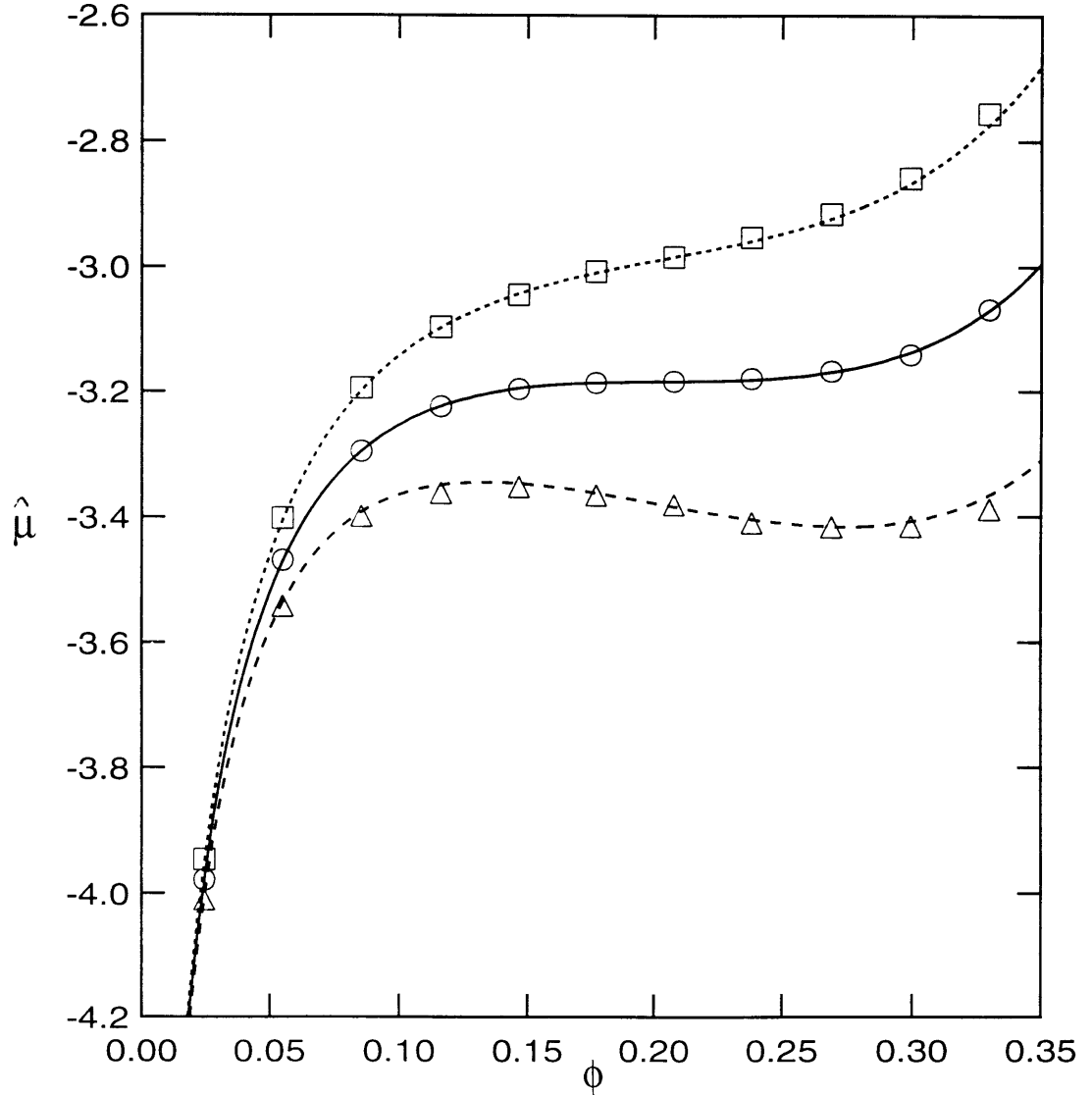


Figure 2-1: Illustration of the temperature expansion method for $\lambda = 1.25$. The open symbols represent the Monte Carlo results for the chemical potential for three different values of the reduced energy $\hat{\epsilon} = 1.318$ (triangles), 1.267 (circles), and 1.216 (squares). The solid line is a fit to the $\hat{\epsilon} = 1.267$ Monte Carlo results using Eq. (2.17) with $n_0 = 4$. The dashed lines are the chemical potentials obtained by extrapolating the $\hat{\epsilon} = 1.267$ chemical potential to $\hat{\epsilon} = 1.318$ (coarse dashed line) and to $\hat{\epsilon} = 1.216$ (fine dashed line).

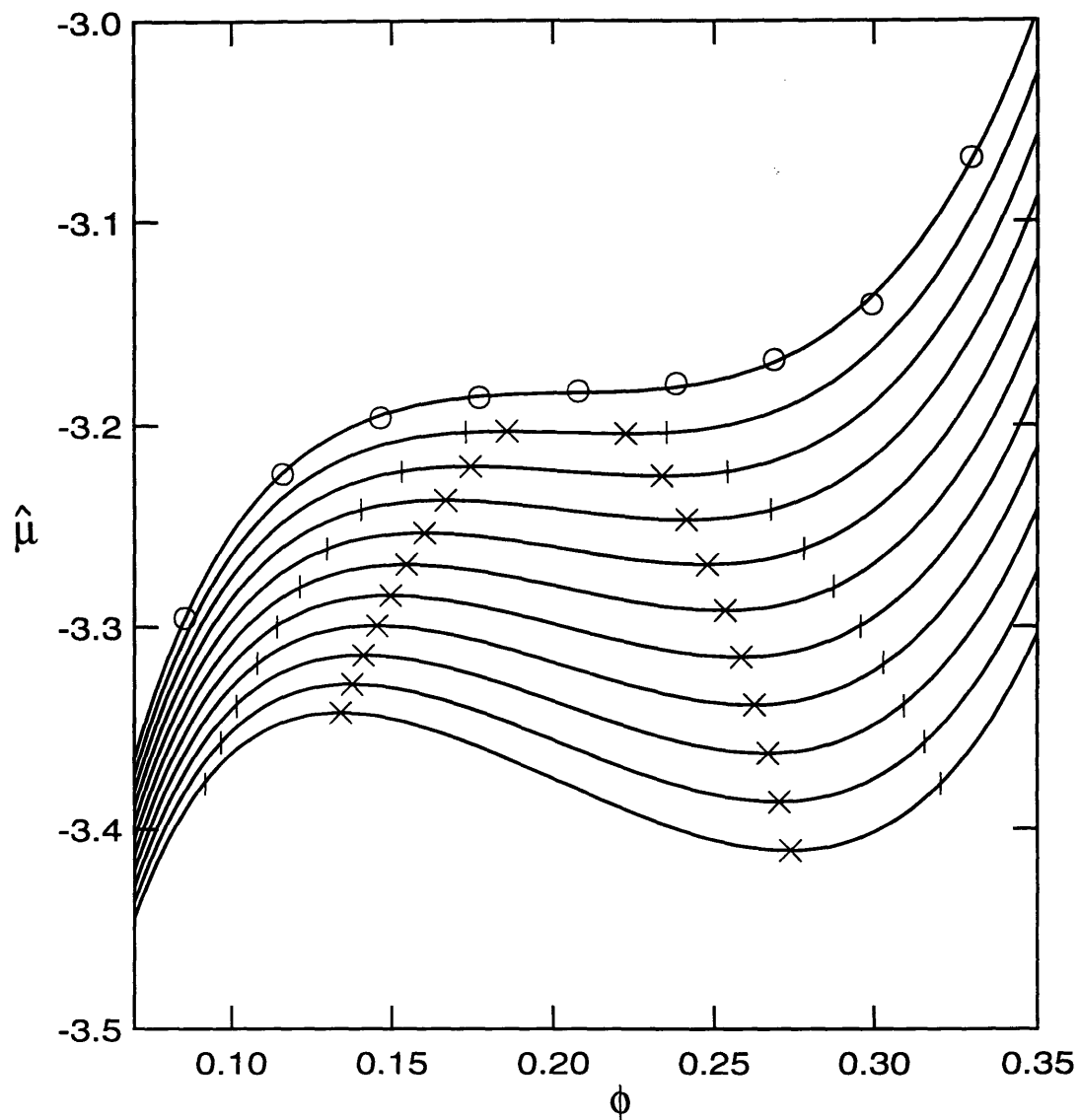


Figure 2-2: Reconstruction of the spinodal and the coexistence curve. The open circles are the Monte Carlo results for the chemical potential with $\lambda = 1.25$ and $\hat{\epsilon} = 1.267$. The isotherms which result from the temperature extrapolation (from $\hat{\epsilon} = 1.267$ to $\hat{\epsilon} = 1.317$ in steps of 0.005) are shown as solid lines. The coexisting points and spinodal points at each temperature are shown as dashes and crosses respectively.

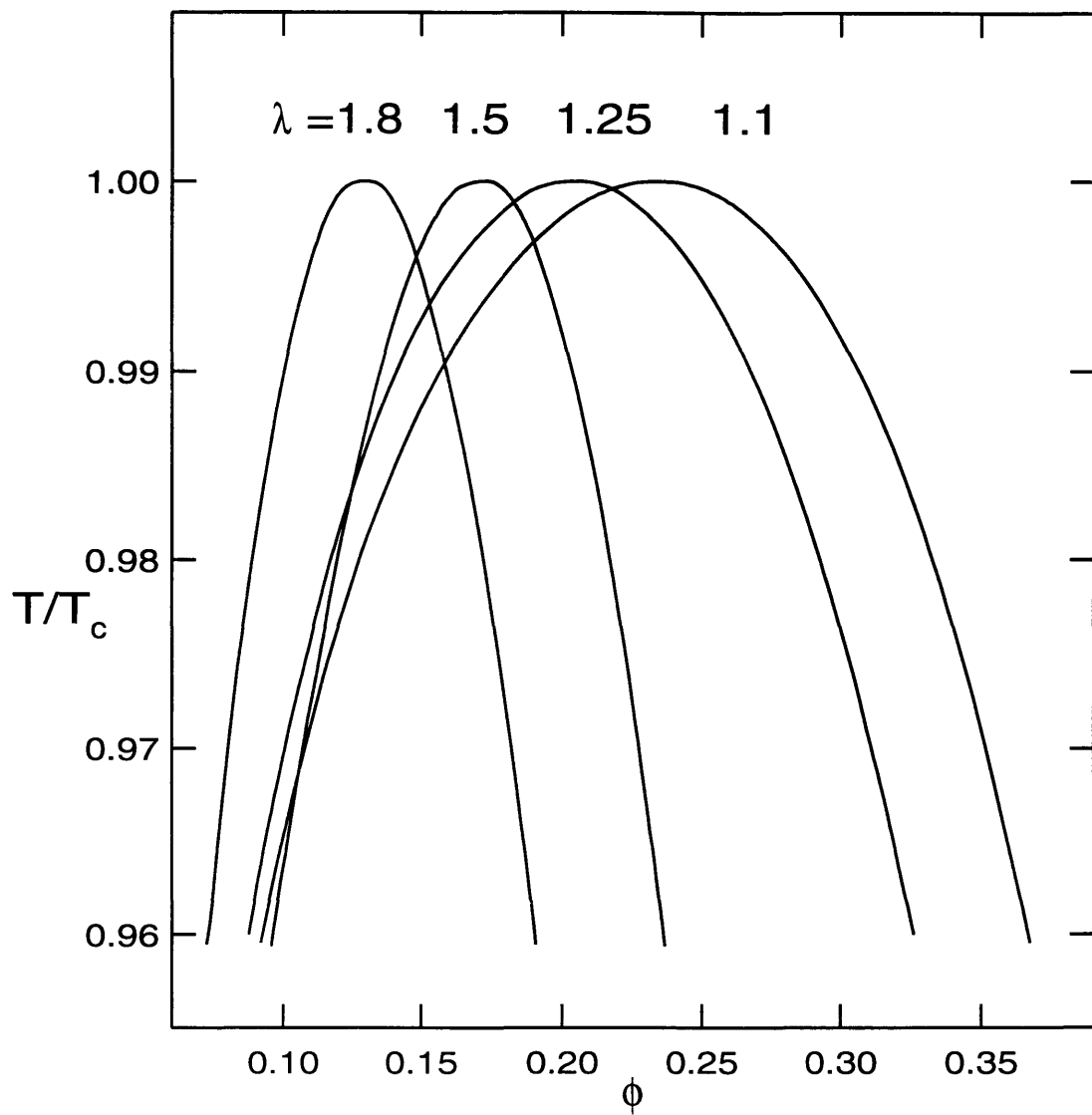


Figure 2-3: Coexistence curves. The curves with progressively larger widths represent the results obtained for the reduced ranges $\lambda=1.8$, 1.5, 1.25 and 1.1.

Table 2.1: Results and parameters from the Monte Carlo simulations at different reduced ranges λ ($\lambda = \infty$ is the mean field limit). The quantities presented are: (i) the critical volume fraction ϕ_c ; (ii) the reduced critical energy $\hat{\epsilon}_c$; (iii) the average number of contacts per particle at the critical point $\bar{\eta}_c$; (iv) the number of successful attempts made K_{tot} in units of 10^6 ; (v) the reduced energy at which the simulation is performed $\hat{\epsilon}_1$; (vi) the diameter of the particles σ ; (vii) the maximum number of particles used in the simulation M_{max} ; and (viii) the maximum volume fraction ϕ_{max} .

λ	ϕ_c	$\hat{\epsilon}_c$	$\bar{\eta}_c$	K_{tot}	$\hat{\epsilon}_1$	σ	M_{max}	ϕ_{max}
∞	0.134	0.000	∞	0.5	0.000	0.16	178	0.38
∞	0.132	0.000	∞	0.9	0.000	0.14	248	0.36
2.40	0.140	0.197	15.51	1.0	0.195	0.16	140	0.30
2.40	0.140	0.197	15.64	4.7	0.196	0.14	208	0.30
2.20	0.135	0.263	11.81	2.9	0.260	0.14	228	0.33
2.20	0.135	0.262	11.59	1.7	0.260	0.16	140	0.30
2.00	0.126	0.361	8.58	4.1	0.357	0.14	218	0.31
2.00	0.126	0.359	8.45	3.0	0.357	0.16	142	0.30
1.80	0.132	0.487	6.71	6.0	0.480	0.16	148	0.32
1.80	0.129	0.483	6.46	7.7	0.480	0.18	118	0.36
1.65	0.146	0.610	5.95	7.3	0.608	0.16	142	0.30
1.65	0.149	0.606	5.95	9.1	0.605	0.18	110	0.34
1.50	0.171	0.763	5.38	12.8	0.760	0.18	110	0.34
1.50	0.166	0.773	5.33	5.5	0.767	0.16	154	0.33
1.40	0.172	0.935	4.85	3.4	0.930	0.14	218	0.31
1.40	0.173	0.922	4.72	20.6	0.920	0.18	100	0.31
1.30	0.194	1.129	4.44	15.5	1.127	0.18	110	0.34
1.30	0.193	1.128	4.41	6.7	1.127	0.18	110	0.34
1.25	0.205	1.269	4.27	21.5	1.267	0.18	110	0.34
1.25	0.206	1.270	4.31	14.0	1.267	0.18	100	0.31
1.20	0.219	1.449	4.17	5.5	1.435	0.16	169	0.36
1.20	0.216	1.443	4.09	17.1	1.435	0.18	110	0.34
1.15	0.227	1.693	3.91	10.7	1.680	0.16	169	0.36
1.15	0.227	1.673	3.83	24.0	1.660	0.18	110	0.34
1.10	0.235	2.035	3.54	31.9	2.015	0.16	169	0.36
1.10	0.244	2.038	3.63	61.6	2.015	0.18	118	0.36
1.05	0.246	2.667	3.16	122.1	2.650	0.18	125	0.38
1.05	0.273	2.665	3.42	81.4	2.650	0.18	125	0.38

2.3. RESULTS AND DISCUSSION

tions for different values of the reduced range, λ . For each value of λ listed in column 1, we present in columns 2–4 the corresponding results we obtained for the critical volume fraction ϕ_c , the critical reduced energy $\hat{\epsilon}_c \equiv \epsilon/kT_c$, and the average number of contacts per particle, $\bar{\eta}_c$, at the critical point. Note that all the results in the table were obtained using $n_0 = 4$. The manner in which we obtained the results in the $\lambda = \infty$ case will be discussed in Sec. 2.3.3.

To gain insight into the accuracy of the results, we varied the conditions under which the simulations were made. We list in columns 5–9 of Table 2.1 the simulation parameters that we varied: K_{tot} , the total number of successful attempts (in units of 10^6) made during the testing of the chemical potential at each volume fraction, the reduced energy, $\hat{\epsilon}_1$, at which the simulation is performed, σ , the diameter of the particles, M_{max} , the maximum number of particles used in the simulation, and ϕ_{max} , the maximum volume fraction at which the simulation was carried out (note that $\phi_{max} = M_{max}\pi\sigma^3/6$). For each reduced range presented in Table 2.1, we show the results obtained with two different sets of simulation parameters. Although it is difficult to evaluate *a priori* the systematic errors inherent in our method, we may estimate *a posteriori* our errors by using the variation in the values of the quantities of interest: the critical volume fraction ϕ_c , the reduced critical energy $\hat{\epsilon}_c$, and the average number of contacts per particle at the critical point $\bar{\eta}_c$. We see that these quantities vary by no more than a few percent between the different runs for a given λ .

Another source of systematic errors which we investigated is that brought about by our particular choice of fit to Eq. (2.17). Different fits will result in different values for the critical volume fraction and the critical temperature. In Table 2.2 we show the values of ϕ_c and $\hat{\epsilon}_c$ obtained using different values of n_0 at three different ranges: $\lambda=1.8$, 1.3 and 1.1. We see that the variation in ϕ_c and $\hat{\epsilon}_c$ due to the change in n_0 is of the order of the errors shown in Table 2.1. Thus we conclude that, for $n_0=3$, 4, or 5, our results are relatively insensitive to the value of n_0 chosen. However, we do find that for n_0 below three the fit does not give an adequate representation of the chemical potential, while for n_0 above five, the fit begins to follow the statistical

errors of the simulation results.

Table 2.2: Variation of the simulation results with the order of the chemical potential fit, n_0 .

λ	n_0	ϕ_c	$\hat{\epsilon}_c$
1.80	3	0.126	0.466
1.80	4	0.129	0.483
1.80	5	0.132	0.484
1.30	3	0.189	1.131
1.30	4	0.194	1.129
1.30	5	0.196	1.132
1.10	3	0.251	2.006
1.10	4	0.244	2.038
1.10	5	0.245	2.038

2.3.2 Other Monte Carlo Results

In view of the non-orthodox nature of our calculational procedure, which involves analytic techniques as well as simulations, it is useful to compare our results with those available from conventional Monte Carlo simulations. We present, in Fig. 2-4, the coexistence curves from our simulations at $\lambda=1.25$ (coarse dashed line) and 1.5 (fine solid line), together with the coexisting points for the same values of λ as obtained by Vega *et al.* (open circles and squares respectively) [47]. We can see that the agreement between the two simulations is satisfactory, even though we extend our coexistence curves to temperatures significantly below the critical point. Note that $T^* = kT/\epsilon$.

We believe that our approach provides a better way to estimate ϕ_c and $\hat{\epsilon}_c$ than the conventional Monte Carlo method. As can be seen from Fig. 2-4, the Gibbs ensemble Monte Carlo simulations of two coexisting phases [47] are impractical to carry out close to the critical point. Therefore, the critical parameters of those calculations must still be obtained from some form of extrapolation. Our simulations are carried out very close to the critical temperature allowing for a better estimation of ϕ_c and $\hat{\epsilon}_c$.

In Figs. 2-5 and 2-6 we compare our deduced critical volume fractions and reduced critical energies with those found by conventional Monte Carlo simulations. In Fig. 2-5, we show our results (solid circles) for the critical volume fraction ϕ_c as a function of the reduced range λ . We also show in Fig. 2-5 the results for ϕ_c , as obtained by other Monte Carlo simulations: (i) Henderson *et al.* [45] use an NVT algorithm (open squares); (ii) Vega *et al.* [47] use a Gibbs ensemble Monte Carlo simulation (open triangles); (iii) Lomba *et al.* [19] use a Gibbs ensemble Monte Carlo simulation but choose a Yukawa potential instead of a square-well (open circles). The corresponding results for the reduced critical energy $\hat{\epsilon}_c$ are shown in Fig. 2-6. We have converted the Yukawa potential parameters into those of an equivalent square-well by taking the depth of the two potentials to be the same and requiring the high temperature limit of the second virial coefficients to be equal. The Yukawa potential results of Lomba *et al.* [49] illustrate that the phase separation phenomena do not depend on

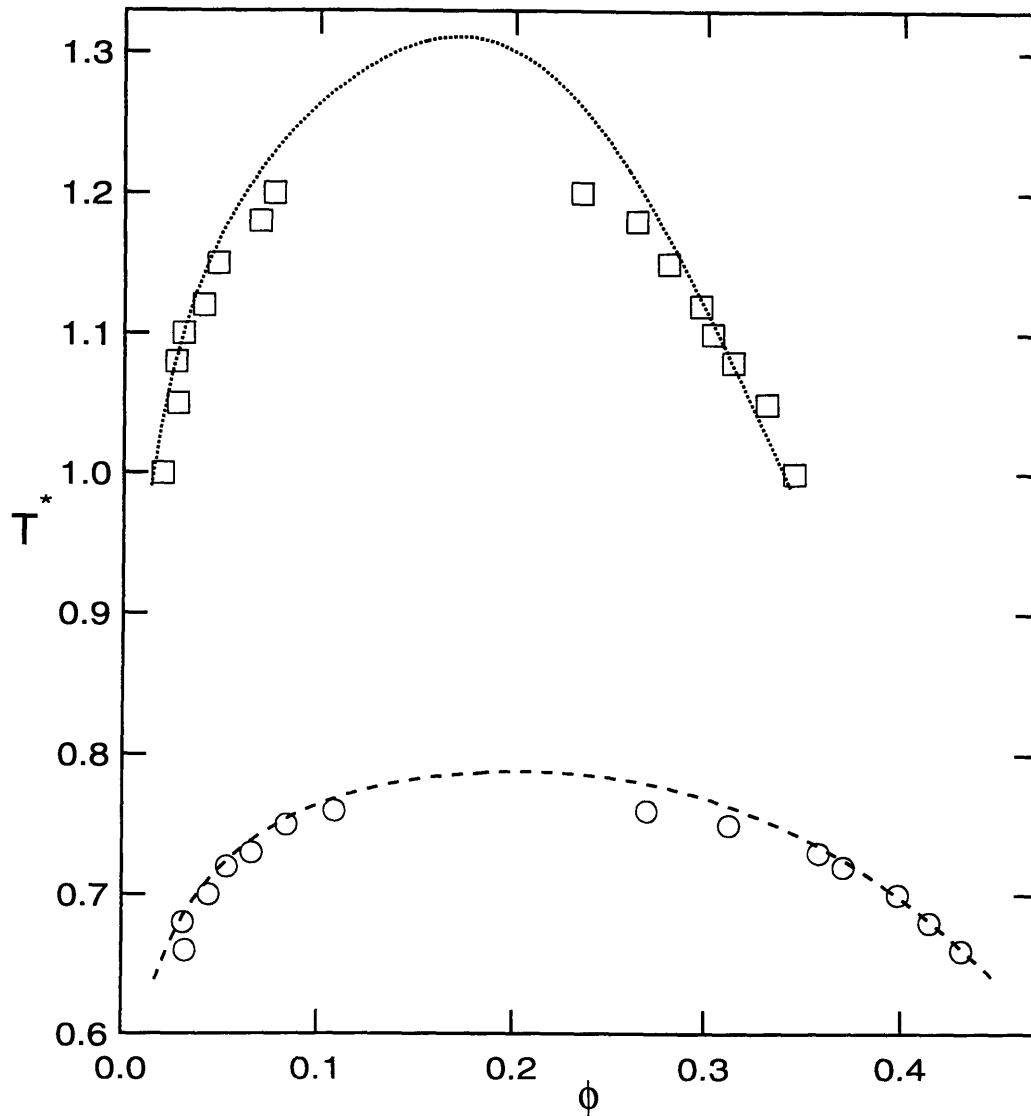


Figure 2-4: Comparison of the coexistence curves. The coexistence curves from our simulations at $\lambda=1.25$ (coarse dashed line) and 1.5 (fine dashed line) are shown together with the coexisting points obtained by Vega *et al.* [47] for the same ranges (open circles and squares respectively). Note that $T^* = kT/\epsilon$.

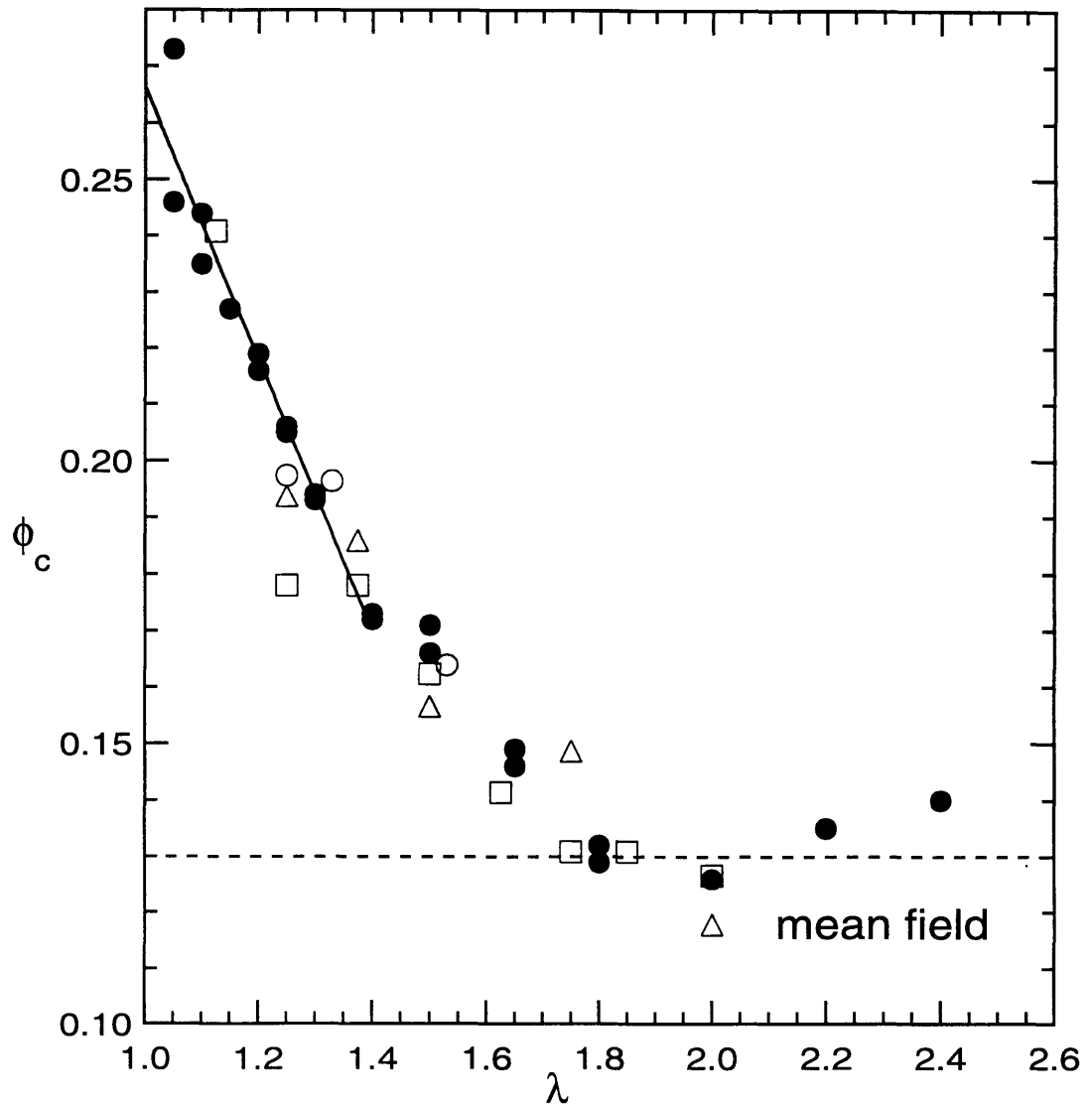


Figure 2-5: Variation of the critical volume fraction with the reduced range. Our results (solid circles) are presented together with those of Henderson *et al.* [45] (open squares), Vega *et al.* [47] (open triangles), and Lomba *et al.* [49] (open circles). The solid line is a linear extrapolation of our results to $\lambda = 1$. The dashed line is the mean field result.

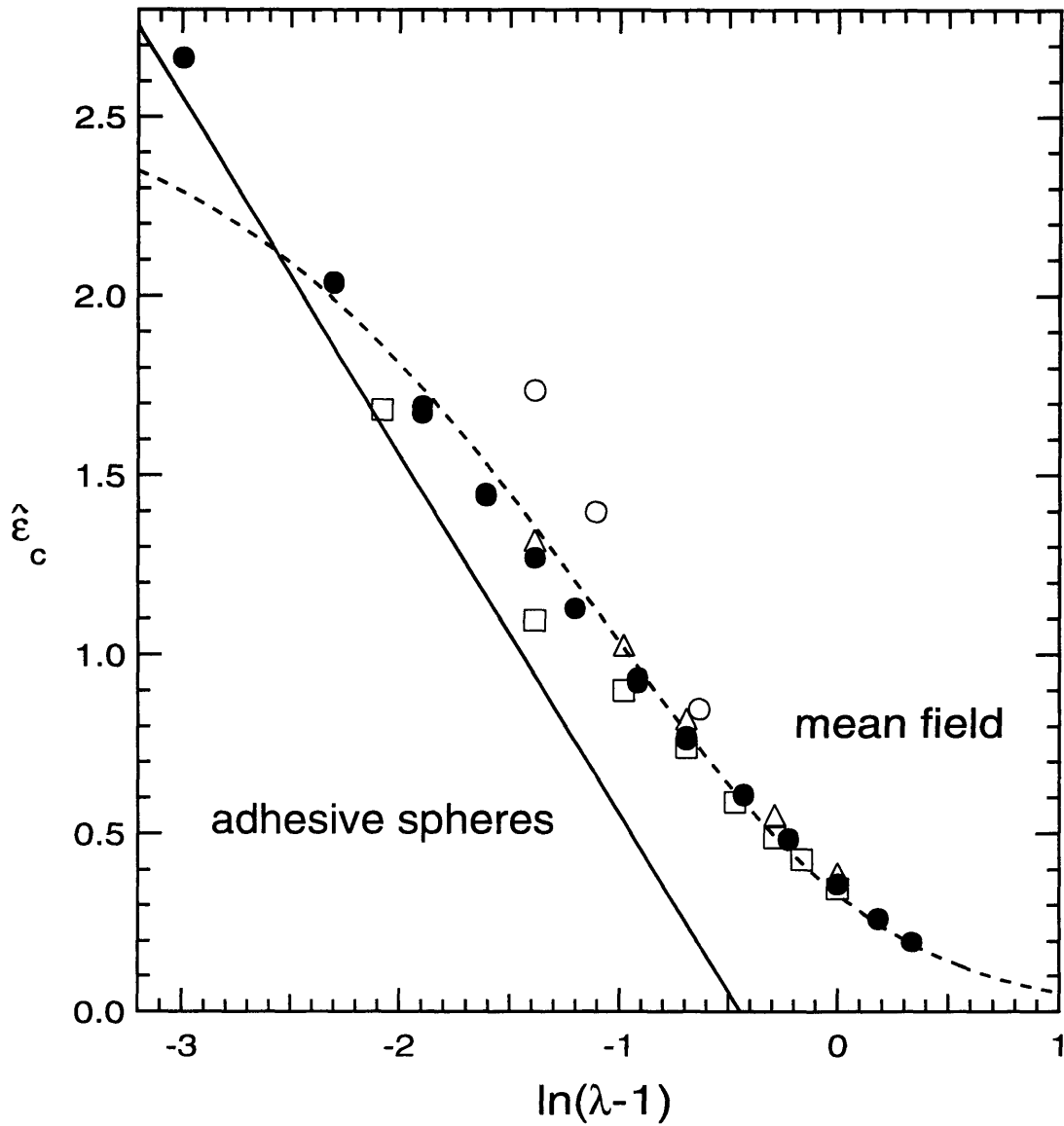


Figure 2-6: Variation of the critical reduced energy with the reduced range λ . Our results (solid circles) are presented together with those of Henderson *et al.* [45] (open squares), Vega *et al.* [47] (open triangles), and Lomba *et al.* [49] (open circles). The solid line is Eq. (2.31) with $\tau_c = 0.13$. The dashed line is Eq. (2.30) with $a_c = 10.6$.

the detailed form of the potential chosen. We see that the mean values we find for ϕ_c and $\hat{\epsilon}_c$ are consistent with those found by others. In addition, the uncertainty in our results is smaller than that obtained by conventional Monte Carlo simulations.

The effectiveness of our approach is especially important for short range potentials which are the focus of our study. Monte Carlo simulations become increasingly time-consuming as $\lambda \rightarrow 1$ [46]. However, the time saved by our use of analytic methods allows us to thoroughly investigate the short range regime as can be seen from Fig. 2-5.

2.3.3 Connection with the Mean Field and Adhesive Sphere Models

As we have seen above, our Monte Carlo calculations provide a description of the phase diagram over a wide domain of λ : $1.05 \leq \lambda \leq 2.40$. It is interesting to examine these Monte Carlo results in the $\lambda \rightarrow \infty$ and $\lambda \rightarrow 1$ limits where analytic solutions are available. The $\lambda \rightarrow \infty$ limit corresponds to mean field theory [61], while the $\lambda \rightarrow 1$ limit corresponds to the adhesive or sticky sphere model [38, 39]. Both of these limiting theories depend on one parameter only. We will show how the general two parameter ($\hat{\epsilon}$ and λ) square-well potential reduces to these different one parameter models. We show that we can recover the well-known mean field results [3, 43] and we determine the domain of λ in which mean field theory becomes a valid approximation. For the $\lambda \rightarrow 1$ limit, the Monte Carlo calculation provides us with important estimates of the critical parameters of the adhesive sphere model, which have been the subject of theoretical uncertainty [39].

The connection between the two parameter square-well potential and the two limiting theories is most readily seen by considering the second virial coefficient $B_2(T)$ of the square-well potential [62], where

$$\begin{aligned}
 B_2 &= -\frac{3}{\pi\sigma^3} \int_0^\infty \{\exp[-u(r)/kT] - 1\} d^3r \\
 &= \underbrace{1}_I - \underbrace{4[\exp(\hat{\epsilon}) - 1](\lambda^3 - 1)}_{II}.
 \end{aligned}
 \tag{2.23}$$

Here term (I) comes from the hard sphere core, while term (II) is the contribution of the attractive square-well.

Recall that the virial coefficient is defined through the equation of state [63] by

$$\frac{\pi_p \Omega_p}{kT} = \phi + \sum_{i=2}^{\infty} B_i \phi^i, \quad (2.24)$$

where π_p is the osmotic pressure of the solution and B_i is the i^{th} virial coefficient. For the interaction between the proteins to provide a physically reasonable equation of state, the attraction term (II) in B_2 (see Eq. (2.23)) must be finite. In the mean field case, this requirement implies that as $\lambda \rightarrow \infty$, we must have $\hat{\epsilon} \rightarrow 0$ (see Eq. (2.23)). Analogously, for the adhesive sphere model, we must take $\hat{\epsilon} \rightarrow \infty$ as $\lambda \rightarrow 1$. Therefore, the second virial coefficients for the two limiting theories are

$$B_2^{mf} = 4 - 4\hat{\epsilon}\lambda^3, \quad \text{for mean field } (\lambda \rightarrow \infty, \hat{\epsilon} \rightarrow 0) \quad (2.25)$$

$$B_2^{ad} = 4 - 12(\lambda - 1)\exp(\hat{\epsilon}), \quad \text{for adhesive spheres } (\lambda \rightarrow 1, \hat{\epsilon} \rightarrow \infty). \quad (2.26)$$

Eqs. (2.25) and (2.26) provide relationships between $\hat{\epsilon}$ and λ for the two limiting theories. Thus, if we define the quantities

$$a = 4\hat{\epsilon}\lambda^3 \quad (2.27)$$

and

$$\tau = \frac{1}{12(\lambda - 1)\exp(\hat{\epsilon})}, \quad (2.28)$$

we see that a is the single parameter which characterizes the mean field theory, while τ is the single parameter which describes the adhesive sphere model. Before we begin the analysis of our results within the context of the mean field and adhesive sphere models, we discuss these two limiting theories more fully.

In the mean field limit a is the well known van der Waals term [61] and is a measure of the strength of the attraction between particles for a long range interaction potential. Explicitly, a is the energy of interaction per particle times the volume over

2.3. RESULTS AND DISCUSSION

which the interaction is felt (in dimensionless units) [36]. Note that in the mean field limit the chemical potential may be written as [3, 43]

$$\hat{\mu}_{mf} = \hat{\mu}_{HS} - 2a\phi. \quad (2.29)$$

Here $\hat{\mu}_{HS}$ is the chemical potential of an assembly of hard spheres. In Refs. [3, 43] the Carnahan-Starling approximation for this term ($\hat{\mu}_{CS}$, defined in Eq. (2.18)) was used to obtain the critical parameters analytically. Our $\lambda = \infty$ entries in Table 2.1 we obtained by calculating $\hat{\mu}_{HS}$ from a simulation of hard spheres. We then used Eq. (2.29) to obtain the results shown in Table 2.1. These agree with those found analytically in the mean field limit [3, 43]. We also find $a(T_c) = 10.6$ as predicted theoretically.

In this short range regime, it is the parameter τ , given in Eq. (2.28), which characterizes the potential. The parameter τ was introduced by Baxter [38] and it is a measure of the stickiness of the adhesive spheres. It may also be thought of as a dimensionless measure of the temperature of the systems, being zero at low temperatures and larger at high temperatures [64]. Here the hard sphere model is regained in the limit of $\tau \rightarrow \infty$: in this limit $B_2^{ad}=4$ (see Eqs. (2.26) and (2.28)), which is the hard sphere result.

We begin our analysis by considering the predictions of these limiting theories for the reduced critical energy, $\hat{\epsilon}_c \equiv \epsilon/kT_c$. In Fig. 2-6 we show our simulation results (solid points) for the reduced critical energy as a function of $\ln(\lambda - 1)$. We see that reduced critical energy increases as the range of the potential decreases. The mean field result for $\hat{\epsilon}_c$ is

$$\hat{\epsilon}_c = \frac{a_c}{4\lambda^3}. \quad (2.30)$$

This last equation is derived from Eq. (2.27) and a_c is the value of the parameter a at the critical point. The dashed line in Fig. 2-6 represents Eq. (2.30) with $a_c = 10.6$, a value determined analytically [3, 43]. We see that mean field theory gives a good estimate for $\hat{\epsilon}_c$ for reduced ranges greater than $\lambda \sim 1.10$ (i.e., $\ln(\lambda - 1) \sim -2.30$).

As λ tends to unity, we may compare our findings with the adhesive sphere results given by Eq. (2.28). At the critical point $\tau = \tau_c$ and the relation between $\hat{\epsilon}_c$ and λ is

$$\hat{\epsilon}_c = -\ln[12\tau_c(\lambda - 1)]. \quad (2.31)$$

Thus, using our short-range results for $\hat{\epsilon}_c$ (shown in Fig. 2-6), we may estimate a value for τ_c . The numerical value of τ_c is not well-established. Watts *et al.* [39] obtained τ_c by using the Percus-Yevick [37] equation to find an analytic solution for the equation of state [38] of adhesive spheres. They undertook a calculation of the equation of state in three distinct ways: through the pressure, compressibility, and energy equations. The pressure equation gives unphysical solutions while the other two equations predict different values for τ_c : 0.0976 (from the compressibility equation) and 0.1185 (from the energy equation). Neither of these two values is consistent with our results for $\hat{\epsilon}_c$ as λ approaches unity. For our Monte Carlo results to asymptotically approach the theoretical predictions, we require $\tau_c > 0.125$. We observe that this lower bound for τ_c is larger than either of the two previous estimates made by Watts *et al.* [39], although the energy equation result $\tau_c=0.1185$ is closer to the lower bound we find. The solid line in Fig. 2-6 shows a comparison of our results with Eq. (2.31) for $\tau_c = 0.13$.

We may also examine the behavior of the critical volume fraction in the context of the two limiting theories. In Fig. 2-5 we show the mean field result for ϕ_c as the horizontal dashed line: $\phi_c=0.130$ [3, 43]. We note that for the mean field theory to give an accurate result for ϕ_c , the reduced range λ must exceed 1.65.

To compare the Monte Carlo findings with the adhesive sphere model, we extrapolate our results for the critical volume fraction to $\lambda=1$. If we perform a linear extrapolation (solid line), we find that $\phi_c(\lambda = 1)$ is 0.266 ± 0.009 . The uncertainty reported in this quantity represents only the statistical errors of our data and does not include any systematic errors. The critical volume fractions predicted by Watts *et al.* [39] are: $\phi_c = 0.121$ (from the compressibility equation) and $\phi_c = 0.320$ (from the energy equation). The large uncertainty in the theoretical result is due to the

increasing flatness of the chemical potential near the critical point as $\lambda \rightarrow 1$. This makes the critical volume fraction very sensitive to the approximations made in the Percus-Yevick scheme. However, as with the results for τ_c , we note that the energy equation prediction is closer to our value for ϕ_c than the prediction from the compressibility equation. We believe that our result, $\phi_c(\lambda = 1) = 0.266 \pm 0.009$, represents a reliable estimate of the critical volume fraction for adhesive spheres. This value of the critical volume fraction may prove useful as a benchmark for further investigations of systems with short range interactions.

Another quantity we may examine to elucidate the connection between the limiting theories and the Monte Carlo results is $\bar{\eta}$, the average number of contacts per particle. We calculate the dependence of $\bar{\eta}$ on volume fraction during the Monte Carlo simulations. This important quantity is the key ingredient in our extrapolation formula (Eq. (2.22)). The low ϕ behavior of $\bar{\eta}$ may be examined theoretically. As is shown in Appendix E, the result is [65]

$$\bar{\eta} = 8(\lambda^3 - 1) \exp(\hat{\epsilon})\phi. \quad (2.32)$$

Equation (2.32) is derived using the Boltzmann distribution and assuming that the particles interact independently. The assumption of independent interactions is also a fundamental postulate in mean field theory and it is justified when the number of possible contacts is large i.e., $\lambda \rightarrow \infty$. Thus, Eq. (2.32) should hold for all ϕ in the mean field limit and where it reduces to $\bar{\eta}_{mf} = 8\lambda^3\phi$.

In Fig. 2-7, we show the average number of contacts per particle $\bar{\eta}$ (open symbols), as a function of ϕ for several of the entries in Table 2.1: (i) $\lambda=1.05$ ($\hat{\epsilon}_1=2.650$, triangles); (ii) $\lambda=1.25$ ($\hat{\epsilon}_1=1.267$, squares); (iii) $\lambda=1.65$ ($\hat{\epsilon}_1=0.605$, bow ties); and (iv) $\lambda=2.20$ ($\hat{\epsilon}_1=0.260$, circles). Recall that $\hat{\epsilon}_1$ is the reduced energy at which the simulations are performed. The straight lines represent Eq. (2.32) with $\hat{\epsilon} = \hat{\epsilon}_1$ at each of the ranges listed above. As expected, Eq. (2.32) fits the Monte Carlo results very well at low ϕ for all values of λ . For a given range λ , the deviation of $\bar{\eta}$ from the direct proportionality to ϕ expressed in Eq. (2.32) is a measure of the departure from the

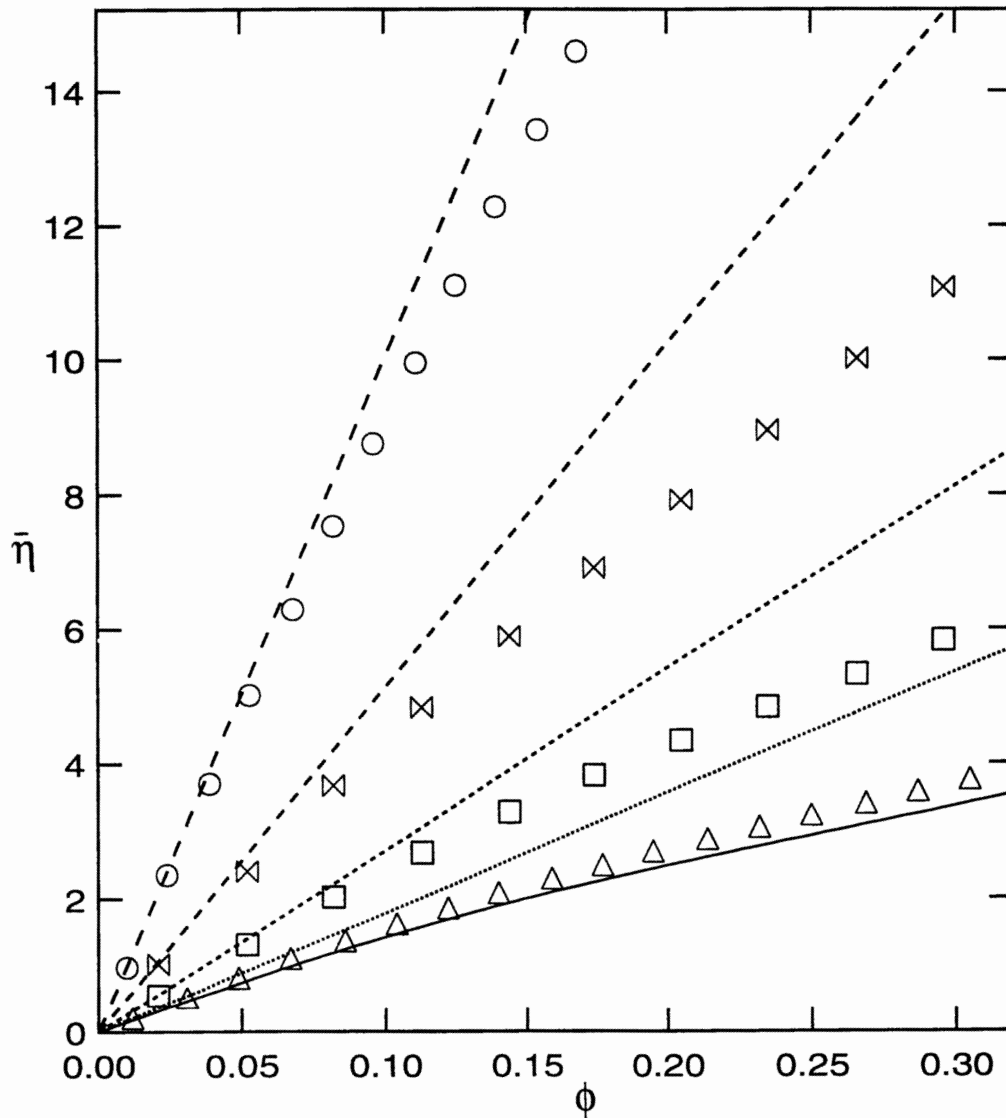


Figure 2-7: The average number of contacts per particle. The average number of contacts per particle, $\bar{\eta}$, is shown as a function of ϕ for several ranges: (i) $\lambda=1.05$, $\hat{\epsilon}=2.650$ (triangles); (ii) $\lambda=1.25$, $\hat{\epsilon}=1.267$ (squares); (iii) $\lambda=1.65$, $\hat{\epsilon}=0.605$ (bow ties); (iv) $\lambda=2.20$, $\hat{\epsilon}=0.260$ (circles). The straight dashed lines represent the low ϕ behavior for the different ranges as given by Eq. (2.32). The solid line is the adhesive sphere result for $\tau = 0.13$.

mean field limit. The solid line in the figure is the analytic result for $\bar{\eta}$ in the adhesive sphere limit [66] with $\tau = 0.13$. At low ϕ , this full expression reduces to $\bar{\eta} = 2\phi/\tau$, a result which can be obtained directly from Eq. (2.32). As we see, the average number of contacts per particle provides direct physical insight into the protein interactions. We will return to it in the next section.

2.3.4 Comparison with Experimental Data for the γ -crystallin Proteins

In this section we compare the coexistence curves generated by the Monte Carlo simulation with the experimentally measured ones [2]. In Fig. 2-8 we present data points of the reduced coexistence curves (T/T_c vs. ϕ) for γ_{IIIa} (circles), γ_{IIIb} (squares), γ_{II} (triangles), and γ_{IVa} (bow ties). The experimentally observed value of ϕ_c is 0.21 ± 0.02 for all the γ -crystallins. From Fig. 2-5, this corresponds to a range of approximately $\lambda=1.25$. Thus, we also show our Monte Carlo results for the coexistence curves at $\lambda=1.25$ (coarse dashed line).

For comparison, we also present the mean field coexistence curve as obtained analytically (fine dashed line). Recall that in the mean field case $\phi_c = 0.130$. Thus one can understand the experimentally observed value of $\phi_c = 0.21$ as arising from the short range character of the interaction potential. In addition, the Monte Carlo results for $\lambda = 1.25$ predict a coexistence curve which is twice as broad as the one obtained by mean field theory. Nevertheless, the predicted width is still about half that found experimentally. Even if we allow for uncertainty in the value of the critical volume fraction, and hence consider smaller values of λ , we still find that the curves generated by Monte Carlo simulation are significantly narrower than the experimental coexistence curves.

The question of the extra width of the coexistence curves notwithstanding, it is important to note that the agreement between the Monte Carlo simulations and the experiment results worsens as λ increases. Therefore, it is safe to conclude that the range of the protein-protein interaction is no greater than $\lambda=1.25$. Since the γ -

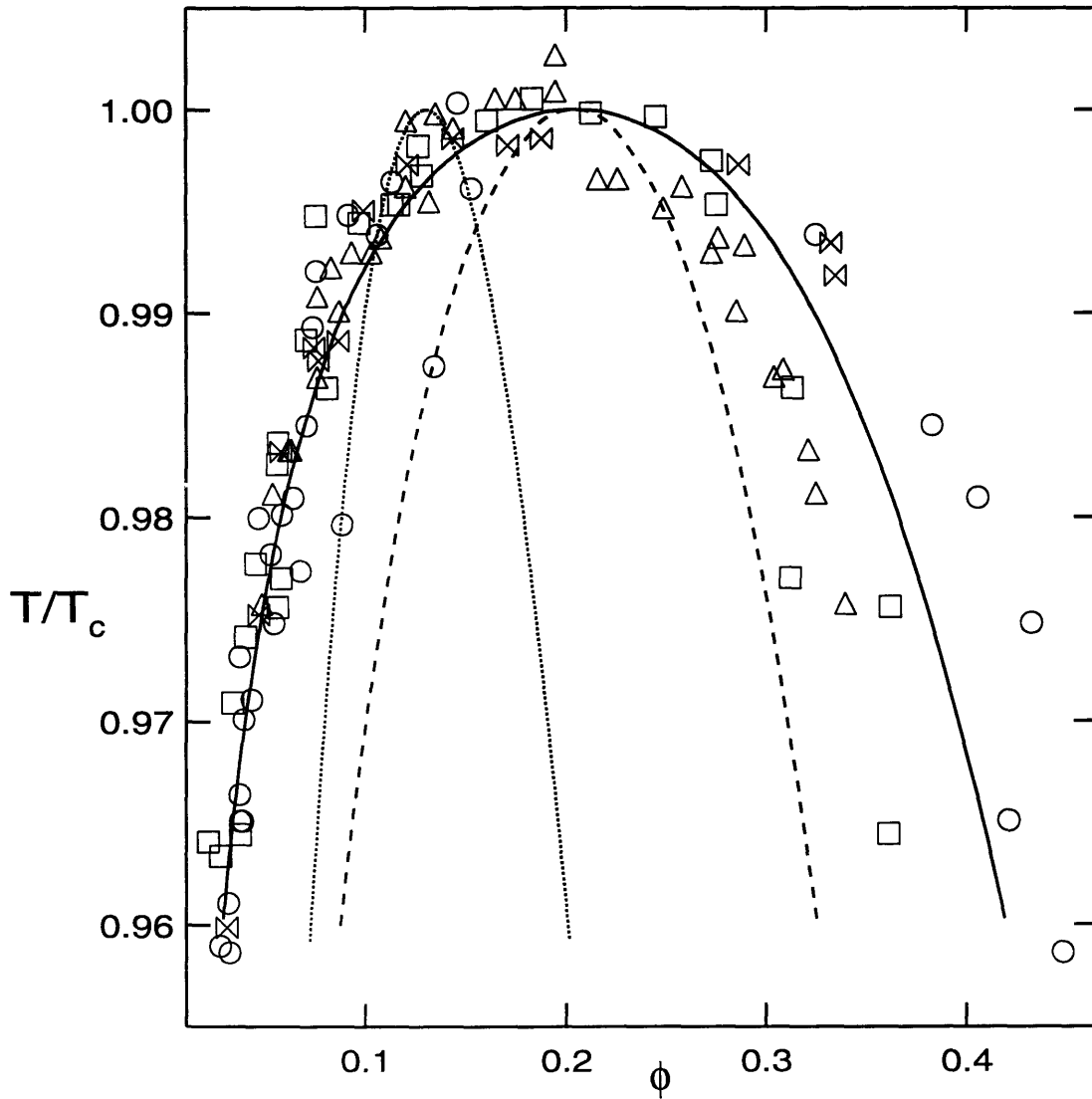


Figure 2-8: Comparison with the experimental results for the γ -crystallins. The coexistence curve generated by the Monte Carlo simulation for $\lambda=1.25$ is shown as a coarse dashed line. The fine dashed line represents the coexistence curve obtained analytically in the mean field limit. The experimental results of Broide *et al.* [2] are presented for γ_{IIIa} (circles), γ_{IIIb} (squares), γ_{II} (triangles), and γ_{IVa} (bow ties). The solid line is the coexistence curve obtained for $\lambda=1.25$ and a temperature dependent interaction energy of the form $\epsilon = kT_c \hat{\epsilon}_c (1 + \frac{T-T_c}{T_c} \kappa)$, with $\kappa = -3$.

crystallins are typically 48Å in diameter, this sets an upper limit of 12Å on the width of the attractive well of the γ -crystallins. This is consistent with the conclusions drawn from structure factor measurements [67]. The information contained in Figs. 2-3, 2-5, and 2-8 shows that λ must exceed 1.65 for mean field theory to provide a satisfactory description of the protein-protein interactions. Our observation above that $\lambda \leq 1.25$ for the γ -crystallins implies that a mean field model is inadequate for an accurate representation of γ -crystallin phase separation.

We now examine possible explanations for the width of the experimentally determined coexistence curves. Since the interaction between the proteins is in fact mediated by the surrounding water, we may consider the energy of interaction between the proteins to be temperature dependent. For example,

$$\epsilon(T) = kT_c \hat{\epsilon}(T_c) \left(1 + \frac{T - T_c}{T_c} \kappa\right), \quad (2.33)$$

where κ is a constant. In Fig. 2-8, the coexistence curve for $\lambda = 1.25$ with $\kappa = -3$ is shown as a solid line. This curve obviously gives a better fit to the data than the temperature independent ($\kappa = 0$) case. It was calculated as follows. In our simulations we obtain the coexisting volume fractions ϕ^I and ϕ^{II} as a function of $\hat{\epsilon} \equiv \epsilon(T)/kT$. We may use Eq. (2.33) to express the reduced temperature T/T_c in terms of $\hat{\epsilon}$, as

$$\frac{T}{T_c} = \frac{1 - \kappa}{(\hat{\epsilon}/\hat{\epsilon}_c) - \kappa}. \quad (2.34)$$

This equation allows use to convert our values of $\hat{\epsilon}$ to reduced temperatures T/T_c for any value of κ . Note that for $\kappa=0$, i.e., a temperature independent energy, we have $T/T_c = (\hat{\epsilon}_c/\hat{\epsilon})$.

We see from Fig. 2-8 that the temperature dependence of the interaction energy in no way affects ϕ_c , but it does increase the width of the coexistence curve. It remains to be seen whether such a strong temperature dependence of $\epsilon(T)$ is physically reasonable.

It is also possible that the observed extra width of the coexistence curve could

result from an anisotropic character of the protein interaction energy. We note from our simulations that, as the range of the interaction decreases, $\bar{\eta}_c$, the average number of contacts per particle at the critical point, decreases (see Table 2.1) while the width of the coexistence curves increases (see Fig. 2-3). From this perspective the large widths observed experimentally correspond to an effective number of contacts which is even smaller than that found for short-range isotropic interactions. Such a decrease in the number of contacts will occur if the true potential is both short-ranged and anisotropic.

From these considerations an interesting point emerges. We can see from the results presented in Table 2.1 that for a system with short range attractions ($\lambda \rightarrow 1$) each particle already makes only about three contacts at the critical point. An anisotropic potential i.e., one for which the attraction between proteins depends on their relative orientation, will cause the average number of contacts per particle to drop even further and may change the interactions of particles to the point where phase separation is replaced by reversible aggregation. The experimental results presented in this section lead us to believe that the proteins we study could be in fact close to this boundary. Work is currently underway to explore the role of anisotropy in the relationship between phase separation and reversible aggregation.

2.4 Summary and Conclusions

We have studied the binary liquid phase separation of aqueous protein solutions by modeling the protein interactions with a square-well potential. We utilize this potential in a hybrid Monte Carlo method which blends simulations with thermodynamic extrapolation techniques. In this method, we use the results of Monte Carlo simulations along a single isotherm to construct an analytic form of the chemical potential for a series of isotherms above and below the critical temperature. This unorthodox Monte Carlo scheme permits us, by the economy of its design, to reconstruct the phase diagram of systems over a wide domain of the reduced range of attraction λ . In particular, we have thoroughly explored potentials in the short range regime, with

2.4. SUMMARY AND CONCLUSIONS

ranges as small as $\lambda = 1.05$. These potentials are especially important for they apply to many colloidal suspensions, including the γ -crystallin protein solutions which we have previously investigated experimentally.

Our results provide insight into the central role played by the range of the interaction in determining the shape and location of the phase boundaries. Indeed, we have found that as the range decreases, the width of the coexistence curve increases and the critical volume fraction shifts to higher values. As part of our analysis, we have demonstrated how the two parameter square-well model reduces to the one parameter mean field model as $\lambda \rightarrow \infty$ and to the one parameter adhesive sphere model as $\lambda \rightarrow 1$. In the mean field limit, we recover the analytic result for the critical volume fraction, $\phi_c(\lambda \rightarrow \infty) = 0.13$. In fact, we find that the mean field model is a valid approximation provided that $\lambda \geq 1.65$. On the other hand, by examining our short range results, we are able to propose a value for the critical volume fraction in the adhesive sphere limit, $\phi_c(\lambda = 1) = 0.266 \pm 0.009$. We have also obtained an estimate for the critical value of the Baxter parameter $\tau_c \approx 0.13$. In view of the uncertainty in previous analytical findings, we believe that our results will be useful benchmarks for future theoretical and experimental studies of the adhesive sphere system.

For the γ -crystallins, we have experimentally observed a critical volume fraction of $\phi_c \approx 0.21$ and very broad coexistence curves. These facts imply that $\lambda \leq 1.25$, that is the width of the attractive well of these proteins is no greater than one quarter of their diameter. Thus, we conclude that the interactions between the proteins fall into the short range regime and cannot accurately be described by a mean field theory. Although our simulation results for the critical volume fraction of short range systems are in agreement with the experimentally observed value, the calculated width of the coexistence curve is still significantly smaller than that found experimentally. We have shown that the extra width of the experimental curves may be explained phenomenologically by employing a temperature dependent depth of the attractive well. However, another possibility is that this additional width may be due to anisotropy in the interaction potential. Such anisotropic interactions are to be expected in protein solutions. The calculational simplicity of our hybrid Monte Carlo method should

CHAPTER 2. THE LIQUID PHASE

facilitate a systematic examination of the effects of such anisotropic interactions.

2.4. SUMMARY AND CONCLUSIONS

Chapter 3

The Solid Phase

Portions of this chapter have already been published in Ref. [19]: N. Asherie, A. Lomakin, and G. B. Benedek. "Phase Diagram of Colloidal Solutions," *Phys. Rev. Lett.* **77**, 4832 (1996).

3.1 Introduction

The phase diagrams of colloidal solutions have been studied for over a century not only because of their great theoretical interest, but also for the many industrial applications of colloids [68]. The most commonly observed phase transition in colloidal solutions is solidification. Upon a change in temperature (or other external condition) the colloidal particles form a condensed phase which may have a regular structure (crystals) or be amorphous (aggregates). A less frequent transition is liquid-liquid phase separation (coacervation). Here the colloidal solution forms two distinct liquid phases: one colloid-rich, the other colloid-poor [69].

These transitions have analogous counterparts in simple molecular fluids. The solidification of colloids is equivalent to the fluid-solid transition while colloidal liquid-liquid phase separation corresponds to gas-liquid coexistence. There is however one striking difference: when most simple fluids are cooled the order of phases observed

3.1. INTRODUCTION

is gas to liquid to solid. In colloidal solutions the colloidal “gas” usually transforms into a solid without passing through a liquid phase.

We illustrate this anomalous order of phases by presenting in Fig. 3-1 the phase diagram of γ_{II} -crystallin protein [2, 3, 4]. This phase diagram is typical of the γ -crystallins (a family of monomeric eye lens proteins) and of other small globular proteins [5, 6, 7]. The circles are points which represent the volume fractions (ϕ) of coexisting protein-rich and protein-poor liquid phases (liquid-liquid coexistence curve). The squares (liquidus line) and the triangle (solidus line) respectively represent the volume fractions of protein in the liquid and solid phases in equilibrium with each other. We see that there is no triple point and the coexistence curve lies below the liquidus line. The γ -crystallins are an unusual colloidal system in that liquid-liquid phase separation may be observed despite it being metastable with respect to solidification. For comparison the phase diagram of argon is shown in Fig. 3-2 [8, 9, 10]. We see that for argon the critical point lies above the triple point (i.e., $T_c > T_t$) indicating the presence of a stable liquid phase. Although most colloidal phase diagrams resemble that shown in Fig. 3-1, a few do show phase behavior closer to that pictured in Fig. 3-2 [21].

Recently, evidence has accumulated that the interaction range plays a significant role in determining the structure of the phase diagram [70, 71]. In colloid-polymer mixtures it has been found that the shape of the phase diagram depends on the ratio of the radius of gyration of the polymer molecules to the radius of the colloidal particles [21, 72, 73]. Theoretical models [74, 75] of this system agree with the experimental observation that with very small polymers (i.e., very short ranges of attraction) there is no colloidal liquid phase [21]. Colloids are not the only system for which the connection between short-range interactions and the structure of the phase diagram has been noted. Fullerenes, macromolecules of carbon, also do not appear to exhibit a liquid phase upon cooling [76]. Simulations of hard spheres with an attractive Yukawa potential have been carried out to investigate the phase diagram of the fullerene C_{60} [48, 49]. It is found that for a sufficiently short-range potential the coexistence curve lies below the liquidus curve.

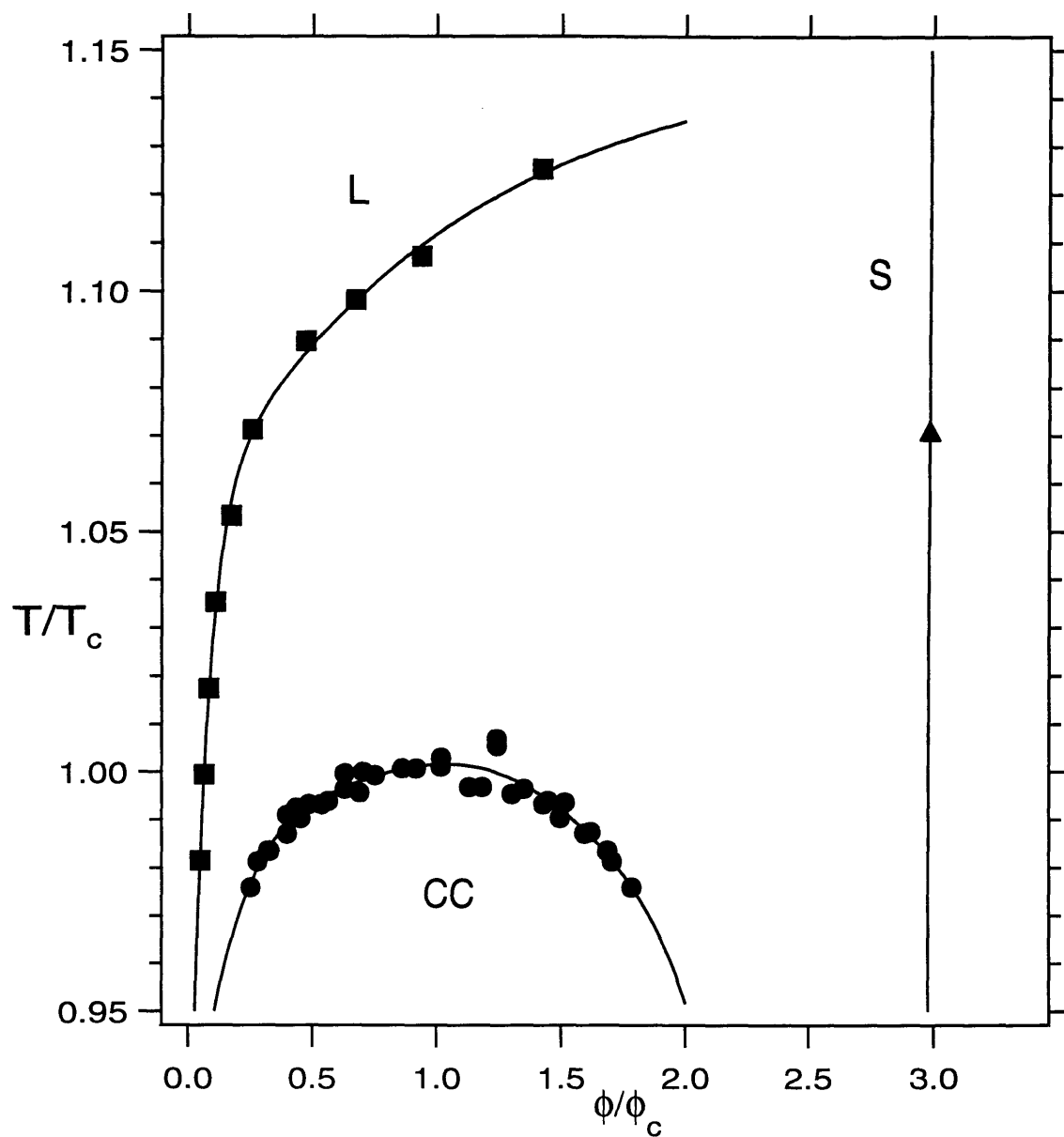


Figure 3-1: The phase diagram of γ_{II} -crystallin [2, 3, 4]. The circles are points on the liquid-liquid coexistence curve (CC). The squares are points on the liquidus line (L). The triangle is a point on the solidus line (S). The lines are guides to the eye. The critical temperature is $T_c = 278.4\text{K}$. The critical volume fraction is $\phi_c = 0.21$.

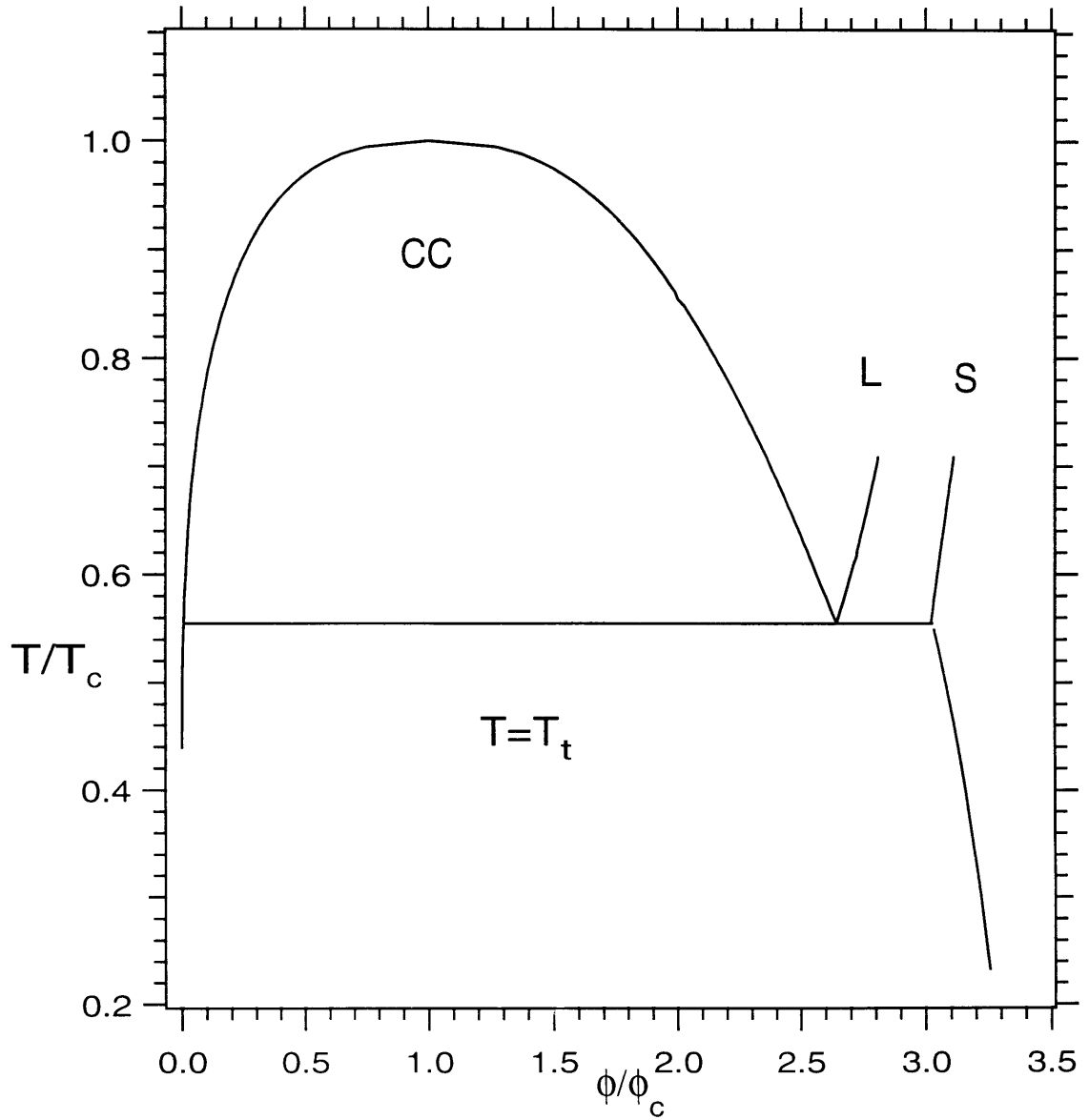


Figure 3-2: The phase diagram of argon [8, 9, 10] showing the coexistence curve (CC) and the liquidus (L) and solidus (S) lines. The critical temperature is $T_c = 150.86\text{K}$. The critical volume fraction is $\phi_c = 0.133$, assuming a hard core diameter $\sigma = 3.162\text{\AA}$ [12]. The triple point temperature is $T_t = 83.78\text{K} = 0.56T_c$.

In this chapter we present a general analysis which explains these individual experimental and theoretical findings. In the Monte Carlo work described in the previous chapter, we studied the liquid-liquid phase separation of globular particles with attractive interactions and obtained numerically the chemical potential of the *liquid* phase. Here we present an analytic expression for the chemical potential of a *solid* with short-range interactions. Our model of the solid incorporates the same essential features that were used to describe the liquid phase: the range λ of the interaction, the interaction energy ϵ and the number of contacts n_s made per particle. Knowing the chemical potentials of both phases we are able to demonstrate how the relative locations of the phase boundaries are related to λ and n_s .

3.2 The Cell Model

We use the Lennard-Jones and Devonshire cell model [77] to obtain an approximate analytic expression for the chemical potential of a solid with short-range interactions. As we did in Chapter 2 (see Eq. (2.2)), we will assume that the effective potential energy $u(r)$ for a pair of proteins (diameter σ) whose centers are separated by a distance r , is of the form of an attractive square-well with a hard core as given by Eq. (3.1) below.

$$u(r) = \begin{cases} +\infty, & \text{for } r < \sigma \\ -\epsilon, & \text{for } \sigma \leq r < \lambda\sigma \\ 0, & \text{for } r \geq \lambda\sigma \end{cases} \quad (3.1)$$

Here λ is the reduced range of the potential well and ϵ is its depth. When the interactions are short-ranged ($\lambda \rightarrow 1$) the thermodynamic properties become independent of the shape of the potential. We work with the square-well potential rather than other potentials [48, 49, 78] because it allows for an unambiguous definition of not only the range of interaction λ , but also of n_s , the number of contacts made per particle in the solid phase. Explicitly, n_s is the number of particles whose centers lie in the region $\sigma \leq r < \lambda\sigma$ from a given particle. As is shown in the Appendix F, the square-well

model produces a very simple form for the chemical potential of the solid μ_s , namely

$$\mu_s = \mu_0 - n_s(\epsilon/2) - kT \ln[(\lambda - 1)^3]. \quad (3.2)$$

The first term on the right-hand side of Eq. (3.2), μ_0 , is the standard part of the chemical potential. The second term is the total energy associated with each particle. The last term is the entropic contribution: the volume accessible to the center of mass of the particle is proportional to $(\lambda - 1)^3$. The proportionality factor, essentially the volume of a unit cell, has been absorbed into μ_0 . As we show in Appendix F, this factor can be calculated within the framework of the cell model [79]. The cell volume, however, is practically constant, reflecting the incompressibility of the solid phase. We may therefore choose a cell volume at zero pressure. The corresponding value of μ_0 is found to be identical to that in the expression $\mu = \mu_0 + kT \ln \phi$, which is appropriate for dilute solutions [80] (see Appendix F).

An important parameter in the chemical potential of the solid is n_s . The value of n_s is not known *a priori*, but determined by the structure of the solid. It is common practice to impose a particular crystal structure upon the solid when calculating the chemical potential by choosing the appropriate integral value for n_s . In the subsequent analysis, however, we will treat n_s as a continuous, phenomenological parameter so as to subsume within the cell model the actual structure and detailed interactions in the solid phase of real colloids.

3.3 Results and Discussion

3.3.1 Construction of the Phase Diagram

The liquidus line is obtained by equating the chemical potentials of the solid and the liquid. The reduced chemical potential of the liquid $\hat{\mu}_l$ was established as a numerical function of ϕ and $\hat{\epsilon}$ in Chapter 2. Here we have added the subscript l to the chemical potential of the liquid to distinguish it from the chemical potential of the solid μ_s . Note that for a protein-water solution $\mu_s \equiv \mu_p(s) - \gamma\mu_w(s)$, where $\mu_p(s)$ and $\mu_w(s)$ are

the chemical potentials of the protein and water in the solid phase. This is analogous to the representation we used for the liquid phase (see Sec. 2.2 and Appendix A). The corresponding reduced chemical potential of the solid is given by $\hat{\mu}_s \equiv [\mu_s - \mu_0]/kT$ and thus

$$\hat{\mu}_s = -n_s(\hat{\epsilon}/2) - 3 \ln[(\lambda - 1)]. \quad (3.3)$$

The liquidus line is then given by

$$\hat{\mu}_s(\hat{\epsilon}; n_s) = \hat{\mu}_l(\phi, \hat{\epsilon}). \quad (3.4)$$

For any solid with a given value of n_s , Eq. (3.4) allows us to construct the liquidus line, using Eq. (3.3) for $\hat{\mu}_s$ together with our previous calculations of $\hat{\mu}_l$. In Fig. 3-3 we show the liquidus lines for several values of n_s at $\lambda = 1.25$: $n_s = 12.0$ (A); $n_s = 11.6$ (B); $n_s = 11.5$ (C). These were constructed as follows. For any given reduced temperature T/T_c , the equivalent reduced energy $\hat{\epsilon} = \hat{\epsilon}_c/(T/T_c)$ was found (this last equation follows from $\hat{\epsilon} \equiv \epsilon/kT$; we take $\hat{\epsilon}_c = 1.269$ for $\lambda = 1.25$ (see Table 2.1)). The value of ϕ which satisfies Eq. (3.4) for that reduced energy $\hat{\epsilon}$ was then calculated. This procedure was repeated for temperatures in the range $1.17 \geq (T/T_c) \geq 0.94$ to generate the liquidus lines shown. The line (D) is the solidus line obtained by using the volume of a unit cell appropriate to a face-centered cubic solid ($n_s = 12$) at zero pressure, as is shown in Appendix F. The coexistence curve (E) is taken from our simulation work in Chapter 2 (see Fig. 2-3). As the value of n_s increases liquid-liquid coexistence becomes metastable with respect to solidification. The shape of the phase diagram for $n_s = 12.0$ (curves A, D, and E) has the same structure as the one found experimentally for the γ -crystallins (Fig. 3-1).

We may consider as a measure of the metastability of the liquid phase the "metastability gap" $(T_L - T_c)/T_c$, where T_c is the critical temperature and T_L is the temperature of the point on the liquidus line at the critical volume fraction ϕ_c . According to Eq. (3.4)

$$\hat{\mu}_s(\hat{\epsilon}_L; n_s) = \hat{\mu}_l(\phi_c, \hat{\epsilon}_L). \quad (3.5)$$

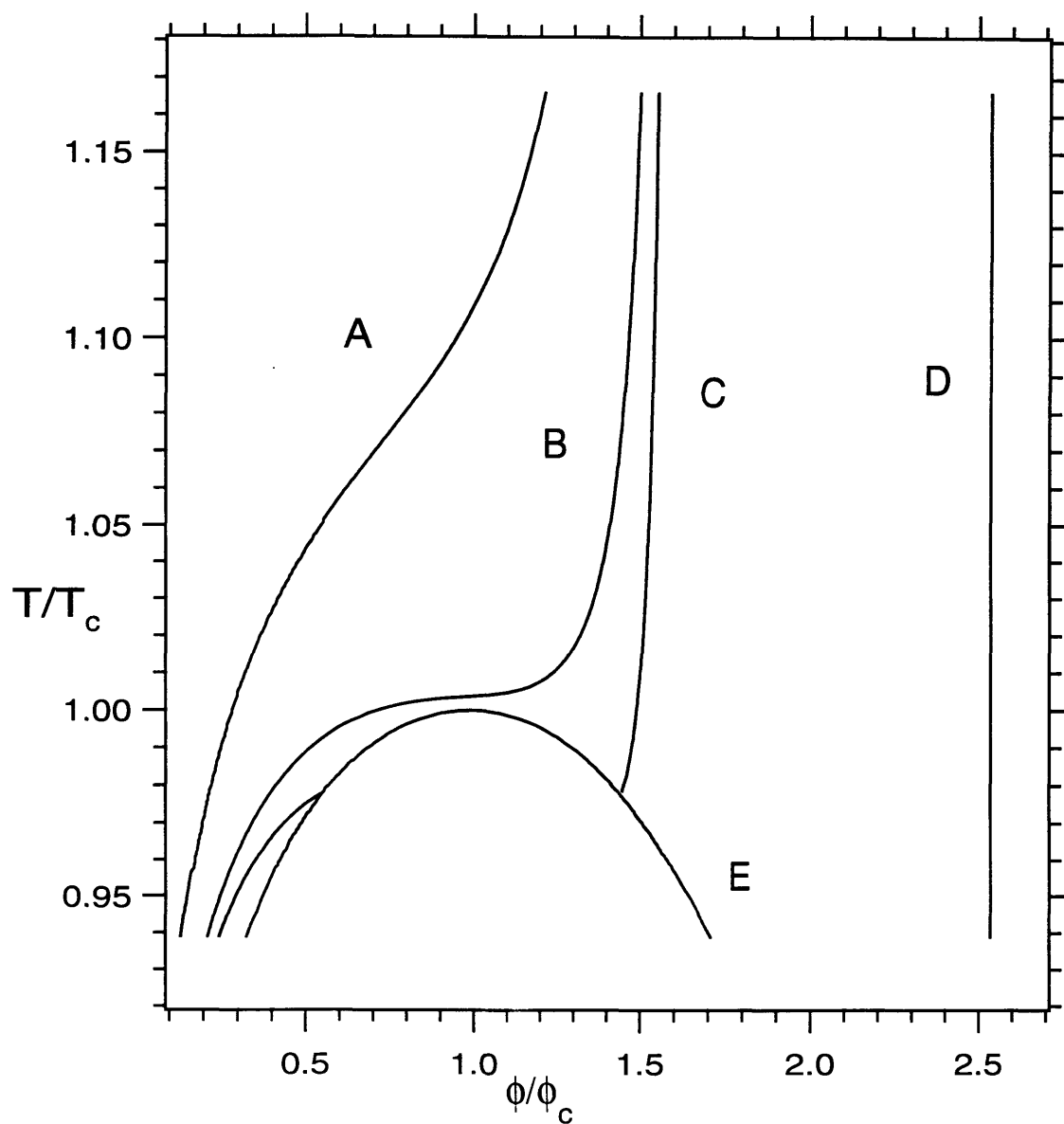


Figure 3-3: The phase diagram for a square-well system for $\lambda = 1.25$. The liquidus lines for three values of n_s are shown: $n_s = 12.0$ (A); $n_s = 11.6$ (B); $n_s = 11.5$ (C). The vertical line (D) is the solidus for $n_s = 12.0$. Curve E is the coexistence curve taken from Fig. 2-3.

with $\hat{\epsilon}_L = \epsilon/kT_L$. As we have shown earlier (see Eq. (2.22)), the chemical potential of the liquid at $\hat{\epsilon}_L$ may be expanded about the reduced critical energy $\hat{\epsilon}_c$

$$\hat{\mu}_l(\phi_c, \hat{\epsilon}_L) = \hat{\mu}_l(\phi_c, \hat{\epsilon}_c) - \frac{(\hat{\epsilon}_L - \hat{\epsilon}_c)}{2} n_l. \quad (3.6)$$

Here n_l is the change in the total number of contacts in the liquid phase upon the addition of an extra particle at the critical point ($n_l \equiv \frac{\partial}{\partial \phi}[\phi\bar{\eta}(\phi, T_c)]$, in the notation of Chapter 2.) Let us define n_s^* as the number of contacts in the solid at which $T_L = T_c$, i.e., the value of n_s for which the liquidus line touches the critical point. Thus n_s^* is given by $\hat{\mu}_s(\hat{\epsilon}_c; n_s^*) = \hat{\mu}_l(\phi_c, \hat{\epsilon}_c)$. Using this definition of n_s^* and substituting Eqs. (3.6) and (3.3) into Eq. (3.5) we obtain

$$\frac{T_L - T_c}{T_c} = \frac{n_s - n_s^*}{n_s^* - n_l} \quad (3.7)$$

From Eq. (3.7) we see that the metastability gap $(T_L - T_c)/T_c$ depends on the parameters n_s , n_s^* , and n_l . The first of these is determined by the structure of the solid. We list the other two in Table 3.1. In this Table we show for each value of λ , the corresponding values of n_s^* and n_l . Since n_s^* and n_l are evaluated at the critical point, these two parameters depend only on λ and not on ϵ . We note that for all of our short range simulations, the quantity $n_s^* - n_l$ is in the range 3.2 – 3.8. We also list the critical volume fraction ϕ_c and the reduced critical energy $\hat{\epsilon}_c$ at each range. Columns 3 – 5 are from the results of Chapter 2 (see Table 2.1), while n_s^* is obtained from a numerical solution of the equation $\hat{\mu}_s(\hat{\epsilon}_c; n_s^*) = \hat{\mu}_l(\phi_c, \hat{\epsilon}_c)$.

For the solid phase to be more stable than any coexisting liquid phases i.e., $T_L > T_c$, we require $n_s > n_s^*$. The maximum number of contacts for hard spheres with short-range interactions is 12. Therefore at any given range the solid will always be stable if $12 \geq n_s > n_s^*$. As the range decreases so does n_s^* and fewer contacts are necessary to form a stable solid. Thus, we expect that as the range decreases, there will be a greater variety of solid structures (characterized by different n_s) which will be more stable than the coexisting liquid phases. As a result, liquid-liquid phase separation will be less likely to be observed.

3.3. RESULTS AND DISCUSSION

Table 3.1: Metastability gap parameters at different reduced ranges λ (See Eq. (3.7)). The quantities presented are: (i) the average number of contacts per particle in the solid n_s^* at the metastability boundary ($T_L = T_c$); (ii) the change in the total number of contacts per particle in the liquid n_l upon the addition of an extra particle; (iii) the critical volume fraction ϕ_c ; (iv) the reduced critical energy $\hat{\epsilon}_c$.

λ	n_s^*	n_l	ϕ_c	$\hat{\epsilon}_c$
1.25	11.59	7.84	0.205	1.269
1.20	11.03	7.47	0.216	1.443
1.15	10.45	7.00	0.227	1.673
1.10	9.75	6.57	0.244	2.038
1.05	8.95	5.62	0.246	2.667

3.3.2 Comparison with Experimental Results

With our approach we may understand the experiments of Ilett *et al.* [21]. These authors study colloid-polymer mixtures with different ranges of interaction, and find at $\lambda \approx 1.25$ a transition from the type of phase diagram shown in Fig. 3-2 to that in Fig. 3-1 (they do not observe the metastable liquid-liquid coexistence). In our analysis this transition occurs when the metastability gap changes sign i.e., when $n_s^* = n_s$. In the solid phase the colloid-polymer system forms close-packed crystals i.e., $n_s = 12$ [81]. Thus from Table 3.1 we expect the transition to occur at $\lambda \approx 1.25$, the value observed experimentally. In their experiments only one range of interaction shorter than the crossover value of $\lambda \approx 1.25$ is studied: $\lambda = 1.08$ [82]. Taking $n_s = 12$, Table 3.1 implies that at this range the metastability gap of the system is so large that the coexistence curve lies outside the experimentally examined region. It would be interesting to search intermediate ranges i.e., $1.08 < \lambda < 1.25$, for metastable liquid-liquid coexistence and compare the values of the metastability gaps with our predictions.

In Chapter 2 we suggested that the large values of ϕ_c and the broad coexistence curves observed for the γ -crystallins imply that these proteins lie in the domain $\lambda \leq 1.25$ [18]. We now find that this is precisely the domain where liquid-liquid coexistence may be metastable, and in fact this metastability is observed. When plotted in reduced units the phase diagrams of each of the γ -crystallins have the same shape with almost identical values of ϕ_c and with approximately the same size metastability gaps (see Fig. 3-4). The interactions between these proteins may therefore be described by potentials with the same range. In addition, all the protein crystals should have the same value of n_s . For $(T_L - T_c)/T_c \approx 0.1$, the relation found experimentally for the γ -crystallins [3], Eq. (3.7) gives $n_s - n_s^* \approx 0.35$. It is because n_s differs little from n_s^* that we observe both liquid-liquid coexistence and solidification for the γ -crystallins.

Related observations are made by Broide *et al.* [5] who find that the metastability gap of lysozyme is unaffected by ionic strength but that it is larger (by 5 – 10°C) for needle-shaped crystals than for prism-shaped crystals. In our analysis these two

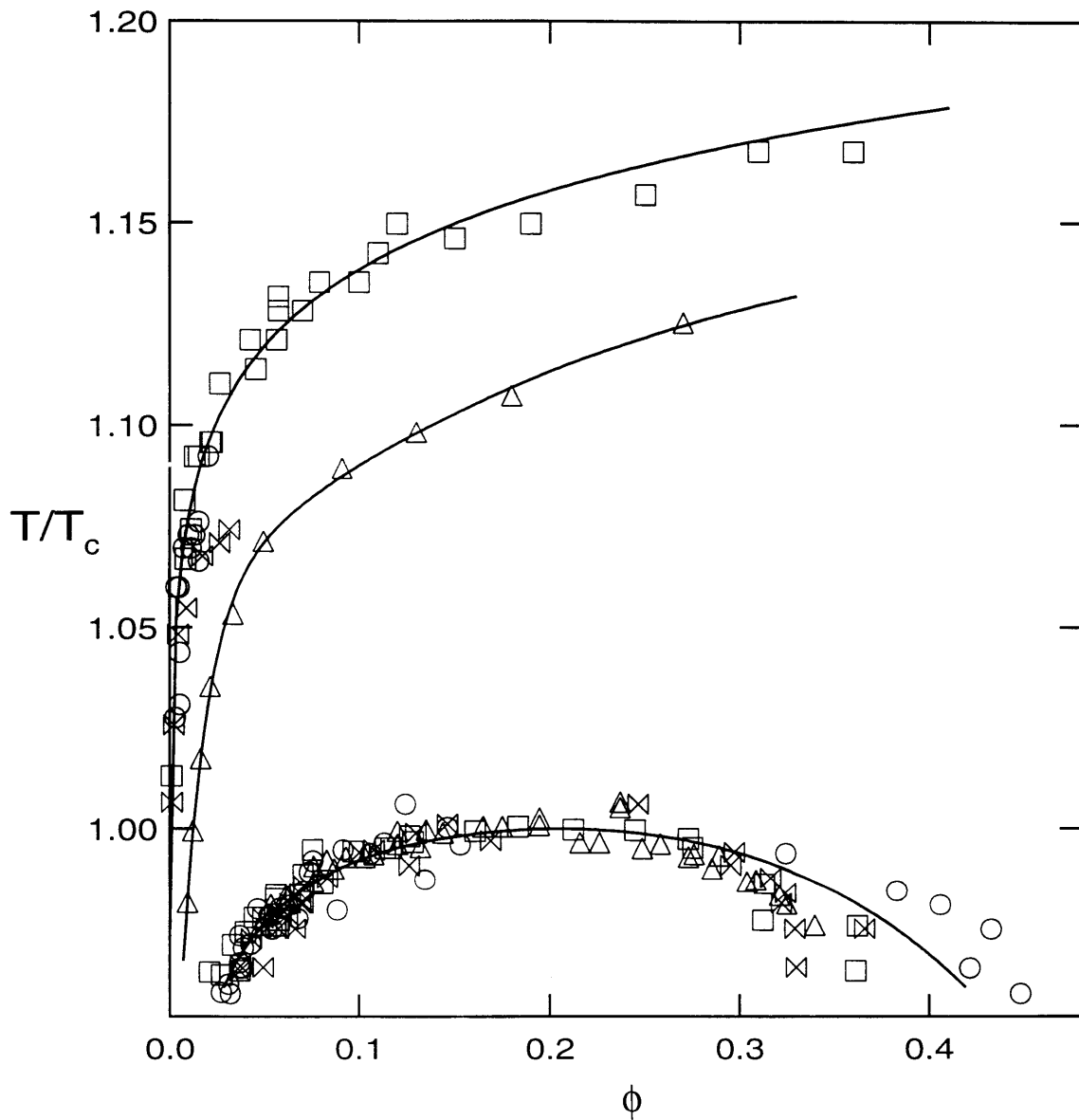


Figure 3-4: The experimentally determined phase diagram for the γ -crystallins from Refs. [2, 3]. We present the liquid-liquid coexistence curve and the liquidus line for γ_{IIIa} (circles), γ_{IIIb} (squares), γ_{II} (triangles) and γ_{IV} (bow ties). The solid lines are guides to the eye.

types of crystals are predicted to have values of $n_s - n_s^*$ differing by approximately 0.05 – 0.1.

The difference seen in Fig. 3-4 between the liquidus lines of γ_{II} (triangles) and γ_{IIIb} (squares) may similarly be a reflection of the different crystal structures of these proteins. However, another factor may be causing this difference. It is known that γ_{II} undergoes rapid aggregation while γ_{IIIb} does not (see Chapter 4). For the measurement of the coexistence curve, this aggregation was suppressed by adding dithiothreitol to the solution of γ_{II} [83]. Such precautions were not taken for the measurement of the γ_{II} liquidus line [3]. It would be necessary to repeat the γ_{II} liquidus line measurements under conditions where the aggregation is suppressed to make a proper comparison with the γ_{IIIb} results.

It has been argued that in the adhesive sphere limit ($\lambda \rightarrow 1, \hat{\epsilon} \rightarrow \infty$) there is no thermodynamically stable liquid phase [84]. Our analysis shows that for a close-packed system the absence of a stable liquid phase already begins when $\lambda < 1.25$.

3.4 Summary

We have presented a simple analytic form for the chemical potential of a short-range solid in terms of the physically significant parameters of the system: the interaction range λ , the interaction energy ϵ and the number of contacts made per particle n_s . We have shown how these parameters determine the order of phases found in the phase diagram. In particular, we have demonstrated that for a given n_s , a sufficiently short range of interaction leads to the metastability of liquid-liquid coexistence with respect to solidification. Conversely, experimental information on the relative order of the phase boundaries of colloidal solutions gives direct insight into the magnitude of the physically important parameters λ and n_s . The theory we have presented appears capable of explaining the phase diagrams observed in a wide variety of colloidal systems including globular protein solutions.

Chapter 4

Aggregation and Phase Separation

4.1 Introduction

In the previous two chapters we have seen how the range of interaction of globular proteins and the number of contacts made in the solid phase determine the positions of the phase boundaries observed. In particular, the fact that the liquid-liquid coexistence curve lies below the liquidus line is a reflection of the shortness of the range of interaction. There is another important phenomenon occurring in globular protein solutions which is related to the short-ranged nature of the protein interactions. This phenomenon is aggregation. Aggregation, unlike liquid-liquid or liquid-solid coexistence, is not a phase transition. Aggregates are amorphous solid structures which are not in thermodynamic equilibrium with the surrounding solution [85]. Aggregation is driven by the specific short-ranged bonding that is present between proteins. Almost all protein solutions are susceptible to aggregation.

In globular protein solutions, aggregation is in competition with liquid-liquid and liquid-solid phase separation. Since aggregates do not require the formation of translationally and rotationally ordered structures, they form more rapidly than crystals. Thus aggregation often preempts crystallization in a liquid-solid transition. Much effort has gone into investigating how to prevent aggregation and promote crystallization of globular proteins [82]. There has been, however, much less work done on the effect of aggregation on the liquid-liquid phase separation of globular proteins.

In this chapter we describe our study of the liquid-liquid coexistence curve of oligomers of globular proteins. For our experimental work we will use proteins from the γ -crystallin family. The γ -crystallins are a homologous family of monomeric proteins found in the mammalian eye lens. They have molecular weights of approximately 21 kDa and the average hydrodynamic radius for all the γ -crystallins is 2.4 nm [89]. As we stated in Sec. 2.1, the γ -crystallins may be divided into two groups: “high- T_c ” proteins, such as γ_{IIIa} (γ_C) and γ_{IVa} (γ_E), which exhibit high critical temperatures ($T_c \approx 38^\circ\text{C}$), and “low- T_c ” proteins, such as γ_{II} (γ_B) and γ_{IIIb} (γ_D), which exhibit low critical temperatures ($T_c \approx 5^\circ\text{C}$). The critical volume fractions of all the γ -crystallins are approximately the same $\phi_c = 0.21 \pm 0.02$ [2].

Phase separation and aggregation of the γ -crystallins are involved in the formation of cataracts [25]. Recently, Pande *et al.* [23] have observed that liquid-liquid phase separation temperature (T_{ph}) of a solution of γ_{II} increased with time when the solution was allowed to stand several weeks at room temperature. These authors argued that this increase in T_{ph} was due to the formation of a new protein species, γ_{III} . They isolated this species and concluded that it consisted of two types of dimers of γ_{II} -crystallin [86]: covalently crosslinked dimers, formed through the oxidation of thiol groups on the protein, and loosely associated dimers. This work stimulated us to develop a model system which could be used to understand how small aggregates affect the liquid-liquid phase separation of globular protein solutions, and in particular to help us interpret the results found for γ_{II} -crystallin. We believe that the oligomers that we have produced are such a system.

We decided to work with the protein γ_{IIIb} -crystallin for the following reasons: γ_{IIIb} is a low T_c protein like γ_{II} . These two proteins have almost identical liquid-liquid coexistence curves (see Figs. 2-8 and 4-7). Unlike γ_{II} , γ_{IIIb} does not form dimers spontaneously by oxidation under the same solution conditions [87]. This allows us to produce oligomers of γ_{IIIb} in a controlled fashion. Work in this laboratory has shown that the thiol groups of the γ -crystallins are susceptible to chemical modifications [25, 86, 88]. In particular, the monofunctional reagent N-ethylmaleimide has proved effective in modifying thiol groups [88]. We therefore chose a bifunctional

analog of this reagent, bismaleimidohexane, which has two maleimide groups and therefore can be used to form oligomers of the protein. The details of the crosslinking procedure are given in Secs. 4.2.2. Our measurements of the coexistence curves of the resulting oligomers are described in Sec. 4.3

4.2 Materials and Methods

4.2.1 Preparation of Pure γ_{IIIb} -Crystallin Solutions

The γ_{IIIb} crystallin used in our study was isolated from 1- to 6-week-old calf lenses, obtained by overnight express from Antech (Tyler, TX). The monomeric γ -crystallin fraction was isolated from the soluble protein fraction by size-exclusion chromatography on Sephadex G-75, as described in Thomson *et al.* [4]. Native γ -crystallin so obtained was further fractionated into γ_{I} , γ_{s} , γ_{II} , γ_{III} and γ_{IV} by cation-exchange chromatography on Sulfopropyl Sephadex C-50, according to Ref. [14]. Anion-exchange chromatography on Diethylaminoethyl (DEAE)-Sephadex was used to fractionate γ_{III} into γ_{IIIa} and γ_{IIIb} as described in Broide *et al.* [2]. Native γ_{III} consists of $\approx 40\%$ γ_{IIIb} and $\approx 60\%$ γ_{IIIa} by weight. Immediately after elution, the pure γ_{IIIb} fraction was transferred into 275 mM sodium acetate buffer, pH 4.8, to prevent possible oxidation of sulfhydryl groups. The purity of the protein samples was at least 98%, based on both cation-exchange High Performance Liquid Chromatography (CAT-HPLC) and size-exclusion High Performance Liquid Chromatography (SX-HPLC). These methods are described in Sec. 4.2.3.

The purified γ_{IIIb} fraction was dialyzed exhaustively into 100 mM sodium phosphate buffer (ionic strength 240 mM, pH 7.1), which contained sodium azide (0.02%). The concentrations of the pure γ_{IIIb} samples were determined by UV absorption at 280 nm, as described by Berland *et al.* [3], using the extinction coefficient $E_{280}^{0.1\%,1\text{cm}} = 2.11$. The protein volume fraction ϕ was obtained from the concentration C (in mg/ml) by using the specific volume \bar{v} with $\phi = \bar{v}C$. We assume that the specific volume of all the γ -crystallins are the same and use the value determined for γ_{II} , $\bar{v} = 0.71\text{cm}^3/\text{g}$ [31]

(we will use the above values of $E_{280}^{0.1\%,1cm}$ and \bar{v} for the oligomers as well). Precautions were taken to obtain aggregate-free and crystal-free solutions [2, 3, 89].

The next two subsections describe the production and characterization of the crosslinked oligomers.

4.2.2 Production of Crosslinked Oligomers

The oligomers of γ_{IIIb} -crystallin were produced by crosslinking the native monomer protein with the homobifunctional reagent bismaleimido-hexane (BMH, Pierce Chemical Company, Rockford, IL) [90, 91, 92]. This reagent is a member of a homologous series of the form $(\text{C}_4\text{H}_2\text{O}_2\text{N})_2(\text{CH}_2)_n$, with $n = 6$ for BMH [93]. The reagent is insoluble in water, but it has been added to aqueous protein solutions as the solid [94]. Since no significant side reactions occur, stoichiometric amounts can be added. For our reactions the mass of BMH used was such that the mole ratio of protein to BMH was unity. Typically 5 mls of γ_{IIIb} at 30–40 mg/ml (in phosphate buffer) were added to the appropriate amount of dry, solid BMH (about 2–3 mgs). The solution was left to react at room temperature while stirring constantly for 3 hours. Since BMH is insoluble in water the solution appeared slightly cloudy even at the beginning of the reaction. However, by the end of the reaction the cloudiness had increased, most probably due to the formation of large aggregates.

The reaction was not quenched, but at the end of the three hours the sample was centrifuged for 30 minutes at 10,000 rpm. The supernatant was removed and filtered (0.22 micron filter). The resulting solution was clear. The composition of this solution was determined by SX-HPLC on a Superdex 200HR column (see Sec. 4.2.3). Typically, the solution contained 30% monomers, 60% dimers and 10% higher oligomers (see Fig. 4-1). Since this monomer fraction may have been modified by the crosslinker (e.g., by an intramolecular crosslink) we will refer to this fraction as “monomer after crosslinking” to distinguish it from the native monomer protein. The individual oligomers were subsequently isolated by low pressure size-exclusion chromatography on a XK26/70 Superdex-75 column (Pharmacia Biotech, Piscataway, NJ) at a flow rate of 1ml/min, with 100 mM sodium phosphate buffer containing sodium azide

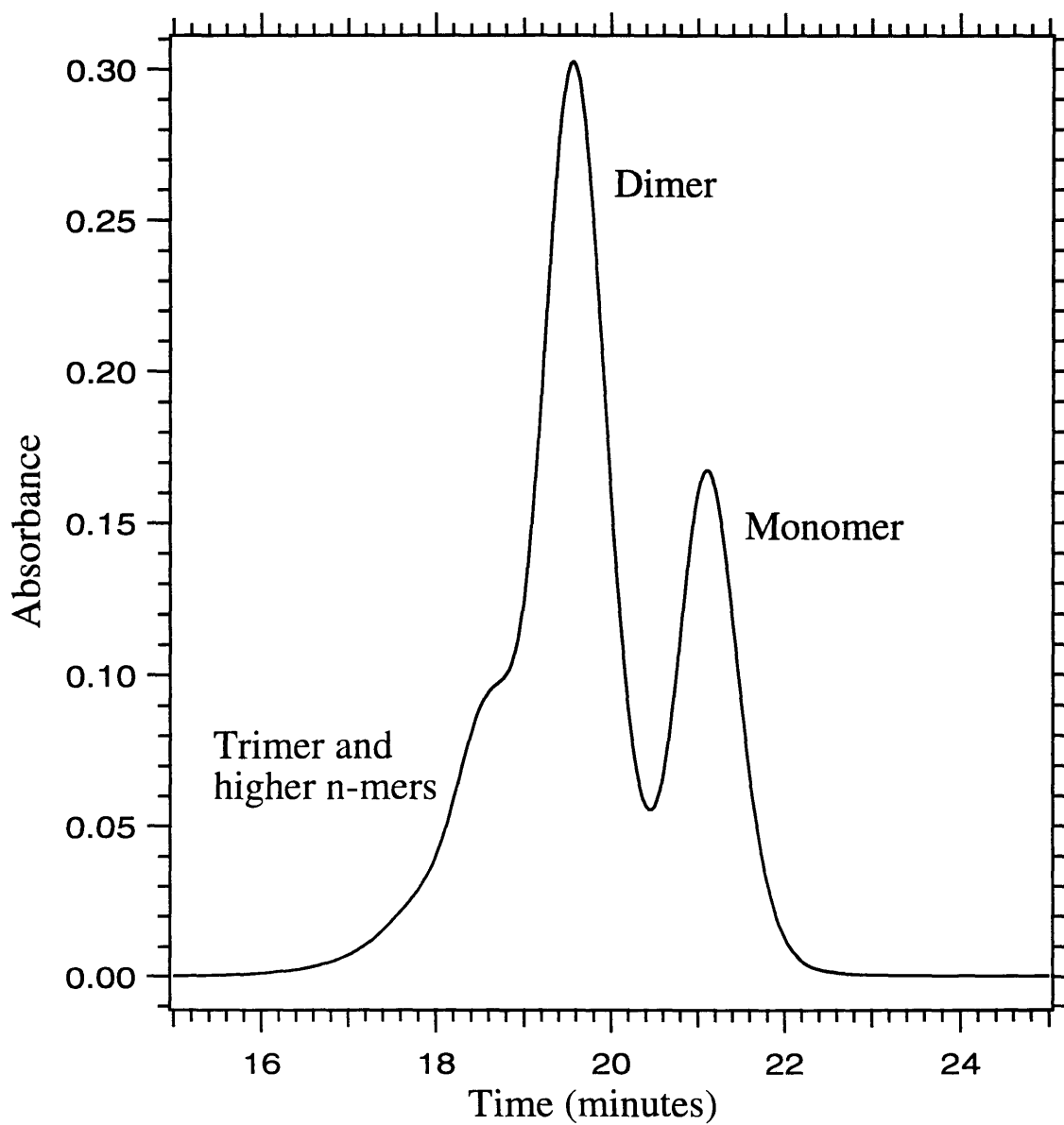


Figure 4-1: Size exclusion chromatography results (on a Superdex 200HR column) for the crosslinked protein after a three hour reaction. The absorbance at 280 nm is shown as function of retention time in minutes.

4.2. MATERIALS AND METHODS

(0.02%). This allowed for a larger scale separation, and with a better resolution, than was possible with the Superdex 200HR column. The final yields of the oligomers after the low pressure separation were approximately 15% monomer, 30% dimer and 3% trimer. These fractions were each at least 99% pure as determined by SX-HPLC. The good resolution of the low pressure column enabled us also to collect a mixture of higher n-mers (trimers, tetramers and pentamers).

The structure of BMH and the expected structure of the dimer are shown in Fig. 4-2. The predominance of the dimer over other oligomers is consistent with the three dimensional x-ray crystal structure of γ_{IIIb} [95]. This protein has five thiol groups, but only one of these has significant surface accessibility ($\sim 7\text{\AA}^2$). Thus oligomers greater than the dimer will be very slow to form. In the next section we describe the characterization of these oligomers.

4.2.3 Characterization of Crosslinked Oligomers

The oligomers were studied by gel electrophoresis (SDS/PAGE), SX-HPLC, Quasielastic Light Scattering (QLS) and CAT-HPLC.

Gel Electrophoresis

SDS/PAGE was carried out on 12% gels in the absence of urea with a Mini-Protean II electrophoresis system (Bio-Rad). Conditions were as described elsewhere [96, 97]. Gels were run with and without dithiothreitol (DTT). This reagent reduces disulphide links and has been shown to dissociate γ_{IIIH} completely into monomers [23]. The electrophoresis results displayed in Fig. 4-3 show that the major dimer (lanes 5 and 6) and trimer bands (lanes 8 and 9) have molecular weights of ≈ 41 kDa and ≈ 62 kDa respectively. This is consistent with the known native monomer molecular weight of ≈ 21 kDa [98]. Tetramers (≈ 85 kDa) and pentamers (≈ 97 kDa) were seen in the high n-mer fraction collected (lane 10). SDS/PAGE experiments following a four hour incubation with 100mM DTT (data not shown), gave results almost identical to those displayed in Fig. 4-3.

Note that, as expected, the control sample (lane 3) appears identical to the native

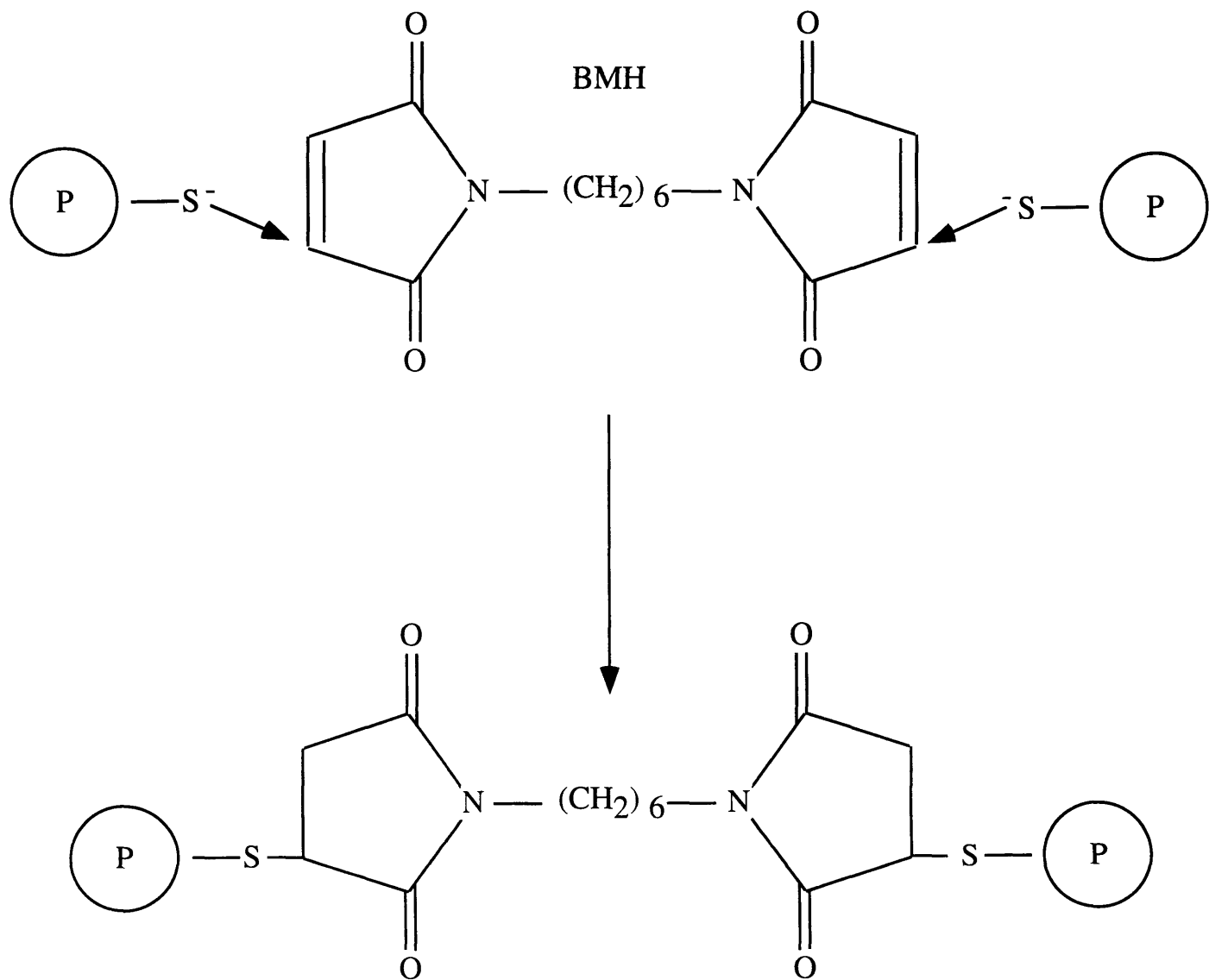


Figure 4-2: The expected reaction of bismaleimido-hexane (BMH) with thiol groups on a protein, schematically shown as P. The final product is the expected structure of the dimer.

4.2. MATERIALS AND METHODS

sample (lane 2). The control was treated exactly as the oligomer samples except that no crosslinker was added to the solution. The native sample was taken from the pure γ_{IIIb} fraction in 275 mM sodium acetate buffer at pH 4.8 (see Sec. 4.2.1).

Size-exclusion HPLC

SX-HPLC was carried out essentially according to Friberg *et al.* [86] with a Superdex 200HR column from Pharmacia Biotech (Piscataway, NJ); the buffer used was 100 mM sodium phosphate (pH 7.1) with 0.02% sodium azide. The results for the molecular weights of the various oligomers were consistent with those found by SDS/PAGE.

Quasielastic Light Scattering

QLS measurements were performed with a 144 channel Langley-Ford model 1097 correlator and a Coherent model Innova 90 Plus argon laser to determine the hydrodynamic radii (R_h) of the oligomers. To calculate the distribution of sizes of the scattering particles, we used the method described in Ref. [99], which was adapted for the analysis of the homodyne correlation function. The conditions of non-negativity and smoothness were superimposed on the size distribution to stabilize this otherwise ill-conditioned problem [100].

Measurements were made at (25°C) on the monomer, dimer, trimer and n-mer fractions. The sample concentrations were 3-5 mg/ml. The results are discussed in detail in Sec. 4.3.3.

Cation-exchange HPLC

CAT-HPLC was performed using a synchropak CM-300 column from SynChrom (Lafayette, IN) as described in Ref. [101]. The results indicated that each oligomer was predominantly (>70%) a single charged species (data not shown).

Our analysis shows that the oligomers of γ_{IIIb} we have formed consist of stable complexes whose size and weight are consistent with that of a γ_{IIIb} monomeric unit. We believe that we have successfully produced crosslinked oligomers of γ_{IIIb} .

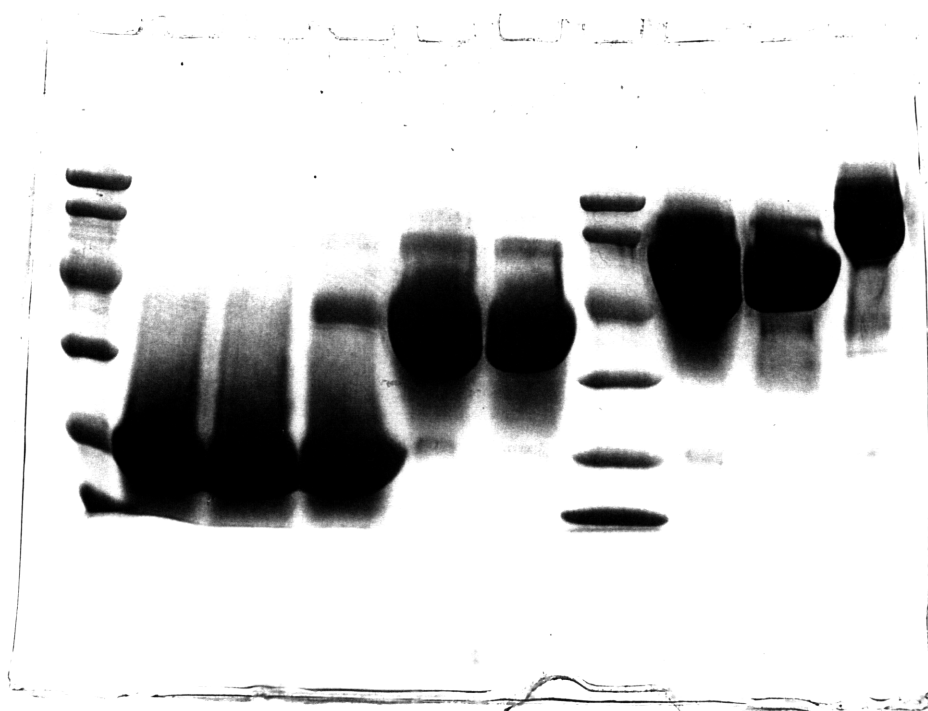


Figure 4-3: SDS/PAGE of the oligomers of γ_{IIB} without DTT. Lanes (left to right): 1 and 7, molecular mass markers (from bottom to top) are 14.4, 21.5, 31.0, 45.0, 66.2 and 97.4 kDa; 2, native γ_{IIB} ; 3, control γ_{IIB} ; 4, monomer fraction after crosslinking; 5 and 6, dimer fraction; 8 and 9, trimer fraction; 10, n-mer mixture.

4.2.4 T_{ph} measurements

We obtained the coexistence curve for the oligomers using the method described elsewhere [35]. We found that concentrating the oligomers with the Ultrafree-4 ultrafiltration devices (Millipore, Bedford, MA; molecular weight cutoffs of 5, 10 and 30 kDa) was easier than with the Centricon-10 and Centricon-30 (Amicon, Beverly, MA) used in previous work [2]. In this work we determined the complete coexistence curve for the γ_{IIIb} dimers. Only a partial coexistence curve could be obtained for the trimers due to the low yields of material. As has been reported previously for high T_c γ -crystallin monomers [2] and for γ_{III} [23], we found that our oligomer samples would often precipitate at high temperatures and high concentrations, making measurements difficult. The coexistence curve for the native monomer γ_{IIIb} has been published from this laboratory [2]. We also present partial coexistence curves for the monomer and n-mer fractions collected after the crosslinking reaction.

4.3 Results and Discussion

4.3.1 Coexistence Curves of Oligomers

In Fig. 4-4 we present the coexistence curves for oligomers of γ_{IIIb} -crystallin in 100mM sodium phosphate solution (pH 7.1). We show our results for the γ_{IIIb} monomers after crosslinking (solid bowties), the γ_{IIIb} dimers (solid circles) and the γ_{IIIb} trimers (solid triangles). The monomer γ_{IIIb} (taken from Ref. [2]) is also shown (open squares) together with our control sample for the monomer (open circles). This control sample was treated exactly as the oligomer samples except that no crosslinker was added to the solution. As can be seen in the figure, we find excellent agreement between our control sample and the results of Broide *et al.* [2].

The solid lines in the figure are fits to the function [102]

$$\frac{\phi^{II} - \phi^I}{\phi_c} = A\left(1 - \frac{T}{T_c}\right)^\beta \quad (4.1)$$

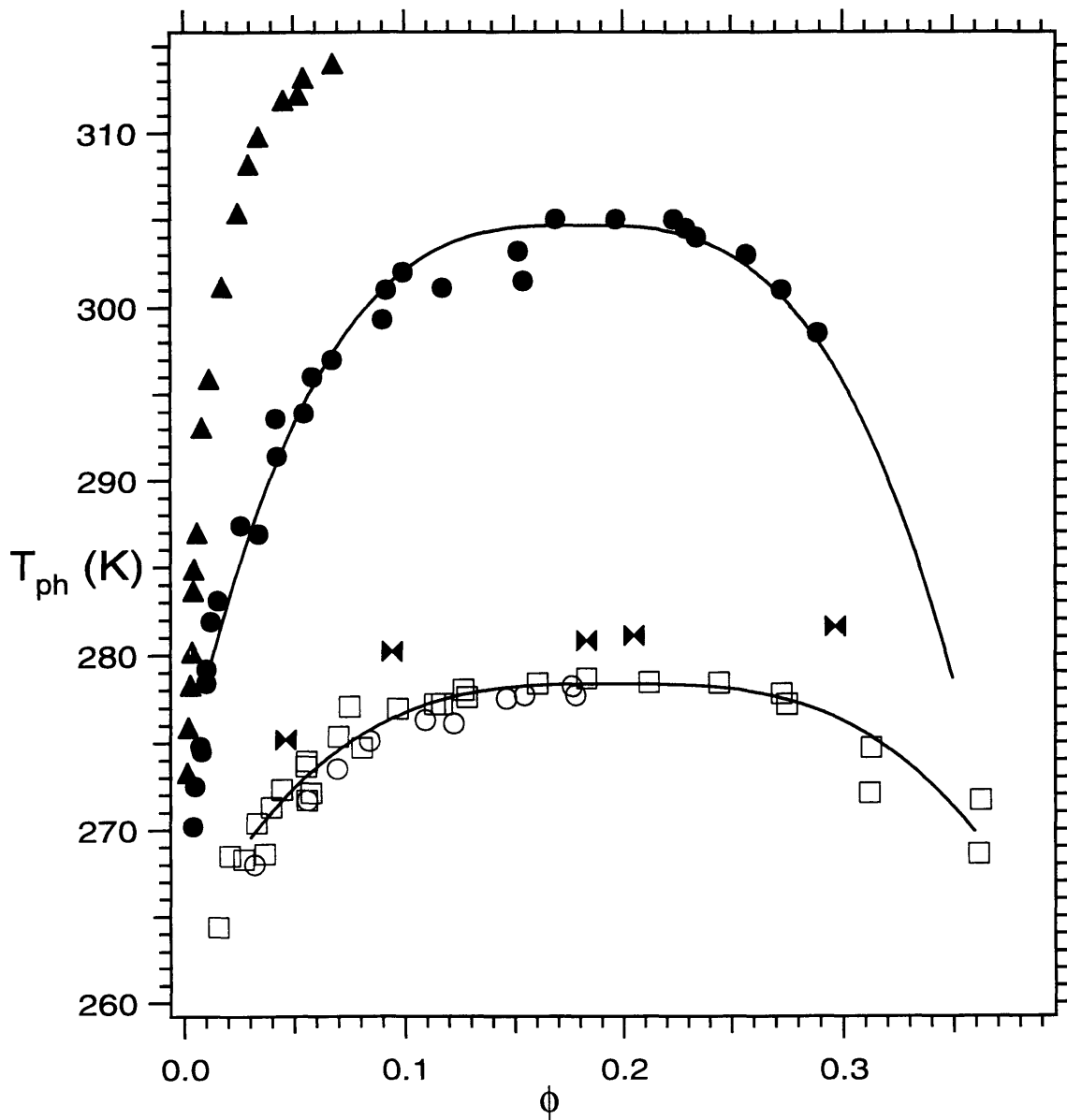


Figure 4-4: Coexistence curves for the oligomers of γ_{IIIb} -crystallin in 0.1 M sodium phosphate (pH 7.1). Native γ_{IIIb} monomers, from Ref. [2] (open squares); control γ_{IIIb} monomers, this report (open circles); γ_{IIIb} monomers after crosslinking (solid bowties); γ_{IIIb} dimers (solid circles); γ_{IIIb} trimers (solid triangles). The solid lines are fits to Eq. (4.1).

4.3. RESULTS AND DISCUSSION

where $\beta = 0.325$, ϕ^I and ϕ^{II} are the volume fractions of the dilute and dense coexisting phases respectively, ϕ_c is the critical volume fraction, T_c is the critical temperature (in Kelvin) and A is a parameter that characterizes the width of the coexistence curve. In Table 4.1 we list the values of the parameters ϕ_c , T_c and A for the native monomer and for the dimer. Note that the parameter A as defined in Eq. (4.1) is a factor of two greater than that used by Broide *et al.* [2].

Table 4.1: Coexistence curve parameters (see Eq. (4.1) in text) for the monomer and dimer.

Oligomer	ϕ_c	T_c (K)	A
Monomer	0.20 ± 0.01	278.4 ± 0.2	5.2 ± 0.2
Dimer	0.18 ± 0.01	304.7 ± 0.5	4.2 ± 0.2

We note that the dimer has a lower critical volume fraction, a higher critical temperature and a narrower coexistence curve than the monomer. As will be discussed in Sec. 4.3.2 this can be understood in terms of the dimer having a longer effective range of interaction than the monomer. Unfortunately, our data does not allow us to determine the values of ϕ_c , T_c and A for the trimers. We however point out that one important trend is seen in all three systems. At a fixed concentration the phase separation temperatures increases with the size of the oligomer.

Another interesting result we find is that the T_{ph} of the monomer samples after crosslinking (solid bowties) is approximately 3°C higher than that of the native monomer, perhaps due to the formation of an intermolecular crosslink. This reaction of the crosslinker with the monomer is an important issue to be investigated, but in the rest of this chapter we will concentrate on the phase separation of the oligomers.

4.3.2 Comparison with Theoretical Models

In Fig 4-5 we show the width of the coexistence curve A and the critical volume fraction ϕ_c for various systems. These are the simple gases nitrogen, carbon monoxide, methane, neon, argon, krypton and xenon [12, 103, 104]; the globular proteins γ -crystallins [2], globulins [105], arachin [106], apoferritin [7], concanavalin [7] and

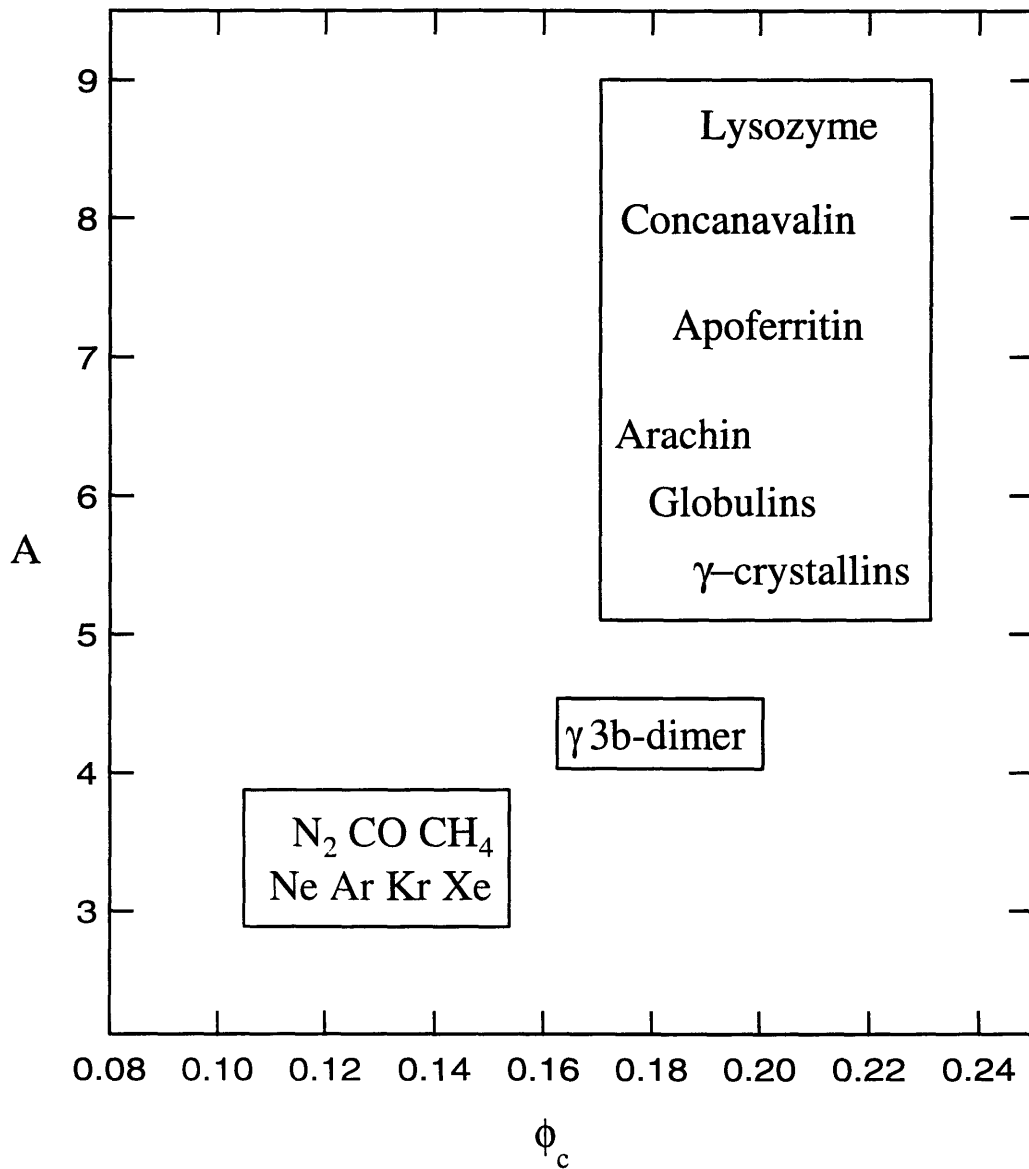


Figure 4-5: The width of the coexistence curve A and the critical volume fraction ϕ_c for various systems. The center of each word lies approximately at the values of A and ϕ_c corresponding to that system.

lysozyme [89]; and the results for the γ_{IIIb} dimer from Table 4.1. The center of each word lies approximately at the values of A and ϕ_c corresponding to that system. The boxes are intended to highlight the fact that the systems cluster into groups. The simple gases all have $\phi_c \approx 0.13$ and $A \approx 3.5$. As was discussed in Chapter 2, these are the values expected for spherical particles with long ranges of interaction [18, 47, 107]. The proteins all have larger values of ϕ_c (0.19-0.21) and A (5.2-8.8). This is due to the short-range nature of the protein interactions. The larger spread in these values for each protein as compared to the simple gases reflects the greater experimental uncertainty in determining ϕ_c and A for proteins.

We see that the γ_{IIIb} dimer lies in between the two groups. This suggests that if we wish to characterize the dimer phenomenologically, it should be described as having an effective range of interaction that falls in between that of monomeric proteins and simple gases. We found in Chapter 2. that as the range of interaction increases so does the critical temperature (see Table 2.1 and Fig. 2-6). Since the dimer may be thought of as a longer effective range of interaction than the monomer, we would expect it to have a higher critical temperature. This is indeed observed experimentally.

We now wish to examine whether the trends we have observed experimentally for the dimers, namely a higher critical temperature, a higher critical volume fraction and a wider coexistence curve than the monomer, are general in nature. We do so by considering several models that take into account the dumbbell shape of the hard core of the dimers, which we did not consider in our qualitative discussion above.

There have been several theoretical investigations into the free energy of dimers consisting of two tangent spheres [108, 109, 110, 111, 112, 113, 114], where both spheres are assumed to have a given range of interaction. Below we present the free energies obtained for three models corresponding to different ranges of interaction: (A) mean field, (B) a square-well potential with $\lambda=1.5$; and (C) adhesive spheres. As we have discussed in Chapter 2, the mean field model is equivalent to a long-ranged ($\lambda \rightarrow \infty$) square-well potential, while for the adhesive sphere the range of interaction is very short ($\lambda \rightarrow 1$).

For each model, we consider the monomer (denoted by the subscript m) and

the corresponding dimer (denoted by the subscript d). We give for each system the reduced free energy per monomeric unit $f \equiv F/N_m kT$. Here N_m is the total number of monomeric units (whether free or bound as a dimer) and F is the Helmholtz free energy of the system. Note that the expressions given below were derived for one component fluids. However, we are assuming the protein-water solution to be incompressible and thus its free energy is equivalent to that of a one component fluid (as is shown in Appendix A). We use the expressions for the free energy of the various models to numerically determine the coexistence curves. This is done by finding the coexisting volume fractions ϕ^I and ϕ^{II} which at each temperature minimize the total free energy of the system.

We wish to point out that using our Monte Carlo method (see Chapter 2) we have already determined the coexistence curves for the monomer systems we are about to consider. For dimer systems, however, simulations are much more difficult to carry out and there are no reliable simulations available. Thus to understand our experimental results we resort to previously published theoretical calculations for the free energy of the dimers. These are described in the various sections below. For consistency we use similar calculations for the free energy of the corresponding monomer systems. As we have mentioned in Chapter 2, theoretical predictions for the free energy of for the monomer systems are only in qualitative agreement with the results from simulations. Nevertheless, as we will see below, they capture the trends observed in the simulations. Thus we believe that predictions of the calculations for dimers will also be qualitatively correct, if not numerically accurate. We use these calculations to illustrate that the relatively short-range of interaction of the monomer plays a role in shaping the phase diagram of the dimer.

A. Mean field

A1. *Monomers.* The free energy for mean field monomers we use is [18]

$$f_m = f_{CS} - a\phi. \quad (4.2)$$

Here f_{CS} is the Carnahan-Starling approximation for the free energy of hard spheres [55], given by

$$f_{CS} = \ln \phi + \frac{3 - 2\phi}{(1 - \phi)^2}. \quad (4.3)$$

Note that f_{CS} represents the entropic contribution to the free energy of the monomers, while $-a\phi$ is the energetic contribution (see Eq. (4.2)). Here a is the well known van der Waals term [61] and was discussed in Sec. 2.3.3. For convenience, we repeat here its definition in terms of the square-well model.

$$a = 4\hat{\epsilon}\lambda^3. \quad (4.4)$$

A2. *Dimers.* The corresponding free energy for mean field dimers is

$$f_d = f_{TS} - a\phi. \quad (4.5)$$

Here f_{TS} is the Tildesley-Street [115] approximation for the free energy of hard di-spheres, given by

$$f_{TS} = \frac{1}{2} \ln \phi + 1.377515 \ln(1 - \phi) + \frac{3.154575 - 2.2031\phi}{(1 - \phi)^2}. \quad (4.6)$$

The factor of one half preceding the $\ln \phi$ term is due to the fact that we normalized the free energy to the number of monomer units. The number of dimers is half the number of the constituent monomers.

B. Square-well potential, $\lambda=1.5$

The free energies we use for both the monomer and dimer are those obtained by second order perturbation theory about the corresponding hard core system (whose properties are known from simulations). We use the expression for the free energy as obtained by Yethiraj and Hall [110]. We have chosen to use the results for $\lambda=1.5$, for this is the only range of interaction which lies in between the mean field and adhesive sphere limits and for which there are published calculations.

B1. *Monomers.* For the monomer we have

$$f_m = f_{CS} - \hat{\epsilon} f_m^{(1)} - \hat{\epsilon}^2 f_m^{(2)}. \quad (4.7)$$

with

$$f_m^{(1)} = -22.563545\phi + \frac{54.5051575 - 131.4520225\phi + 80.95349\phi^2}{(1 - \phi)^3} \quad (4.8)$$

and

$$f_m^{(2)} = \frac{0.6320017 + 12.80252795\phi + 230.5261307\phi^2}{(1 + 6.75237\phi)^3}. \quad (4.9)$$

B2. *Dimers.* The free energy of the dimers is

$$f_d = f_{TS} - \hat{\epsilon} f_d^{(1)} - \hat{\epsilon}^2 f_d^{(2)}. \quad (4.10)$$

where

$$f_d^{(1)} = -28.72802589\phi + \frac{56.35234651 - 134.3272537\phi + 81.56952133\phi^2}{(1 - \phi)^3} \quad (4.11)$$

and

$$f_d^{(2)} = \frac{0.58829498 + 14.59145098\phi + 366.8184896\phi^2}{(1 + 8.26765\phi)^3}. \quad (4.12)$$

C. Adhesive spheres

C1. *Monomers.* For the monomers we use the results of Watts *et al.* [39]. As mentioned in Sec. 2.3.3, these authors calculate the pair distribution of adhesive spheres using the Percus-Yevick equation. They use the energy equation to find an expression for the free energy of adhesive spheres. This is

$$f_m = f_{CS} - 6 \int_{\tau}^{\infty} [\alpha - (\alpha^2 - \gamma)^{1/2}] \frac{d\tau'}{\tau'}, \quad (4.13)$$

where

$$\alpha = \tau + \frac{\phi}{1 - \phi} \quad (4.14)$$

and

$$\gamma = \frac{\phi(1 + \phi/2)}{3(1 - \phi)^2}. \quad (4.15)$$

Recall that the parameter τ is defined as (Eq. (2.28))

$$\tau = \frac{1}{12(\lambda - 1)\exp(\hat{\epsilon})}. \quad (4.16)$$

Note that we choose to work with the energy equation (see Sec 2.3.3), for it leads to values of ϕ_c and τ_c which are close to those we find by simulation.

C2. *Dimers*. For the dimer we use the expression given by Banaszack *et al.* [112] for chains of adhesive spheres, which they obtained from the theory of association of Wertheim [114, 116]. The free energy of adhesive dimers is found to be

$$f_d = f_m - \frac{1}{2} \ln(\phi y_m) \quad (4.17)$$

where

$$y_m = \frac{6\tau}{\phi} [\alpha - (\alpha^2 - \gamma)^{1/2}] \quad (4.18)$$

is the cavity function for adhesive monomers.

Coexistence Curves

In Fig. 4-6 we show the coexistence curves corresponding to the above six systems (A1, A2, B1, B2, C1 and C2). The temperatures are given in reduced units T/T_c^m in terms of the critical temperature of the monomer T_c^m , taken to be the same for all three monomer systems. To convert the parameter τ to a real temperature, Eq. (4.16) was used with $\lambda = 1.05$. This leads to a value of $\hat{\epsilon}_c = 2.66$, consistent with our Monte Carlo simulations (see Table 2.1).

We see from Fig. 4-6 that the value of ϕ_c and the width of the coexistence curve for the dimers increase as the range of the constituent monomer decreases. This is consistent with our previous findings for the monomer. The critical parameters for each system are given in Table 4.2. In column 2 we present the range of interaction of the constituent monomer. In column 3 and 4 we show, respectively, the value of ϕ_c

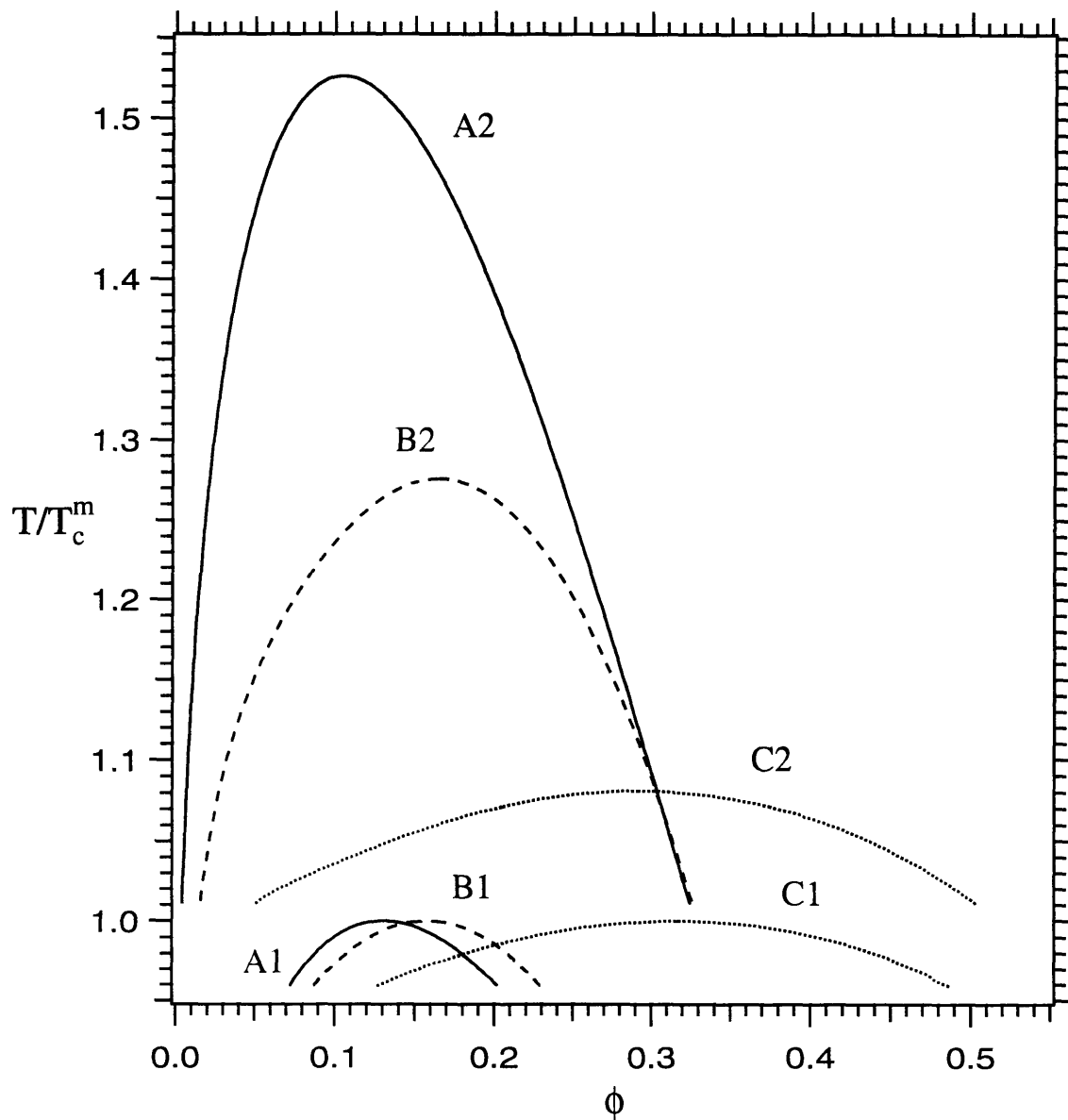


Figure 4-6: Coexistence curves for (1) monomers and (2) dimers according to various models. A: mean field (solid lines); B: square-well potential with $\lambda=1.5$ (coarse dashed lines); C: adhesive spheres (fine dashed lines).

4.3. RESULTS AND DISCUSSION

Table 4.2: Coexistence curve parameters for the curves presented in Fig. 4-6.

Model	Range	ϕ_c	critical temperature	T_c^d/T_c^m
Mean field monomer (A1)	$\lambda \rightarrow \infty$	0.130 (0.130)	$a_c=10.6$ (10.6)	—
Mean field dimer (A2)	—	0.105	$a_c=6.95$	1.53
Square-well monomer (B1)	$\lambda = 1.5$	0.157 (0.171)	$\hat{\epsilon}_c=0.745$ (0.763)	—
Square-well dimer (B2)	—	0.165	$\hat{\epsilon}_c=0.596$	1.28
Adhesive monomer (C1)	$\lambda \rightarrow 1$	0.316 (0.266)	$\tau_c=0.1164$ (0.13)	—
Adhesive dimer (A2)	—	0.291	$\tau_c=0.1419$	1.08

and the critical temperature (expressed as the dimensionless temperature appropriate for each system) for the the different systems. The values in parenthesis shown in columns 3 and 4 are those we obtained from our simulation results in Chapter 2 (see Table 2.1 and Sec. 2.3.3) for the monomer systems. The ratio of critical temperature of the dimer T_c^d to the critical temperature of the monomer T_c^m is given in column 5.

When we compare the dimers with their constituent monomers we see that the critical temperature of any given dimer is higher than that of its constituent monomer [see columns 4 and 5; recall that $a \propto 1/T$ (Eq. (4.4)); $\hat{\epsilon} \propto 1/T$ (see Sec. 2.2); and τ increases monotonically with T (Eq. (4.16))]. This increase of the critical temperature of the dimer can be interpreted, as we did previously, as the dimer having an effectively larger range of interaction. From column 5, we see that although T_c^d/T_c^m is always greater than unity, it decreases as the range of interaction of the constituent monomer decreases.

We also see that ϕ_c (column 3) of the dimer tends to be smaller than that of the monomer. This is again consistent with our phenomenological description that the effective range of the dimer is longer than that of the monomer. As we have seen in Chapter 2, ϕ_c decreases as the range of interaction increases. We note that the value of $\phi_c = 0.165$ for the $\lambda=1.5$ dimer is actually larger than that of the monomer $\phi_c = 0.157$. There is however some theoretical uncertainty in these results, for they depend on which analytic functions are used to express the various terms in the free energies given above. We find from our Monte Carlo simulations that for a monomer with $\lambda=1.5$, $\phi_c = 0.171$. By comparing the Monte Carlo results with the theoretical

results for ϕ_c and $\hat{\epsilon}_c$ in Table 4.2, we see that prediction of the theories used for the monomers are qualitatively correct, although not in strict numerical agreement with the simulations. We expect the theories for the dimers to be of a similar accuracy.

The trends shown by these theoretical results are consistent with our experimental findings given in Sec. 4.3.1. The γ_{IIIb} dimers have a lower value of ϕ_c and a higher critical temperature than the monomer, as is found theoretically for the dimer-monomer pairs. Experimentally we observe $T_c^d/T_c^m = 1.09$ (Table 4.1). This is very close to the value of 1.08 found theoretically for the adhesive dimers and contrasts sharply with the value of 1.53 found in the mean field limit (see Table 4.2). Of course, the experimentally observed change in the critical temperatures between the dimer and the monomer may be due in part to chemical changes brought about by the crosslinker. However, we believe that the relatively small change in the critical temperature of the dimer is one more phenomenon which reflects the short-range nature of the protein interactions.

4.3.3 Effect of Oligomers on the Phase Behavior of Globular Proteins

In Fig 4-7 we compare the coexistence curves for the oligomers of γ_{IIIb} -crystallin with those of γ_{II} -crystallin. The native γ_{IIIb} and γ_{II} monomers are shown as open squares and open triangles respectively. We note that the two coexistence curves are almost identical. The solid symbols represent our results for the γ_{IIIb} oligomers: dimers (circles), trimers (triangles) and a mixture of higher n-mers (squares). We estimated (from our Superdex 75 separation—see Sec. 4.2.2) that the n-mer mixtures contained approximately 10% trimers, 70% tetramers and 20% pentamers. The crossed squares are the results for the partial coexistence curve of γ_{III} taken from Ref. [23].

We can see from this figure, that although the coexistence curves of the γ_{IIIb} and γ_{II} monomers are practically the same, the coexistence curve of γ_{III} is very different from that of the γ_{IIIb} dimer, even though γ_{III} was suggested to consists of dimers of γ_{II} [86]. In fact the partial coexistence curve of γ_{III} resembles more closely that

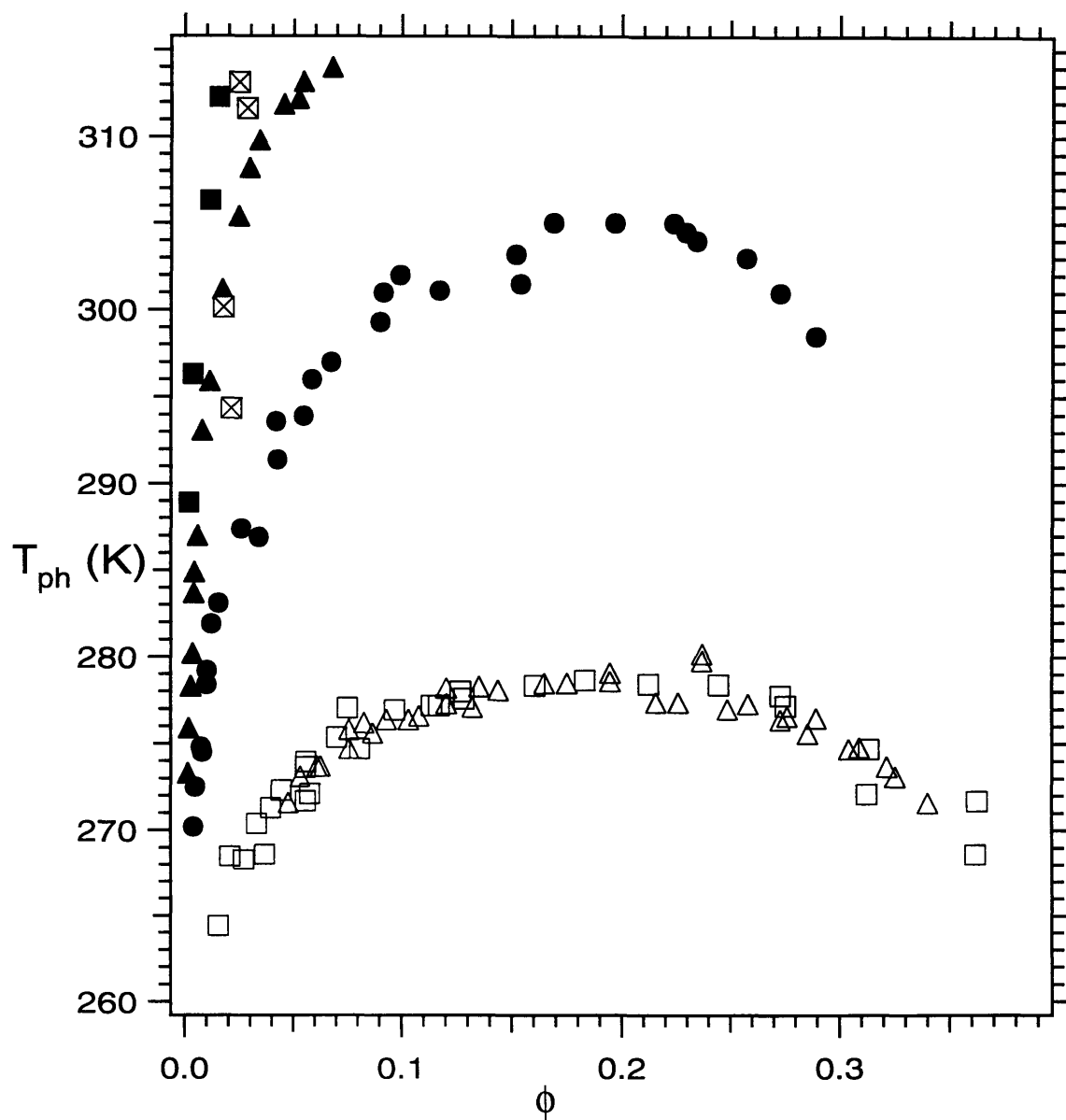


Figure 4-7: Comparison of the coexistence curves for γ_{IIIb} and γ_{II} oligomers. Native γ_{IIIb} and γ_{II} monomers, from Ref. [2] (open squares and open triangles respectively); γ_{IIIb} dimers (solid circles); γ_{IIIb} trimers (solid triangles); γ_{IIIb} n-mers (solid squares); γ_{IIIH} from Ref. [23] (crossed squares).

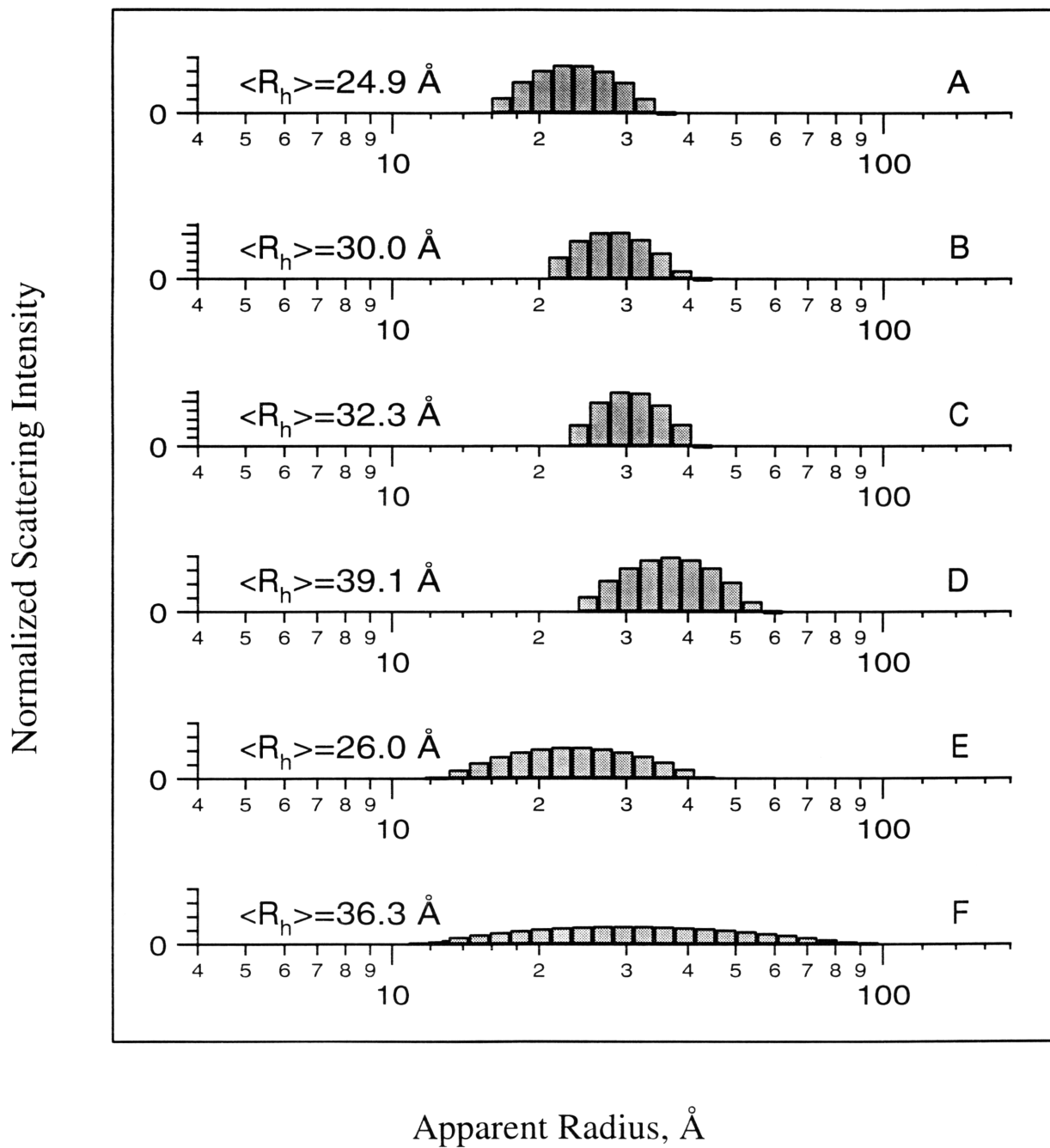


Figure 4-8: QLS data for γ_{IIIb} - and γ_{II} -crystallins. Normalized scattering intensity is plotted against apparent R_h . (A) γ_{IIIb} monomer. (B) γ_{IIIb} dimer. (C) γ_{IIIb} trimer. (D) γ_{IIIb} n-mer mixture. (E) γ_{II} monomer. (F) γ_{III} .

of the γ_{IIIb} n-mers than that of the γ_{IIIb} dimers. This suggests that γ_{IIIH} , reported previously to consist of dimers, may in fact contain some n-mers.

There is further evidence that γ_{IIIH} may actually contain some higher n-mers. In Fig 4-8 we show the normalized scattering intensity plotted against the apparent hydrodynamic radius R_h as obtained by QLS. The results for the γ_{IIIb} oligomers are shown in Figs. 4-8(A-D): native monomer (A); dimer (B); trimer (C) and n-mer mixture (D). The average hydrodynamic radius $\langle R_h \rangle$ increases with size as expected. In Figs. 4-8E and F we present the results for the γ_{II} monomer and for γ_{IIIH} , respectively, taken from Ref. [23]. We see that the average R_h of the γ_{II} monomer is 26.0Å not far from the average R_h of the γ_{IIIb} monomer, 24.9Å. The hydrodynamic radius for γ_{IIIH} however, differs substantially from that of the γ_{IIIb} dimer. The average hydrodynamic radius of γ_{IIIH} is approximately 36Å, significantly larger than the 30Å found for the γ_{IIIb} dimer. In addition, the distribution of particle sizes is much broader for γ_{IIIH} (10-100Å) than is seen for the γ_{IIIb} dimer (20-40Å). We note that for the n-mer mixture of γ_{IIIb} the average hydrodynamic radius is 39Å with a spread in sizes (25-55Å) which is broader than that of the γ_{IIIb} dimer. Thus, it is the n-mer mixture and not the γ_{IIIb} dimers, which most closely resembles γ_{IIIH} in the distribution of particle sizes.

In Table 4.3 we show the hydrodynamic radii expected in the limit of zero concentration for oligomers made up of tangent spheres. This table has been adapted from the calculations of García de la Torre and Bloomfield [117]. Note that $R_h(m)$ is the hydrodynamic radius of the spherical monomer.

For the γ_{IIIb} dimer we find experimentally that $\langle R_h \rangle / \langle R_h(m) \rangle \approx 1.2$, while for γ_{IIIH} we have $\langle R_h \rangle / \langle R_h(m) \rangle \approx 1.4$. These values are both close to those expected theoretically for a dimer (1.38). However, the distribution of particles sizes is very different for the two samples. The largest particles seen in the γ_{IIIb} sample has only $R_h / \langle R_h(m) \rangle \sim 1.6$. The results presented in Table 4.3 suggest that this corresponds to the presence of some trimers in the dimer sample, which is consistent with the gel electrophoresis results (see Fig. 4-3). In contrast, we see that γ_{IIIH} contains particles with sizes up to $R_h / \langle R_h(m) \rangle \sim 3.8$, which correspond to

Table 4.3: Hydrodynamic radii in the limit of zero concentration for oligomers made up of spheres [117]. $R_h(m)$ is the hydrodynamic radius of the spherical monomer.

No. of subunits (n)	Geometry	$R_h/R_h(m)$
1	Sphere	1.00
2	Dimer	1.38
3	Triangle	1.61
3	Linear	1.71
4	Square	1.82
4	Tetrahedron	1.77
4	Linear	2.00
5	Pentagon	2.04
5	Bipyramid	1.92
6	Hexagon	2.25
6	Octahedron	2.02
6	Trigonal prism	2.07
8	Cube	2.31

very large particles; from Table 4.3 we see that octamers are expected to only have a size of $R_h / \langle R_h(m) \rangle = 2.31$. Thus by comparing the QLS results together with the values presented in Table 4.3 for R_h , we can conclude that the γ_{IIIH} sample contains a significant proportion of large oligomers while the γ_{IIIB} sample is mostly dimers. Note that for the γ_{IIIB} n-mer sample, $\langle R_h \rangle / \langle R_h(m) \rangle \approx 1.6$, which is consistent with this sample being composed of a mixture of trimers, tetramers and pentamers.

We have seen in Sec. 4.3.1 that at a fixed concentration T_{ph} increases with the number of subunits in the oligomer. We have argued above that γ_{IIIH} actually contain higher n-mers as well. Thus, we may attribute the difference in the coexistence curves between the γ_{IIIB} dimers and γ_{IIIH} to the presence of n-mers in the latter. Of course, it is possible that part of the difference in the phase behavior may due to the different chemical nature of the protein oligomers, leading to different energies of interaction. The γ_{IIIB} dimer contains a crosslinker while the γ_{IIIH} dimer is thought to be a mixture of covalent dimers held together by a disulphide link as well as associated (non-covalent) dimers [86]. Our results however strongly suggest that these associated dimers can actually form larger oligomers which lead to the observed QLS and coexistence curve data.

4.4 Conclusions

We have produced, through chemical crosslinking, oligomers of the protein γ_{IIIb} -crystallin. We have identified a crosslinking reagent and a set of experimental conditions which enable us to generate large amount of crosslinked oligomers. We have presented the coexistence curves of the dimer and trimer species. We have observed that the dimer has a higher critical temperature, a lower critical volume fraction and a narrower coexistence curve than the native monomer. Our results for the trimer and higher n-mers have shown that at a fixed concentration the phase separation temperature increases with the size of the oligomer. We have found that our experimentally observed increase in the critical temperature of the dimer is consistent with the monomer having a short range of interaction.

We have shown how our results for the dimer can be explained by considering the dimer to have a longer effective range of interaction than the monomer. We have used our results to investigate the effect of aggregation on liquid-liquid phase separation in globular protein solutions. We have observed that the presence of oligomers can dramatically raise the phase separation temperatures. We have compared our results with those obtained for γ_{II} -crystallin where the formation of a new species γ_{III} is believed to raise the phase separation temperature of the protein solution. This work further strengthens the results obtained previously for γ_{II} . However, we have shown here that γ_{III} does not solely consist of dimers of γ_{II} as previously suggested, but that these dimers are accompanied by the formation of larger associated aggregates.

The model system we have presented here allows us to produce large quantities of pure and stable oligomers. With this system we will now be able to construct multicomponent mixtures of monomers and oligomers which mimic the mixtures of monomeric protein and aggregates which form naturally in many biological settings. We believe that the model experimental system we have developed opens the door to a systematic investigation of the effect of aggregation on phase separation in globular protein solutions.

Chapter 5

Conclusions

In this thesis, I have shown how the phase behavior of globular proteins can be understood in terms of their microscopic interactions.

In Chapter 2, I described how we studied the liquid-liquid phase separation of globular proteins. The proteins are modeled as hard spheres with an attractive square-well potential of depth ϵ and reduced range λ . We compared the experimentally determined phase diagram with the results of a modified Monte Carlo procedure which combines simulations with analytic techniques. The simplicity and economy of the procedure made it practical to investigate the effect on the phase diagram of an essentially continuous variation of λ in the domain $1.05 \leq \lambda \leq 2.40$. In particular, we were able to study, without great computational cost, the domain most relevant to globular proteins, that of short ranges of interaction. We calculated the coexistence curves and found that they are in good agreement with the information available from previous standard Monte Carlo simulations conducted at a few specific values of λ . We analyzed the experimental data for the critical volume fractions of the γ -crystallins and determined the actual range of interaction appropriate for these proteins. A comparison of the experimental and calculated widths of the coexistence curves suggests a significant contribution from anisotropy in the real interaction potential of the γ -crystallins. The dependence of the critical volume fraction ϕ_c and the reduced critical energy $\hat{\epsilon}_c$ upon the reduced range λ were also analyzed in the context of two limiting cases: mean field theory ($\lambda \rightarrow \infty$) and the Baxter adhesive sphere model

($\lambda \rightarrow 1$). Mean field theory failed to describe both the value of ϕ_c and the width of the coexistence curve of the γ -crystallins. This is consistent with the observation that mean field is no longer applicable when $\lambda \leq 1.65$. In the opposite case, $\lambda \rightarrow 1$, we obtained the critical parameters for ranges much shorter than those investigated in the literature. This allowed us to reliably determine the critical volume fraction in the adhesive sphere limit: $\phi_c(\lambda = 1) = 0.266 \pm 0.009$.

In Chapter 3, I described how we continued our investigation of the phase diagram of globular proteins by studying the liquid-solid transition. Here we presented an analytic expression for the chemical potential of a solid with short-range interactions by using a simple cell model. We then used this chemical potential, together with the chemical potential of the liquid phase we had obtained previously, to show how the relative positions of the phase boundaries are related to the range of interaction and the number of contacts made per particle in the solid phase. In particular, we explained how the “metastability gap” of globular protein solutions (i.e., the difference between the temperature at which a stable solid phase appears and that at which coexisting liquid phases are first stable) is a manifestation of the shortness of the range of interaction and of the number of contacts made in the solid phase. The theory presented successfully describes the features of the phase diagrams observed in a wide variety of colloidal systems, including globular proteins.

In Chapter 4, I discussed the effect of oligomers on the phase diagram of globular protein solutions. We produced dimers, trimers and larger n-mers of globular proteins by crosslinking native γ_{H1B} -crystallin protein. We determined the experimental conditions which maximized the yield of the dimer species. We presented results for the liquid-liquid coexistence curves of the dimers and trimers. We found that at a given concentration, the trimers phase separated at a higher temperature than the dimers, which in turn phase separated at a higher temperatures than the native monomer protein. In addition, we observed that the critical volume fraction and the width of the coexistence curve for the dimers were smaller than those of the monomers. We analyzed the results within the context of the model used to describe protein monomers. We found that the protein oligomers may be described phenomenologically as having

longer effective ranges of interaction than the monomer protein. The experimental findings were then used to illustrate the relationship between aggregation and phase separation in globular protein solutions.

The work in this thesis has addressed several important issues regarding the phase behavior of globular proteins. I suggest below possible further work.

I have discussed in this thesis that the interactions of globular proteins are anisotropic. I believe that it would be interesting to extend the work presented in Chapters 2 and 3 to include anisotropic interactions. Anisotropic potentials should be investigated not only because they are more realistic, but also because they may shed light on the connection between crystallization, liquid-liquid phase separation and aggregation. How is the anisotropy of the interaction related to crystallization, liquid-liquid phase-separation and aggregation? As illustrated by the work I have done, isotropic interactions allow all three types of behavior to be observed. However, for highly anisotropic interactions (e.g., a sphere with two narrow cones of attraction so that each particle only interacts with a neighbor exactly to its right and left), liquid-liquid phase separation cannot occur. This is because with such anisotropic interactions only chains of particles may form, not the drop-like clusters needed to create new liquid phases. Therefore, by systematically making the interaction between particles more and more anisotropic, it should be possible to observe a change in behavior from liquid-liquid phase separation to aggregation.

It is also important to carry out more experiments to study how the microscopic interactions affect the phase diagrams of proteins. I believe that the work which I described in Chapter 4, related to the phase diagram of crosslinked oligomers, should be extended. If sufficient protein is available, it would be interesting to measure the complete coexistence curve of the trimers (and those of higher n-mers). This study could then be continued to characterize the behavior of other phase separating globular proteins, such as lysozyme or other members of the γ -crystallin family, to see if the universality of the monomer phase diagram is repeated for the oligomers. Another possible extension would be to explore the role of the crosslinker. At its simplest, this would include the specific chemical effects of a given crosslinker on a

protein, and hence on its phase behavior. However, a more appealing possibility is to vary the length of the spacer arm in the crosslinker and study the effect of the distance between the constituents of the oligomer on the phase behavior. I would expect this to produce dramatic results, for in the limit of a very long spacer arm the oligomers should behave like a polymer even though each of the constituents is short-ranged.

I hope that the work presented in this thesis will encourage others to explore the phase behavior of solutions of globular proteins.

Appendix A

The Incompressible Protein-Water Solution

In this Appendix we show the equivalence of the thermodynamic descriptions of an incompressible protein-water solution and that of a compressible one component fluid.

Consider a solution of N_p protein molecules and N_w water molecules at pressure P and temperature T . Let the chemical potential of a protein molecule be μ_p and let that of a water molecule be μ_w . We assume that the protein-water solution is incompressible. Therefore the volume of the system may be written as

$$V = N_p\Omega_p + N_w\Omega_w. \quad (\text{A.1})$$

Here Ω_p is the volume of a protein molecule and Ω_w is that of a water molecule.

The equivalence between the incompressible protein-water solution and a compressible one component fluid is most easily seen by considering the Helmholtz free energy $F(T, V, N_p, N_w)$

$$dF(T, V, N_p, N_w) = -SdT - PdV + \mu_p dN_p + \mu_w dN_w. \quad (\text{A.2})$$

where S and P are, respectively, the entropy and pressure of the system.

Equation (A.1) allows us to eliminate the variable N_w from the free energy. From

this equation we write

$$dN_w = \frac{1}{\Omega_w} dV - \gamma dN_p, \quad (\text{A.3})$$

where $\gamma \equiv \Omega_p/\Omega_w$ is the ratio of the volume of one protein molecule to the volume of a water molecule.

Substituting Eq. (A.3) into Eq. (A.2) we obtain

$$dF(T, V, N_p, N_w) = -SdT - \left(P - \frac{\mu_w}{\Omega_w}\right)dV + (\mu_p - \gamma\mu_w)dN_p. \quad (\text{A.4})$$

The quantity $\pi \equiv -\mu_w/\Omega_w$ is the osmotic pressure of the solution.

We now compare Eq. (A.4) with the free energy of a compressible fluid of N_p particles at temperature T with volume V . This free energy is

$$dF(T, V, N_p) = -SdT - PdV + \mu dN_p. \quad (\text{A.5})$$

Here S is the entropy and P is the pressure of the fluid. The chemical potential of the fluid is μ .

By comparing Eqs. (A.4) and Eqs. (A.5) we see that the free energy of an incompressible protein-water solution is formally equivalent to the free energy of a compressible one component fluid consisting of N_p particles in a volume V at temperature T . The effective chemical potential $\mu_{eff} = \mu_p - \gamma\mu_w$ is equivalent to the chemical potential μ of the fluid. The total pressure $P + \pi$ of the protein-water solution corresponds to the pressure P of the fluid.

Note that since the protein-water solution is incompressible its free energy is insensitive to external pressure P applied to it. We may therefore consider the external pressure to be zero. In this case, the equivalent pressure in the one component fluid is simply the osmotic pressure π . The relationship between the two systems is summarized in Table A.1.

Table A.1: Corresponding variables for the incompressible protein-water solution (at zero external pressure) and the compressible one component fluid systems.

Incompressible protein-water solution	Compressible one component fluid
Number of protein molecules, N_p	Number of particles, N_p
Helmholtz free energy, $F(N_p, V, T)$	Helmholtz free energy, $F(N_p, V, T)$
Chemical potential, $\mu_p - \gamma\mu_w$	Chemical potential, μ
Osmotic pressure π	Pressure P

Appendix B

Widom's Formula

Here we present the original derivation of Widom's formula [54] for the chemical potential of a system of N identical particles, without internal degrees of freedom, in a volume V at temperature T . The equivalence of the thermodynamic descriptions of this system and that of the protein-water solution is discussed in Appendix A.

In the canonical ensemble the chemical potential is given by

$$\mu = \left. \frac{\partial F}{\partial N} \right|_{V,T} = F(N+1, V, T) - F(N, V, T), \quad (\text{B.1})$$

where F is the Helmholtz free energy.

The Helmholtz free energy F is defined as $F(N, V, T) = -kT \ln Z_N(V, T)$, where Z_N is the partition function of a system of N particles. The chemical potential can therefore be written as

$$\mu = -kT \ln \left(\frac{Z_{N+1}(V, T)}{Z_N(V, T)} \right). \quad (\text{B.2})$$

The ratio of the two partition functions is given by

$$\frac{Z_{N+1}}{Z_N} = \frac{1}{(N+1)} \frac{\int \exp(-\beta E_{N+1}\{r_i, p_i\}) \prod_{i=1}^{N+1} d^3 r_i d^3 p_i}{\int \exp(-\beta E_N\{r_i, p_i\}) \prod_{i=1}^N d^3 r_i d^3 p_i}, \quad (\text{B.3})$$

where E_N and E_{N+1} are the energies of the system with N and $N+1$ particles respectively. The position and momenta of the particles are denoted by $\{r_i, p_i\}$. The

factor $1/(N+1)$ is due to the indistinguishability of the particles. As usual $\beta = 1/k_B T$

Now we split the energies into kinetic (K) and potential (W) contributions i.e., $E_N\{r_i, p_i\} = K_N\{p_i\} + W_N\{r_i\}$ and similarly for E_{N+1} . We then perform the integration over the momenta coordinates to obtain (assuming $N \gg 1$)

$$\frac{Z_{N+1}}{Z_N} = \left(\frac{1}{NV_F} \right) \frac{\int \exp(-\beta U\{r_i\}) \exp(-\beta W_N\{r_i\}) \prod_{i=1}^{N+1} d^3 r_i}{\int \exp(-\beta W_N\{r_i\}) \prod_{i=1}^N d^3 r_i}. \quad (\text{B.4})$$

Here V_F is the Fermi volume, which comes from the extra momentum integration in the numerator. We have also written $W_{N+1}\{r_i\} = W_N\{r_i\} + U\{r_i\}$. Thus, $U\{r_i\}$ is the extra potential energy associated with adding the $(N+1)^{\text{th}}$ particle into the original N particle system.

Clearly $W_N\{r_i\}$ is a function of $\{r_1, r_2, \dots, r_N\}$, while $U\{r_i\}$ is a function of $\{r_1, r_2, \dots, r_N, r_{N+1}\}$. However, if we assume that the system is translationally invariant, then the potential energies are only functions of the distances between the particles i.e., of $|r_i - r_j|$ where $i \neq j$. Thus, by changing coordinates so that the origin is centered at the position of the $(N+1)^{\text{th}}$ particle, namely $r'_i = r_i - r_{N+1}$, with $i = 1, 2, \dots, N$, we can write Eq. (B.4) as

$$\frac{Z_{N+1}}{Z_N} = \left(\frac{1}{NV_F} \right) \frac{\int \exp(-\beta U\{r'_i\}) \exp(-\beta W_N\{r'_i\}) (d^3 r_{N+1}) \prod_{i=1}^N d^3 r'_i}{\int \exp(-\beta W_N\{r'_i\}) \prod_{i=1}^N d^3 r'_i}. \quad (\text{B.5})$$

and hence

$$\frac{Z_{N+1}}{Z_N} = \left(\frac{V}{NV_F} \right) \frac{\int \exp(-\beta U\{r'_i\}) \exp(-\beta W_N\{r'_i\}) \prod_{i=1}^N d^3 r'_i}{\int \exp(-\beta W_N\{r'_i\}) \prod_{i=1}^N d^3 r'_i}. \quad (\text{B.6})$$

Let us define

$$\langle \exp(-\beta U) \rangle \equiv \frac{\int \exp(-\beta U\{r'_i\}) \exp(-\beta W_N\{r'_i\}) \prod_{i=1}^N d^3 r'_i}{\int \exp(-\beta W_N\{r'_i\}) \prod_{i=1}^N d^3 r'_i}. \quad (\text{B.7})$$

The canonical average symbolized by $\langle \rangle$ may be interpreted as follows. Imagine that the system of N particles, with potential energy $W_N\{r'_i\}$, and which is in thermodynamic equilibrium, is suddenly frozen into one of its possible equilibrium con-

APPENDIX B. WIDOM'S FORMULA

figurations. Now add to the frozen system another particle, identical to the others, and let it wander amongst the fixed particles, measuring at each point the value of $\exp(-\beta U\{r'_i\})$. The average value of this function is represented by $\langle \rangle$.

Using our definition in Eq. (B.7) we can write Eq. (B.6) as

$$\frac{Z_{N+1}}{Z_N} = \left(\frac{V}{NV_F} \right) \langle \exp(-\beta U) \rangle. \quad (\text{B.8})$$

We can insert Eq. (B.8) into Eq. (B.2) to obtain

$$\mu = -kT \ln \left[\left(\frac{V}{NV_F} \right) \langle \exp(-\beta U) \rangle \right]. \quad (\text{B.9})$$

Now we use the definition of the volume fraction $\phi \equiv N\Omega_p/V$ —where Ω_p is the volume of one particle—to rewrite Eq. (B.9) as

$$\mu = \mu_0 + kT \ln \phi - kT \ln \langle \exp(-U/kT) \rangle, \quad (\text{B.10})$$

where

$$\mu_0 \equiv -kT \ln(\Omega_p/V_F), \quad (\text{B.11})$$

is the standard part of the chemical potential. Equation (B.10) is directly analogous to Eq. (2.11) used in Sec. 2.2.



Appendix C

Chemical Potential Extrapolation

In this Appendix we derive Eq. (2.20), namely

$$\frac{\partial \hat{\mu}}{\partial \hat{\epsilon}} = \frac{1}{\epsilon} \left(\frac{\partial \bar{E}_{int}}{\partial N_p} \right)_{V,T}, \quad (\text{C.1})$$

which is used in Sec. 2.2 to derive the chemical potential extrapolation formula (Eq. (2.22)).

The chemical potential is given by Eq. (2.16):

$$\hat{\mu} \equiv \ln \phi - \ln \langle \exp(\nu \hat{\epsilon}) \rangle. \quad (\text{C.2})$$

Here ν is the number of new contacts made by the test protein, given by (see Eq. (2.14))

$$\nu \equiv N_{con}(N_p + 1) - N_{con}(N_p). \quad (\text{C.3})$$

The canonical average of the second term in Eq. (C.2) may be written explicitly as (cf. Eq. (B.7))

$$\begin{aligned} \langle \exp(\nu \hat{\epsilon}) \rangle &= \langle \exp\{[N_{con}(N_p + 1) - N_{con}(N_p)]\hat{\epsilon}\} \rangle \\ &= \frac{\int \exp[N_{con}(N_p + 1)\hat{\epsilon}] \prod_{i=1}^{N_p} d^3 r'_i}{\int \exp[N_{con}(N_p)\hat{\epsilon}] \prod_{i=1}^{N_p} d^3 r'_i}. \end{aligned} \quad (\text{C.4})$$

Recall that N_{con} is a function of the positions of the particles.

Substituting Eq. (C.4) into Eq. (C.2) gives

$$\hat{\mu} = \ln \phi - \ln \left\{ \int \exp[N_{con}(N_p + 1)\hat{\epsilon}] \prod_{i=1}^{N_p} d^3 r'_i \right\} + \ln \left\{ \int \exp[N_{con}(N_p)\hat{\epsilon}] \prod_{i=1}^{N_p} d^3 r'_i \right\}. \quad (C.5)$$

Differentiating Eq. (C.5) with respect to $\hat{\epsilon}$ we obtain

$$\frac{\partial \hat{\mu}}{\partial \hat{\epsilon}} = -(\langle N_{con}(N_p + 1) \rangle - \langle N_{con}(N_p) \rangle) = -\frac{\partial \langle N_{con}(N_p) \rangle}{\partial N_p}. \quad (C.6)$$

The average interaction energy \bar{E}_{int} is given by $\bar{E}_{int} = -\langle N_{con}(N_p) \rangle \epsilon$ (see Eq. (2.3)). Thus Eq. (C.6) may be written as

$$\frac{\partial \hat{\mu}}{\partial \hat{\epsilon}} = \frac{1}{\epsilon} \left(\frac{\partial \bar{E}_{int}}{\partial N_p} \right)_{V,T}. \quad (C.7)$$

Equation (C.7) is the desired result.

Appendix D

Flowchart of the Monte Carlo Algorithm

In this Appendix we present the flowcharts which schematically describe the Monte Carlo algorithm developed by Dr. Aleksey Lomakin (see also Sec. 2.2).

Figure D-1 illustrates the main operations of the algorithm. These are (the numbers refer to the boxes in the figure):

1. The initial configuration is set up. The number of particles N determines the value of ρ for the simulation as we fix the diameter of the particles beforehand. The variable SUM accumulates the sum of the exponents $\exp(\nu\hat{\epsilon})$, where $\hat{\epsilon}$ is the reduced energy and ν is the number of contacts made by the test particle (see Eq. (2.15)). The quantity $\exp(\nu\hat{\epsilon})$ is used to calculate the chemical potential (cf. Eq. (2.16)). The variable K keeps track of the number of attempts made to add a test particle.
2. A particle (the i^{th} particle) is chosen at random. The number of contacts η_1 made by this particle is calculated and stored. It is by accumulating information on the number of contacts made by each particle that $\bar{\eta}$, the average number of contacts per particle, is calculated. Recall that this quantity appears in the extrapolation formula for the chemical potential Eq. (2.22).
3. A position X,Y,Z in the simulation box is selected at random.

The main procedure of the Monte Carlo algorithm

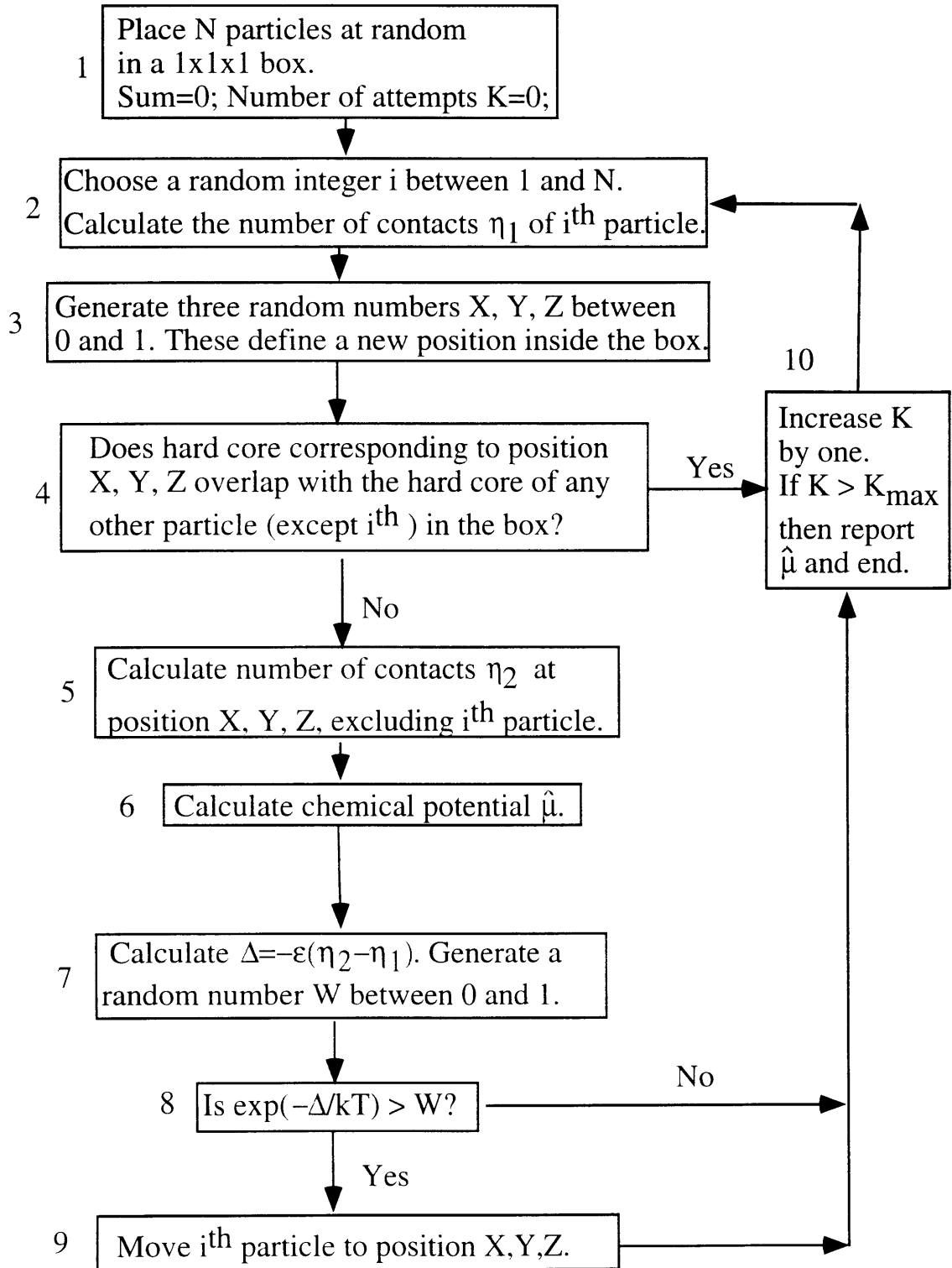


Figure D-1: A flowchart for the main routine of the Monte Carlo algorithm.

4. An attempt is made to add a test particle to the the selected position. If there is an overlap with the hard core of any particle (except the i^{th} particle of step 2) the attempt to add the particle is deemed unsuccessful and the program returns to step 2 (via step 10—see below).
5. The number of contacts made by the test particle η_2 is calculated. The i^{th} particle is excluded from this calculation.
6. The chemical potential is calculated. For lack of space, this step is shown in detail in Fig. D-2 and explained below.
- 7.-9. The Metropolis equilibration algorithm. The difference in energy Δ between the test particle and the i^{th} particle is calculated. If this is negative then the condition in step 8 will always be true and the i^{th} particle is moved to the position of the test particle. If $\Delta > 0$ then the move is accepted with a probability of $\exp(-\Delta/kT)$.
10. After both an unsuccessful addition of the test particle (at step 4) or a successful addition which leads to the Metropolis algorithm (steps 7-9) the variable K , which records the number of attempts to add a test particle, is increased by one. If the total number of attempts K to add the test particle is greater than a specified maximum K_{max} (typically 10^7), the program reports the value of the chemical potential and terminates.

Figure D-2 illustrates the calculation of the chemical potential in step 6. The steps in this procedure are (the numbers refer to the boxes in the figure):

1. The procedure first program checks if the number of attempts K to add a test particle is greater than a specified minimum K_0 (typically, 10^4). If it is, then the procedure continues in its calculation of the chemical potential. If $K < K_0$ then the chemical potential is not calculated and the routine continues with a Metropolis equilibration algorithm (steps 7-9). This is done to ensure that the chemical potential calculations are not biased by the choice of initial conditions.

Step 6: calculation of the chemical potential

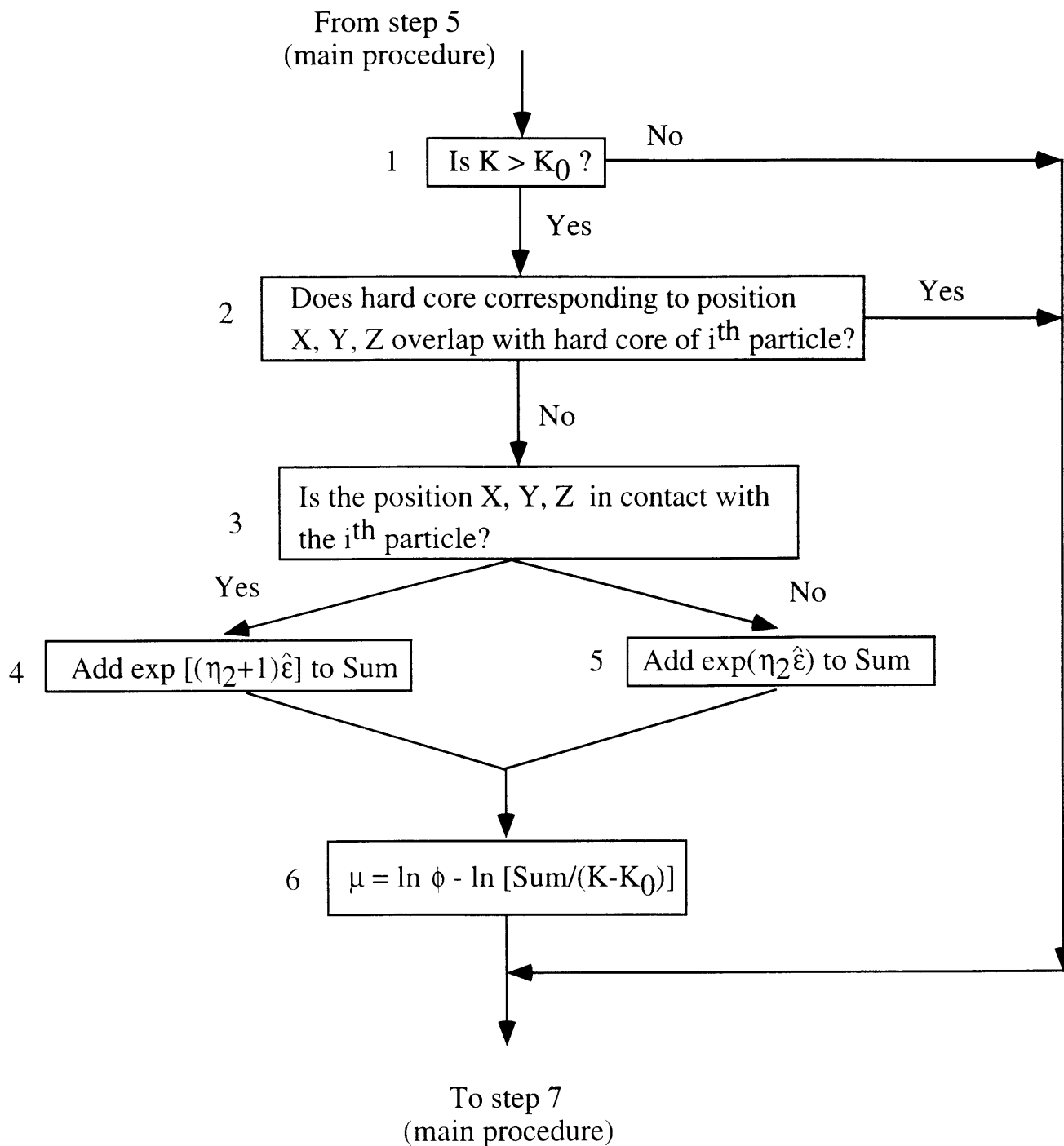


Figure D-2: A flowchart for the calculation of the chemical potential.

APPENDIX D. FLOWCHART OF THE MONTE CARLO ALGORITHM

2. A check is made to see if the hard core of the test particle overlaps with that of the i^{th} particle chosen previously. If it does, the program returns to the main procedure (the equilibration algorithm in steps 7–9) and there is no contribution to SUM. If there is no overlap, the chemical potential calculation continues.
3. A check is made to see if the test particle and i^{th} particle are in contact.
4. If the two particles are in contact, then the total number of contacts made by the test particle ν is $\eta_2 + 1$ (η_2 is the number of contacts of the test particle as determined in step 5 of the main procedure). SUM is increased by $\exp[(\eta_2 + 1)\hat{\epsilon}]$.
5. If the two particles are not in contact, then the total number of contacts made by the test particle ν is η_2 . SUM is increased by $\exp(\eta_2\hat{\epsilon})$.
6. After either step 4 or 5 the chemical potential is calculated. Note that once the system is in equilibrium, then $\text{SUM}/(K - K_0)$ is equal to $\langle \exp(\nu\hat{\epsilon}) \rangle$ (cf. Eq. (2.16)). The program then returns to the equilibration algorithm (steps 7–9 of the main procedure).

The savings in computational time with respect to conventional algorithms comes from using the position of the test particle both as a means to calculate the chemical potential (in step 6 of the main procedure) and as a way to equilibrate the system (steps 7–9 in the main procedure).

Appendix E

The Number of Contacts

In this Appendix we derive Eq. (2.32), which describes the low ϕ behavior of $\bar{\eta}$, the average number of contacts per particle. The result is [65]

$$\bar{\eta} = 8(\lambda^3 - 1) \exp(\hat{\epsilon})\phi. \quad (\text{E.1})$$

Equation (E.1) is derived as follows. In the low ϕ limit we may assume that the particles make contacts independently of each other. That is, the probability two particles are in contact is independent of the presence of any third particle. Thus $\bar{\eta}$, the average number of contacts made by each particle, may be written as

$$\bar{\eta} = p_{con}N. \quad (\text{E.2})$$

where p_{con} is the probability that the particle makes one contact and N is the total number of particles in the system.

Now the probability p_{con} will be proportional to the Boltzmann factor $\exp(\hat{\epsilon})$, where $\hat{\epsilon} = \epsilon/kT$. for the attraction between particles increases the chance that they will be near one another. However, for particles to be in contact their centers must lie with a distance $\sigma \leq r < \lambda\sigma$ of each other (see Eq. (2.2)). Thus the probability that two particles are in contact is given by

$$p_{con} = \frac{v_{con}}{V} \exp(\hat{\epsilon}). \quad (\text{E.3})$$

Here V is the total volume of the system and v_{con} is the volume in which the center of a particle may be located while still being in contact with a given particle. For a contact the centers of two particles must lie within a distance $\sigma \leq r < \lambda\sigma$ of each other. Therefore v_{con} is given by

$$v_{con} = \frac{4\pi\sigma^3}{3}(\lambda^3 - 1) = 8\Omega_p(\lambda^3 - 1), \quad (\text{E.4})$$

with $\Omega_p \equiv \pi\sigma^3/6$, the volume of one particle, since the hard core radius of a particle is $\sigma/2$.

Therefore, substituting Eqs. (E.3) and (E.4) into Eq. (E.2) we obtain

$$\bar{\eta} = 8\Omega_p(\lambda^3 - 1) \frac{N}{V} \exp(\hat{\epsilon}). \quad (\text{E.5})$$

Using the definition of $\phi \equiv N\Omega_p/V$, this may be rewritten as

$$\bar{\eta} = 8(\lambda^3 - 1) \exp(\hat{\epsilon})\phi. \quad (\text{E.6})$$

which is the desired result.

Appendix F

The Chemical Potential of the Solid

In this Appendix we present a derivation of the chemical potential of the solid as given by Eq. (3.2) in the context of a simple cell model [79]. The solid consists of N identical particles, in a volume V , at temperature T . This is analogous to the representation we used for the liquid phase (see Sec. 2.2). The equivalence of the thermodynamic descriptions of this system and that of the protein-water solution is discussed in Appendix A.

As stated in Section 3.2, we will assume that the particles interact in a pairwise additive fashion through a square-well potential of the form

$$u(r) = \begin{cases} +\infty, & \text{for } r < \sigma \\ -\epsilon, & \text{for } \sigma \leq r < \lambda\sigma \\ 0, & \text{for } r \geq \lambda\sigma \end{cases} \quad (\text{F.1})$$

The equation to be derived is

$$\mu_s = \mu_0 - n_s(\epsilon/2) - kT \ln[(\lambda - 1)^3]. \quad (\text{F.2})$$

The cell model is based on the idea that each molecule in a solid is confined by its neighbors. We picture the neighboring molecules as forming a cage or cell in which

the central molecule moves. The size of the cell at equilibrium is determined by the external pressure p , through $(\partial F/\partial V)|_{N,T} = -p$, where F is the Helmholtz free energy of the system. As we discussed in Sec. 3.2, we will neglect the small compressibility of the solid and calculate the partition function, and hence the chemical potential of the solid, at zero pressure.

In the simple cell model which we adopt, these are the fundamental assumptions:

1. The solid consists of a regular lattice of particles.
2. Each particle moves independently in a cell, caged in by its neighbors, in a potential provided by the neighbors when they are all at their lattice positions.

Since the particles are assumed to move independently, the partition function of the whole system $Z(N, V, T)$ is

$$Z(N, V, T) = [Z_1(V, T)/V_F]^N. \quad (\text{F.3})$$

where V_F is the Fermi volume and $Z_1(V, T)$ is the configurational integral for one particle. The configurational integral is defined as

$$Z_1(V, T) \equiv \int \exp[-w(\rho)/kT] d^3\rho. \quad (\text{F.4})$$

Here $w(\rho)$ is the potential provided by the neighbors and ρ is the radial distance a particle moves from the center of its cell with all of its neighbors remaining at their lattice positions. Note that, following the approximations made by Lennard-Jones and Devonshire, we have written the potential energy $w(\rho)$ in its spherically averaged form. It has been shown the error introduced by this approximation is small [79]. The integration in Eq. (F.4) is to be taken throughout the interior of the cell.

If we can evaluate Z_1 , we can calculate the partition function and hence the chemical potential of the solid. To do so we need to (i) specify the geometry of the integration and (ii) represent the quantity $w(\rho)$ so as to calculate Z_1 . For convenience, we shall assume that the lattice of particles is a face-centered cubic lattice i.e., each particle has twelve nearest neighbors. Our derivation is easily modified for other

geometries. We will take distance between adjacent particles (the lattice spacing) to be $\alpha\sigma$. Since the particles are spheres of diameter σ , α is simply the ratio of the lattice spacing to the diameter of the particle. Clearly, $\alpha > 1$. We will determine the equilibrium value of α by minimizing the free energy of the solid at zero pressure.

We are now going to compute the configurational integral Z_1 (see Eq. (F.4)) as a function of α . We will start from the close-packed solid with $\alpha = 1$ and then let α increase monotonically. The pair potential is assumed to be short-ranged, so that only nearest-neighbors interact. When a particle is at the center of its cell it is in contact with all of its neighbors. Recall (see Sec. 2.1) that two particles are said to be in contact if their centers lie within a distance $\sigma \leq r < \lambda\sigma$ of each other. Thus $w(\rho = 0)$, the potential energy of a particle when it and its neighbors are at their cell centers, may be written as

$$w(\rho = 0) = -n_s\epsilon/2, \tag{F.5}$$

with n_s the number of nearest neighbors. For our geometry $n_s = 12$, but we leave n_s as a general quantity since this derivation will be applicable to other lattice geometries.

Note also that for this geometry of the solid, it is the hard core repulsion of the neighbors which limits the displacement of particles in their cells. In Fig. F-1A we show the solid for the case where $\alpha \approx 1$. The large solid circles represent the hard cores of the particles, diameter σ . The small solid circle encloses the region where the center of the central particle may move without overlapping the hard core of a neighbor. Recall that, due to the hard core repulsion, the closest the center of the central particle may be to the center of any of its neighbors is a distance σ . We call the region enclosed by the small solid circle the “sphere of excursion”. This sphere has a radius of $(\alpha - 1)\sigma$. The dashed circle encloses the region where the central particle could move and still be in contact with all of its neighbors. We call this region the “sphere of attraction”. This sphere has a radius of $(\lambda - \alpha)\sigma$.

In Fig. F-1B the sphere of excursion and sphere of contact are shown separately. The particle can only move in the sphere of excursion. Therefore, the hatched region

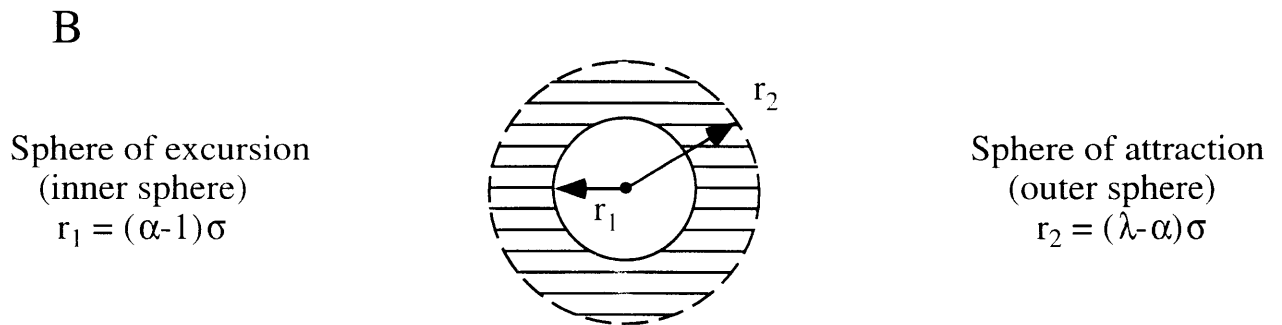
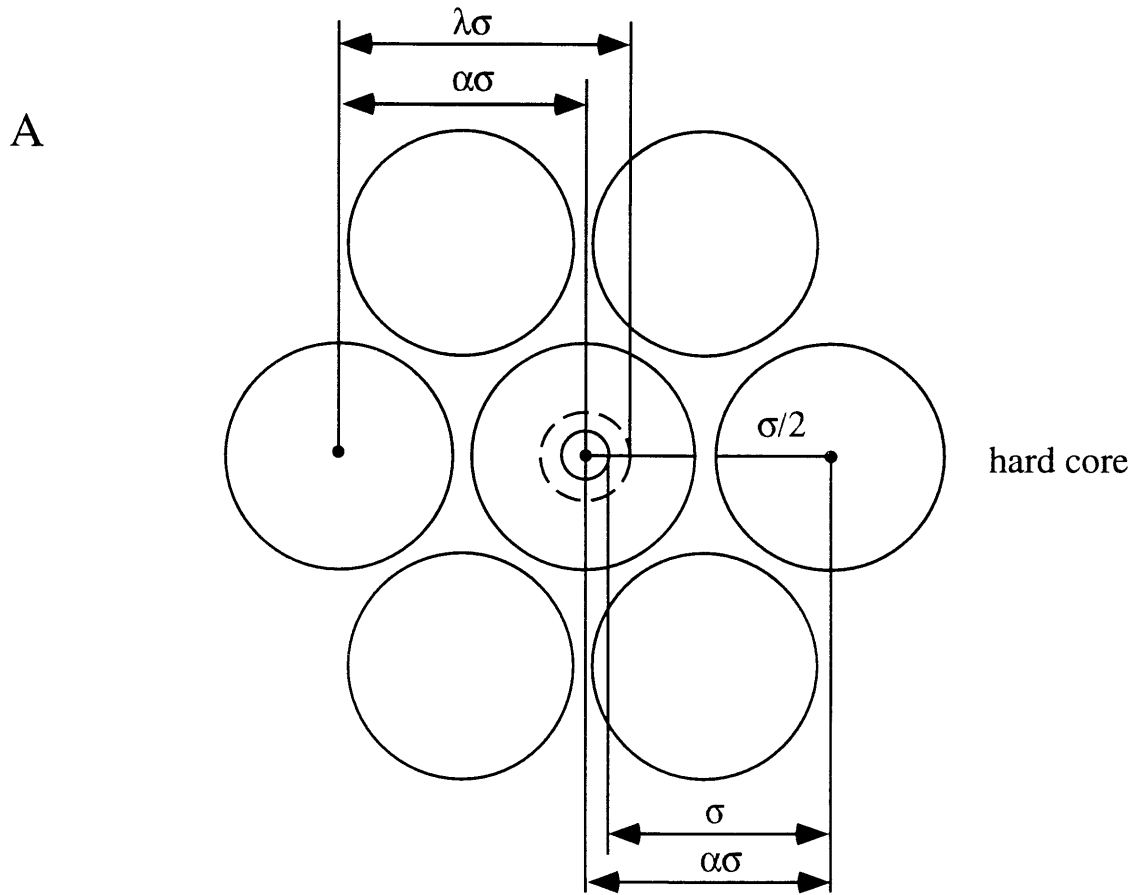


Figure F-1: (A) The solid for the case where $\alpha \approx 1$ and $\lambda > 1$. The large solid circles represent the hard cores of the particles. The small solid circle encloses the region where the center of the central particle may move without overlapping the hard core of a neighbor ("sphere of excursion"). The dashed circle encloses the region where the central particle could move and still be in contact with all of its neighbors ("sphere of attraction"); (B) The sphere of excursion and sphere of attraction shown separately. The hatched region is excluded to the particle for it cannot move beyond the sphere of excursion.

is forbidden to the particle. Since the sphere of attraction is greater than that of excursion, the particle makes all n_s contacts throughout the sphere of excursion. Therefore $w(\rho) = -n_s\epsilon/2$ for $\rho < (\alpha - 1)\sigma$. Hence for the domain of α shown in Fig. F-1 the configurational integral is easily calculated from Eq. (F.4) as

$$Z_1^+ = 4\pi \int_0^{(\alpha-1)\sigma} \exp[-w(\rho)/kT] \rho^2 d\rho = \frac{4\pi}{3} \sigma^3 (\alpha - 1)^3 \exp(n_s \hat{\epsilon}/2), \quad (\text{F.6})$$

where we have used the notation $\hat{\epsilon} \equiv \epsilon/kT$. Here the superscript + is used to denote the case where the sphere of excursion is smaller than the sphere of attraction. This domain applies as α increases relative to λ until the point is reached that the sphere of excursion attains the same radius as the sphere of attraction. Since the radii of these sphere are, respectively, $(\alpha - 1)\sigma$ and $(\lambda - \alpha)\sigma$, the condition for congruence is $\alpha - 1 = \lambda - \alpha$ or $\alpha = (\lambda + 1)/2$. Therefore the upper limit of validity of Eq. (F.6) is $\alpha = (\lambda + 1)/2$.

We now consider the domain where $\alpha > (\lambda + 1)/2$. The geometry of this domain is illustrated in Figs. F-2A and B. The symbols are defined as for Fig. F-1. In this domain the sphere of attraction has a smaller radius than that of the sphere of excursion. As a result, the configurational integral has a different dependence on α than for the domain $\alpha < (\lambda + 1)/2$. Within the sphere of attraction ($\rho < (\lambda - \alpha)\sigma$), the particle makes all of its attractive contacts and thus $w(\rho) = -n_s\epsilon/2$. In the region outside the sphere of attraction, but within the sphere of excursion (the gray area in Fig. F-2B), the particle will lose of some of its contacts. Therefore in this region we may write $w(\rho) = -n_s\epsilon/2 + \kappa\epsilon$, where $\kappa > 0$, and $\kappa\epsilon$ represents the average attractive energy lost in this region.

The configurational integral in the $\alpha > (\lambda + 1)/2$ domain contains two terms, corresponding to the white and gray regions in Fig. F-2B. They are (using Eq. (F.4))

$$Z_1^- = 4\pi \exp[n_s \hat{\epsilon}/2] \int_0^{(\lambda-\alpha)\sigma} \rho^2 d\rho + 4\pi \exp[n_s \hat{\epsilon}/2] \exp[-\kappa \hat{\epsilon}] \int_{(\lambda-\alpha)\sigma}^{(\alpha-1)\sigma} \rho^2 d\rho. \quad (\text{F.7})$$

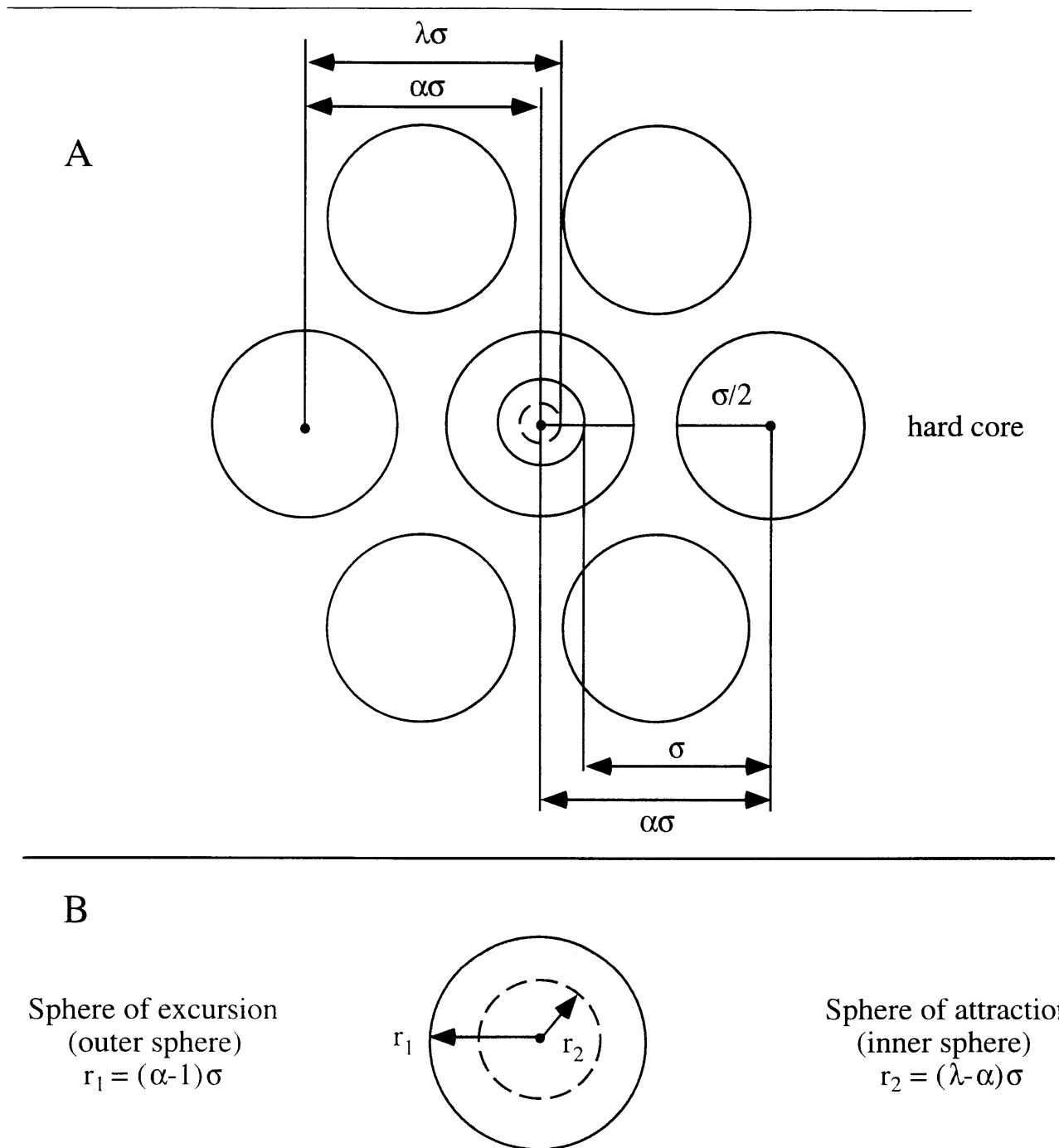


Figure F-2: (A) The solid for the case where $\alpha - 1 > \lambda - \alpha$, i.e., $\alpha > (\lambda + 1)/2$. The large solid circles represent the hard cores of the particles. The small solid circle represents the sphere of excursion. The dashed circle represents the sphere of attraction; (B) The sphere of excursion and sphere of attraction shown separately. In the gray region the particle does not feel the attraction of all of its neighbors.

Here the superscript $-$ is used to denote the regime $\alpha > (\lambda + 1)/2$.

We are interested in the short-range ($\lambda \rightarrow 1$) and hence deep well ($\hat{\epsilon} \rightarrow \infty$) limit of the pair potential, for this is limit which corresponds to the interaction of globular proteins. In this limit the second term in Eq. (F.7), which comes from the gray region in Fig. F-2, makes only a small contribution to the configurational integral (since it is smaller than the first term by a factor of $\sim \exp(-\kappa\hat{\epsilon})$) and Eq. (F.7) becomes approximately

$$Z_1^- = \frac{4\pi}{3} \sigma^3 (\lambda - \alpha)^3 \exp[n_s \hat{\epsilon}/2]. \quad (\text{F.8})$$

Substituting Eq. (F.6) or Eq. (F.8) together into Eq. (F.3) we obtain the partition function in the two domains. namely

$$Z^+(N, \alpha, T) = \left[\frac{4\pi}{3} \frac{\sigma^3}{V_F} (\alpha - 1)^3 \exp(n_s \epsilon/2kT) \right]^N. \quad (\text{F.9})$$

for $\alpha \leq (\lambda + 1)/2$ and

$$Z^-(N, \alpha, T) = \left[\frac{4\pi}{3} \frac{\sigma^3}{V_F} (\lambda - \alpha)^3 \exp(n_s \epsilon/2kT) \right]^N. \quad (\text{F.10})$$

for $\alpha > (\lambda + 1)/2$.

The behavior of Z as a function of α is shown in Fig. F-3. It can be seen that the partition function is a maximum i.e., the free energy is a minimum at $\alpha = (\lambda + 1)/2$. Thus, at zero pressure, the solid will adopt a cell size of $\alpha = (\lambda + 1)/2$. This is the cell size for which the sphere of excursion and sphere of attraction are equal in size. Note that $\partial Z/\partial \alpha$ is discontinuous at $\alpha = (\lambda + 1)/2$. This is because the pair potential is not a smooth function of particle position and because we have neglected correlations between the motions of particles.

Therefore, using our approximations, the partition function of the solid is

$$Z(N, p = 0, T) = \left[\frac{4\pi}{3} \frac{\sigma^3}{V_F} \left(\frac{\lambda - 1}{2} \right)^3 \exp(n_s \epsilon/2kT) \right]^N. \quad (\text{F.11})$$

This result is obtained by substituting $\alpha = (\lambda + 1)/2$ into either Eq. (F.9) or Eq. (F.10).

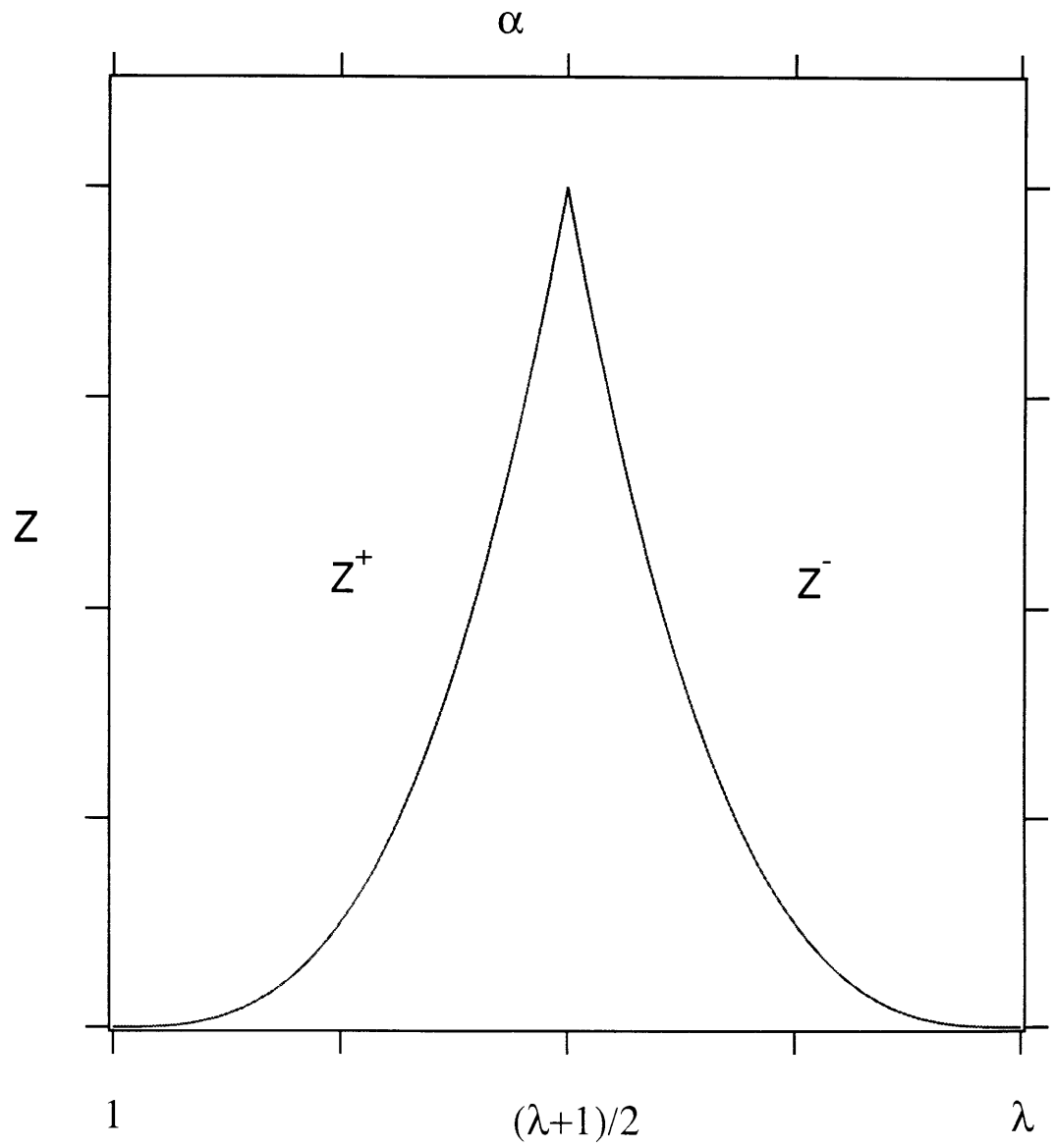


Figure F-3: A schematic illustration of the partition function Z as a function of cell size α in the $\hat{\epsilon} \rightarrow \infty$ limit.

Since we are working at constant pressure (and $p = 0$), the free energy of the solid, given by $G = -kT \ln Z$, is the Gibbs free energy. From Eq. (F.11) we have

$$G(N, p = 0, T) = -NkT \ln\left(\frac{\Omega_p}{V_F}\right) - Nn_s(\epsilon/2) - NkT \ln[(\lambda - 1)^3] \quad (\text{F.12})$$

The chemical potential of the solid μ_s is $\mu_s = G/N$. Using Eq. (F.12), we write

$$\mu_s = -kT \ln\left(\frac{\Omega_p}{V_F}\right) - n_s(\epsilon/2) - kT \ln[(\lambda - 1)^3]. \quad (\text{F.13})$$

Here $\Omega_p = \pi\sigma^3/6$ is the volume of one particle. We now may define $\mu_0 \equiv -kT \ln(\Omega_p/V_F)$ as the standard chemical potential of the solid and we note that it is equal to the standard part of the chemical potential of the liquid (see Eq. (B.11)). Then Eq. (F.13) may be rewritten as

$$\mu_s = \mu_0 - n_s(\epsilon/2) - kT \ln[(\lambda - 1)^3]. \quad (\text{F.14})$$

This equation is the same as Eq. (3.2). As we have discussed in Sec. 3.2, the first term on the right-hand side of Eq. (F.14) is the standard part of the chemical potential. The second term is the total energy associated with each particle and the last term is the entropic contribution. It is a measure of the volume accessible to the center of mass of the particle.

Addendum: The Volume Fraction of the Solid

We may obtain ϕ_s , the volume fraction of the solid at equilibrium, as follows.

The volume fraction ϕ_s is given by

$$\phi_s = \Omega_p/v_s. \quad (\text{F.15})$$

where $v_s \equiv V/N$ is the volume per particle in the solid.

At close packing we have

$$\phi_{cp} = \Omega_p/v_{cp}. \quad (\text{F.16})$$

where $\phi_{cp} = 0.7405$ is the volume fraction of closed-packed hard spheres and v_{cp} is the corresponding volume per particle. Thus from Eqs. (F.15) and (F.16) we write

$$\phi_s = \phi_{cp}(v_{cp}/v_s). \quad (\text{F.17})$$

Now from geometrical considerations $v_s/v_{cp} = \alpha^3$, since the volume per particle scales as the cube of the lattice spacing, and for $\alpha = 1$, $v_s = v_{cp}$. Therefore from Eq. (F.17) the volume fraction of the solid, in terms of α , is

$$\phi_s = \phi_{cp}/\alpha^3 \quad (\text{F.18})$$

As we showed above, the free energy of the solid is a minimum at $\alpha = (\lambda + 1)/2$. Therefore the corresponding volume fraction is

$$\phi_s = 8\phi_{cp}/(\lambda + 1)^3, \quad (\text{F.19})$$

and since we have neglected the compressibility (or equivalently the thermal expansivity) of the solid, ϕ_s is independent of temperature.

In Chapter 2 we saw that in order to obtain the experimentally observed value of the critical volume fraction for the γ -crystallins λ should be 1.25. Using this value in Eq. (F.19), together with $\phi_{cp} = 0.7405$, we find $\phi_s = 0.520 = 2.5\phi_c$, as is shown in Fig. 3-3. This is smaller than the value found experimentally for the γ -crystallins, $\phi_s = 0.62 = 3.0\phi_c$ (see Fig. 3-1). We believe that the difference is a reflection of the fact that the real interactions between the proteins are anisotropic. When we assume that the interactions are isotropic, we are neglecting orientational degrees of freedom. These would alter the entropy term in Eq. (F.14) and affect the position of the liquidus and solidus lines.

Bibliography

- [1] C.K. Mathews and K.E. van Holde, *Biochemistry* (Benjamin/Cummings, Redwood City, 1990), Chapter 6, pp. 185–190.
- [2] M.L. Broide, C.R. Berland, J. Pande, O.O. Ogun, and G.B. Benedek, *Proc. Natl. Acad. Sci. USA* **88**, 5660 (1991).
- [3] C.R. Berland, G.M. Thurston, M. Kondo, M.L. Broide, J. Pande, O.O. Ogun, and G.B. Benedek, *Proc. Natl. Acad. Sci. USA* **89**, 1214 (1992).
- [4] J.A. Thomson, P. Schurtenberger, G.M. Thurston, and G.B. Benedek, *Proc. Natl. Acad. Sci. USA* **84**, 7079 (1987).
- [5] M.L. Broide, T. Tominc, and M. Saxowsky, *Phys. Rev. E* **53**, 6325 (1996).
- [6] D. Rosenbaum, P.C. Zamora, and C.F. Zukoski, *Phys. Rev. Lett.* **76**, 150 (1996); D. Rosenbaum and C.F. Zukoski, *J. Cryst. Growth* **169**, 752 (1996).
- [7] M. Muschol, M.R. Banish, B.R. Thomas, and F. Rosenberger in *Photon Correlation and Scattering*, Vol. 14, 1996 OSA Technical Digest Series (Optical Society of America, Washington DC, 1996), pp. 145-147; M. Muschol and F. Rosenberger, *J. Chem. Phys.* **107**, 1953 (1997).
- [8] V.A. Rabinovich, A.A. Vasserman, V.I. Nedostup, and L.S. Veksler, *Thermophysical Properties of Neon, Argon, Krypton, and Xenon* (Hemisphere, Washington, 1988), pp. 380-383.
- [9] O.G. Peterson, D.N. Batchelder, and R.O. Simmons, *Phys. Rev.* **150**, 703 (1966).
- [10] P. Flubacher, A.J. Leadbetter, and J.A. Morrison, *Proc. Phys. Soc.* **78**, 1449 (1961).
- [11] J.P.K. Doye and D.J. Wales, *J. Phys. B* **29**, 4859 (1996).
- [12] J.O. Hirschfelder, C.F. Curtiss, and R.B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 2nd printing, 1964), p.160.

BIBLIOGRAPHY

- [13] P.C. Weber in *Advances in Protein Chemistry* **41**, eds. C.B. Anfinsen, J.T. Edsall, F.M. Richards, D.S. Eisenberg (Academic Press, San Diego, 1991).
- [14] G.B. Benedek. *Invest. Ophthalmol. Vis. Sci.* **38**, 1911 (1997).
- [15] V.B. Tolstoguzov. *Food Hydrocolloids* **2**, 339 (1988).
- [16] H. Walter and D.E. Brooks. *FEBS Letters* **361**, 135 (1995).
- [17] C.K. Mathews and K.E. van Holde, *Biochemistry* (Benjamin/Cummings, New York, 1990), Chapter 2. pp. 30–40.
- [18] A. Lomakin. N. Asherie. and G.B. Benedek, *J. Chem. Phys.* **104**, 1646 (1996).
- [19] N. Asherie, A. Lomakin. and G.B. Benedek, *Phys. Rev. Lett.* **77**, 4832 (1996).
- [20] H. Verduin and J.K.G. Dhont. *J. Colloid Interface Sci.* **172**. 425 (1995).
- [21] S.M. Ilett. A. Orrock. W.C.K. Poon. and P.N. Pusey. *Phys. Rev. E* **51**, 1344 (1995).
- [22] J.A. Glasel and M.P. Deutscher (Eds.), *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Academic Press, San Diego, 1995).
- [23] J. Pande. A. Lomakin. B. Fine. O. Ogun, I. Sokolinski, and G.B. Benedek, *Proc. Natl. Acad. Sci. USA* **92**. 1067 (1995).
- [24] W.C.K. Poon. *Phys. Rev. E* **55**. 3762 (1996).
- [25] J. Pande, C. Berland. M. Broide, O. Ogun, J. Melhuish, and G. Benedek, *Proc. Natl. Acad. Sci. USA* **88**. 4916 (1991).
- [26] G.B. Benedek. in *Human Cataract Formation*. CIBA Foundation Symposium 106 (Pitman Publication Ltd., London, 1984) p.237.
- [27] W.A. Eaton and J. Hofrichter. in *Advances in Protein Chemistry*, edited by J.T. Edsall, F.M. Richards. and D.S. Eisenberg (Academic Press, San Diego, 1990), Vol. 40 p. 63; W.A. Eaton and J. Hofrichter, *Science* **268**. 1142 (1995).
- [28] G.G. Glenner. *The New England Journal of Medicine* **302**. 1283 (1980); **302**. 1333 (1980).
- [29] G.B. Benedek. *Appl. Opt.* **10**. 459 (1971).
- [30] R.J. Siezen. M.R. Fisch. C. Slingsby. and G.B. Benedek. *Proc. Natl. Acad. Sci. USA* **82**. 1701 (1985).
- [31] P. Schurtenberger. R.A. Chamberlin. G.M. Thurston. J.A. Thomson, and G.B. Benedek, *Phys. Rev. Lett.* **63**. 2064 (1989).
- [32] M.H.G.M. Penders and A. Vrij, *Adv. Colloid Interface Sci.* **36**, 185 (1991).

BIBLIOGRAPHY

- [33] M.H.G.M. Penders, A. Vrij, and R. Van der Haegen, *J. Colloid Interface Sci.* **144**, 86 (1991).
- [34] C. Liu., A. Lomakin, G.M. Thurston, D. Hayden, A. Pande, J. Pande, O. Ogun, N. Asherie, and G.B. Benedek, *J. Phys. Chem.* **99**, 454 (1995).
- [35] C. Liu., N. Asherie, A. Lomakin, J. Pande, O. Ogun, and G.B. Benedek, *Proc. Natl. Sci. USA* **93**, 377 (1996).
- [36] F. Reif, *Fundamentals of Statistical and Thermal Physics*, (McGraw-Hill Book Company, New York, 1965), Chapter 8, pp. 306–312 and Chapter 10, pp. 422–428.
- [37] J.K. Percus and G.J. Yevick, *Phys. Rev.* **110**, 1 (1958).
- [38] R.J. Baxter, *J. Chem. Phys.* **49** 2770 (1968).
- [39] R.O. Watts, D. Henderson, and R.J. Baxter, *Adv. Chem. Phys.* **21**, 421 (1971).
- [40] Y. Song and E.A. Mason. *J. Chem Phys.* **91**. 7840 (1989).
- [41] J.P. Hansen and I.R. McDonald. *Theory of Simple Liquids*, 2nd ed. (Academic Press, London, 1990). Chapters 5 and 6.
- [42] F. del Rio and L. Lira. *Mol. Phys.* **61**, 275 (1986).
- [43] V.G. Taratuta. A. Holschbach, G.M. Thurston, D. Blankschtein, and G.B. Benedek, *J. Phys. Chem* **94**. 2140 (1990).
- [44] A. Rotenberg. *J. Chem. Phys.* **43**. 1198 (1965).
- [45] D. Henderson. O.H. Scalise. and W.R. Smith, *J. Chem. Phys.* **72**, 2431 (1980).
- [46] W.G. Kranendonk and D. Frenkel, *Mol. Phys.* **64**. 403 (1988).
- [47] L. Vega, E. de Miguel. L.F. Rull. G. Jackson, and I.A. McLure, *J. Chem. Phys.* **96**, 2296 (1992).
- [48] M.H.J. Hagen and D. Frankel. *J. Chem. Phys.* **101**, 4093 (1994).
- [49] E. Lomba and Noé G. Almarza. *J. Chem. Phys.* **100**, 8367 (1994).
- [50] A.Z. Panagiotopoulos. *Mol. Phys.* **61**. 813 (1987).
- [51] A.Z. Panagiotopoulos. N. Quirke. M. Stapelton, and D.J. Tildesley. *Mol. Phys.* **63**, 527 (1988).
- [52] E.A. Guggenheim. *Thermodynamics*, 4th ed. (North-Holland Publishing Company, Amsterdam, 1959). Chapter 4, pp. 156–160; J. Clerk-Maxwell, *Nature* **11**, 357 (1875).

BIBLIOGRAPHY

- [53] I. Prigogine and R. Defay, *Chemical Thermodynamics* (Longmans, London, 1965), Chapter 16, pp. 237–246.
- [54] B. Widom, *J. Chem. Phys.* **39**, 2808 (1963).
- [55] N.F. Carnahan and K.E. Starling, *J. Chem. Phys.* **51**, 635 (1969).
- [56] M.P. Allen and D.J. Tildesley, *Computer Simulations of Liquids* (Clarendon, Oxford, 1989), Chapter 1, pp. 24–27.
- [57] N.A. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [58] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, London, 1996), Appendix C, pp. 363–376.
- [59] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, London, 1996), Chapter 3, pp. 38–42.
- [60] M.P. Allen and D.J. Tildesley, *Computer Simulations of Liquids* (Clarendon, Oxford, 1987), Chapter 7, pp. 171–173.
- [61] F. del Rio and L. Lira, *J. Chem. Phys.* **87**, 7179 (1987).
- [62] L.E. Reichl, *A Modern Course in Statistical Physics* (Univ. of Texas Press, Austin, 1980), Chapter 11, pp. 362–367.
- [63] J.P. Hansen and I.R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic Press, London, 1990), Chapter 4.
- [64] B. Barbov, *J. Chem. Phys.* **61**, 3194 (1974).
- [65] C. Kittel and H. Kroemer, *Thermal Physics*, 2nd ed. (W.H. Freeman and Company, New York, 1980), Chapter 3, pp. 57–63.
- [66] Y.C. Chiew and E.D. Glandt, *J. Phys. A* **16**, 2599 (1983).
- [67] B.M. Fine, A. Lomakin, O.O. Ogun, and G.B. Benedek, *J. Chem. Phys.* **104**, 326 (1996).
- [68] D. H. Everett, *Basic Principles of Colloid Science* (Royal Society of Chemistry, London, 1988).
- [69] D. A. Young (Ed.), *Colloid Stability*, Faraday Discussions of the Chemical Society, Vol. 65 (Chemical Society, London, 1978).
- [70] A. Daanoun, C.F. Tejero, and M. Baus, *Phys. Rev. E* **50**, 2913 (1994).
- [71] C. Rascón, G. Navascués, and L. Mederos, *Phys. Rev. B* **51**, 14899 (1995).
- [72] F. Leal Calderon, J. Bibette, and J. Biais, *Europhys. Lett.* **23**, 653 (1993).

BIBLIOGRAPHY

- [73] P.N. Pusey, W.C.K. Poon, S.M. Ilett, and P. Bartlett, *J. Phys.* **6**, A29 (1994).
- [74] H.N.W. Lekkerkerker, W.C.K. Poon, P.N. Pusey, A. Stroobants, and P.B. Warren, *Europhys. Lett.* **20**, 559 (1992).
- [75] A.P. Gast, C.K. Hall, and W.B. Russel, *J. Colloid Interface Sci.* **96**, 251 (1983).
- [76] P.A. Heiney, J.E. Fischer, A.R. Mcghee, W.J. Romanow, A.M. Denenstein, J.P. Mccauley, A.B. Smith, D.E. Cox, *Phys. Rev. Lett.* **66**, 2911 (1991).
- [77] J.E. Lennard-Jones and A.F. Devonshire. *Proc. R. Soc. London* **A163**, 53 (1937).
- [78] J.P.K. Doye and D.J. Wales, *Science* **271**, 484 (1996).
- [79] J.A. Barker. *Lattice Theories of the Liquid State* (Pergamon Press, Oxford, 1963).
- [80] E.A. Guggenheim, *Thermodynamics*, 5th ed. (North-Holland, Amsterdam, 1967), pp. 220–224.
- [81] P.N. Pusey, W. van Meegen, P. Bartlett, B.J. Ackerson, J.G. Rarity, and S.M. Underwood. *Phys. Rev. Lett.* **63**, 2753 (1989).
- [82] W.C.K. Poon, J.S. Selfe, M.B. Robertson, S.M. Ilett, A.D. Pirie, and P.N. Pusey, *J. Phys. II* **3**, 1075 (1993).
- [83] B.M. Fine. Ph. D. thesis (Massachusetts Institute of Technology, Cambridge, 1994).
- [84] G. Stell, *J. Stat. Phys.* **63**, 1203 (1991).
- [85] N.J. Darby and T.E. Creighton. *Protein Structure* (IRL Press, Oxford, 1993).
- [86] G. Friberg, J. Pande, O. Ogun, G.B. Benedek. *Curr. Eye Res.* **15**, 1182 (1996).
- [87] J. Pande. (private communication).
- [88] J. Pande, O. Ogun, C. Nath and G.B. Benedek, *Exp. Eye. Res.* **57**, 257 (1993).
- [89] C. Liu, Ph. D. thesis (Massachusetts Institute of Technology, Cambridge, 1995).
- [90] F. Yi, B.M. Denker, and E.J. Neer, *J. Biol. Chem.* **266**, 3900 (1991).
- [91] L.L. Chen, J.J. Rosa, S. Turner, and R.B. Pepinsky, *J. Biol. Chem.* **266**, 18237 (1991).
- [92] R.B. Pepinsky, L.L. Chen, W. Meier, and B.P. Wallner, *J. Biol. Chem.* **266**, 18244 (1991).
- [93] S. S. Wong. *Chemistry of Protein Conjugation and Cross-linking* (CRC Press, Boca Raton, 1991), p. 108.

BIBLIOGRAPHY

- [94] F. Wold in *Methods in Enzymology* **25**, eds. C.H.W. Hirs and S.N. Timasheff, (Academic Press, New York, 1972), pp. 623-651.
- [95] Yu. N. Chirgadze, H.P.C. Driessen, G. Wright, C. Slingsby, R.E. Hay, and P.F. Lindley, *Acta Cryst.* **D52**, 712 (1996).
- [96] U.K. Laemmli, *Nature (London)* **227**, 680 (1971).
- [97] K. Weber and M. Osborn, *J. Biol. Chem.* **244** 4406 (1969).
- [98] P.F. Lindley, M.E. Narebor, L.J. Summers, and G.J. Wistow in *The Ocular Lens: Structure, Function, and Pathology* ed. Harry Maisel, (Marcel Dekker, New York, 1985), pp. 123-167.
- [99] T.G. Braginskaya, , P.D. Dobichin, M.A. Ivanova, V.V. Kublin, A.V. Lomakin, V.A. Noskin, G.E. Shmelev, and S.P. Tolpina, *Phys. Scripta* **28**, 73 (1983).
- [100] E.R. Pike in *Scattering Techniques Applied to Supramolecular and Nonequilibrium Systems*, eds. S.H. Chen, B. Chu, and R. Nossal (Plenum, New York, 1981) pp. 179-200.
- [101] R.J. Seizen, E.D. Kaplan, and R.D. Anello, *Biochem. Biophys. Res. Commun.* **127**, 153 (1985).
- [102] H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (New York, Oxford University Press, 1971).
- [103] E.A. Guggenheim. *Thermodynamics*, 5th ed. (North-Holland, Amsterdam, 1967), p. 160.
- [104] M. Tribus. *Thermostatistics and Thermodynamics* (Van Nostrand, Princeton, 1961), p. 216ff.
- [105] V.B. Tolstoguzov. *Food Hydrocolloids* **2**, 195 (1988).
- [106] M.P. Tombs, B.G. Newsom, and P. Wilding, *Int. J. Peptide Protein Res.* **6**, 253 (1974).
- [107] E. de Miguel, *Phys. Rev. E* **55**, 1347 (1997).
- [108] G. Jackson, W.G. Chapman, and K.E. Gubbins, *Mol. Phys.* **65**, 1 (1988).
- [109] A. Yethiraj and C.K. Hall, *Mol. Phys.* **72**, 619 (1990).
- [110] A. Yethiraj and C.K. Hall, *J. Chem. Phys.* **95**, 8494 (1991).
- [111] H. S. Gulati and C.K. Hall, *J. Chem. Phys.* **107**, 3930 (1997).
- [112] M. Banaszak, Y.C. Chiew, and M. Radosz, *Phys. Rev. E* **48**, 3760 (1993).
- [113] Yu.V. Kalyuzhnyi and P.T. Cummings, *Mol. Phys.* **87**, 249 (1996).

BIBLIOGRAPHY

- [114] M.S. Wertheim, *J. Chem. Phys.* **85**, 2929 (1986) and references therein.
- [115] D.J. Tildesley and W.B. Street **41**, 85 (1980).
- [116] W.G. Chapman, G. Jackson and K.E. Gubbins, *Mol. Phys.* **65**, 1057 (1988).
- [117] J. García de la Torre and V.A. Bloomfield, *Quarterly Review of Biophysics* **14**, 81 (1981).

BIOGRAPHICAL NOTE

Neer Ruben Asherie was born in New Rochelle, New York on December 29, 1971. He attended high school at the Sir James Henderson British School of Milan. There he completed the two-year "A" level program in Mathematics, Further Mathematics, Physics, and Chemistry in only one year graduating in June 1988. He was accepted to the University of Cambridge where he read Natural Sciences (Physical) at Gonville and Caius college. While there he captained the college chess club and received a First in all parts of his Tripos Examination. He specialized in mathematical physics and graduated with a B.A. in Natural Sciences in June 1991.

Since September 1991, Neer has been a graduate student in the Department of Physics at the Massachusetts Institute of Technology. For three years he was a research assistant in theoretical plasma physics for Professor Bruno Coppi, studying cyclotron radiation emitted from fusion plasmas. Since June 1994, he has been pursuing his doctorate degree under Professor George B. Benedek exploring the thermodynamics of protein solutions. He has studied mixtures of proteins both theoretically and experimentally and has helped to develop a Monte Carlo computer simulation to study protein liquid-liquid phase separation. He is currently investigating the role of the interaction range in shaping the phase diagrams of proteins and colloids.

Neer has been active in teaching at M.I.T. He has been a teaching assistant for graduate plasma physics courses as well as a recitation instructor for undergraduate mechanics. He was a teaching assistant for Professor Benedek in an undergraduate statistical mechanics course for which he was awarded the Buechner Teaching Prize in 1994.

Neer plans to obtain his Ph.D. in physics in October 1997 and thence become rich and famous. Plan B is to pursue a career in physics.