Title
Fluctuations in Water and their Relation to the Hydrophobic Effect

Author
Varilly, Patrick Stephen

Publication Date
2011

Fluctuations in Water and their Relation to the Hydrophobic Effect

by

Patrick Stephen Varilly

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David Chandler, Co-chair
Associate Professor Jan T. Liphardt, Co-chair
Associate Professor Phillip L. Geissler
Associate Professor Joel Moore

Fall 2011

Fluctuations in Water and their Relation to the Hydrophobic Effect

Abstract

Fluctuations in Water and their Relation to the Hydrophobic Effect

by

Patrick Stephen Varilly

Doctor of Philosophy in Physics

University of California, Berkeley

Professor David Chandler, Co-chair
Associate Professor Jan T. Liphardt, Co-chair

The hydrophobic effect, or the tendency for oil and water not to mix, is a fundamental force that strongly influences the shape, behavior and assembly of solutes in solution. Hydrophobicity emerges from the collective behavior of large numbers of solvent molecules, so its accurate treatment is challenging. A decade ago, Lum, Chandler and Weeks (LCW) addressed this challenge indirectly by modeling how solvent density fluctuations couple to external solutes and constraints, and then inferring hydrophobic behavior from the resulting mean solvent density. LCW theory is successful because it distinguishes between, and separately models, small-length-scale and large-length-scale density fluctuations. In this thesis, we develop methods for probing the statistics of large-length-scale density fluctuations in computer simulations of water. We use these tools to study solvation phenomena in model systems and in proteins, in bulk water and near surfaces, and we rationalize these phenomena in terms of LCW ideas. Building on these ideas and on past efforts by others, we construct a tractable, efficient and accurate theory of solvation on a coarse-grained lattice. The final theory allows us to model the solvation behavior of uncharged, static solutes of arbitrary shape, and we outline the steps necessary to model charged, dynamic solutes in the future. A unifying thread in our solvation studies is the importance of fluctuations of liquid-vapor interfaces. At the end of this thesis, we describe how these fluctuations may play a role in water evaporation.

To Laura

Contents

Acknowledgments

I owe an enormous debt of gratitude to everyone who has shaped me as a human being and as a researcher, and it is both an honor and a pleasure to be able to thank them here.

First and foremost, I would like to thank my family for their unconditional love, support and encouragement. My Dad, for having the patience and perseverance to teach me the better part of everything I know, and for showing me by example what it means to pursue the things you love. My Mom, for teaching me what is truly valuable, and who I wish could still be here to see the fruits of her labor. My big brother Tony, who I am lucky to count as one of my best friends, and whose advice, support and companionship have been invaluable for decades. My sister-in-law Sarah, who inspires me to improve the world through engagement. And the rest of my family in Costa Rica, Fio, Pao, Edwin and Mima, who help me remember where home is, and who I am eternally sorry to see far less often than I would like to.

Academically, I'm immensely thankful for the guidance, encouragement and support of my advisor, David Chandler. Over the past three years, I have learned from him a great deal about statistical mechanics and liquid-state theory, but more importantly, I have benefited greatly from his mentoring. He has taught me how to do research, how to write about it, how to argue a case persuasively, and how to deal with criticism, founded and unfounded. The unique work environment that he has set up for the group and the time and dedication he has afforded to all of us have been invaluable. It is a privilege to be a member of the extraordinary group of people that he has trained throughout his lifetime.

Equally, I wouldn't be where I am now if I hadn't benefited from exceptional guidance before. In particular, Pardis Sabeti has profoundly influenced me by getting me excited about science, and by being willing to take a risk on me when I was still a freshman nearly a decade ago. Her joy, energy and dedication to science are admirable and contagious, and I strive to live up to her example. Similarly, my undergraduate advisor, John Negele, has been a role model to me with regards to the value of precision, thoroughness, perseverance and hard work.

At Berkeley, I've had the fortune of benefiting from interacting with Phill Geissler, whose academic and career advice have been tremendously valuable. I was also lucky to meet Jan Liphardt and Joel Moore, who have allowed me the freedom to explore what a physicist could do outside of the physics department. I learned from Teresa Head-Gordon and Berend Smit the fundamentals of molecular dynamics and Monte Carlo simulations. I'm also grateful to Jhih-Wei Chu for the opportunity to give an invited talk at the Boston ACS meeting in August 2010, where most the lattice LCW model of Chapter 4 was first presented to a large audience. Finally, I am thankful to Richard Saykally and Ron Cohen for our interactions with respect to water evaporation, and for taking the time to listen to our suggestions and try them out. Outside of Berkeley, I have benefited from interacting with Shekhar Garde and Mike Hagan.

As for my fellow students and postdocs, I owe the greatest gratitude to Amish Patel, together with whom most of the work in this thesis was performed. Amish has served as an untiring source of ideas, advice and encouragement throughout the past three years, and is primarily responsible for introducing me to the ideas of LCW theory and how they explain the hydrophobic effect. I was lucky for my time in the group to have coincided with his own. When I entered the group, Amish and Adam Willard were my role models of success in the Chandler group, and I have tried my best to follow in their footsteps. I am also grateful to my officemate of three years, David Limmer, with whom I've shared a great many laughs and discussed much science. The others now in the group, Yael Elmatad, Aaron Keys and Ulf Pedersen, as well as others who have since left, like Thomas Speck, Lester Hedges and Kafui Tay, made my time here highly enjoyable, personally and academically. I'm equally thankful for having shared the Pitzer center with a diverse crew of creative people from whom I've learned a lot, most notably Michael Grünwald, Carl Rogers, Pat Shaffer, Todd Gingrinch, Anna Schneider, Ayelet Benjamini, Chris Ryan, Asaph Widmer-Cooper, Joyce Noah-Vanhoucke, Lutz Maibaum and Eric Sundstrom. Equally, the extended visits of Benjamin Rotenberg and Tuomas Knowles were very enriching. Finally, I thank Walter Drisdell in Richard Saykally's group for introducing me to the water evaporation experiments of Chapter 6.

Many other people outside the Pitzer Center have, at various times, had a positive impact on my time at Berkeley, including Steve Byrnes, Jonas Kjäll, Thierry Botter, Jon Blazek, Gilly Elor, Calvin Berggren, Michelle Yong, Chris Smallwood, Punit Gandhi and John Paulin. I would also like to extend my gratitude to the staff in the physics and chemistry departments that have made everything easier and smoother than it might otherwise have been. In this sense, I am particularly thankful for the constant workings of Anne Takizawa, but also those of Donna Sakima and Claudia Lopez, as well as the administrative assistants in the Pitzer center, Cheryl Higley, Lisa Littlejohn, Cheryn Gliebe and Sonia Waters, and our computer support guru, Jon Forrest.

The best, as always, is saved for last. I cannot imagine what life would have been like without meeting Laura. Even from afar, she has kept me company and given me unyielding love and support, day and night, month after month, year after year. The past five years have been the most memorable in my life owing entirely to her. She is what makes waking up every day a joy, and she is who I think when I fall asleep. I can't thank you enough for the trust you've put in me and the patience you've had in waiting for me. The long wait is finally over!

Chapter 1

Introduction

Water is a fascinating liquid. Unlike most simple liquids, its structure is not determined by packing constraints, but is instead dictated by a strong, dynamic network of hydrogen bonds, the result of strong electrostatic interactions between small, polar molecules [13]. Most biological processes, as well as many important chemical and industrial ones, take place in aqueous solution. In these processes, water is not just a passive solvent, but an active participant, whose unusual interaction with other solutes is key to the functions that these processes accomplish.

One of water's many unusual properties, which we explore in detail in this thesis, is the hydrophobic effect, or the tendency for oil and water to demix.[1] Qualitatively, the effect arises because favorable hydrogen bonding interactions between waters induce them to interact with themselves rather than with oil. Quantitatively, however, the effect is much more subtle. We discuss three such subtleties below.

It has long been observed that water hydrates small solutes differently than large solutes. On the one hand, gases of small molecules are less soluble in warm water than in cold water [105], a fact that renders warm beer unpleasant and makes warm oceans less capable of dissolving carbon dioxide. On the other hand, larger solutes, like micron-sized oil droplets, are more soluble in warm water, not less, which is why it's more effective to wash dishes and clothes with warm water. For simple solutes, this length-scale dependence can be explained in terms of hydrogen bonding. Small solutes can be accommodated in water without breaking hydrogen bonds, so the primary cost to solvating them is a reduction in entropy of the surrounding water. Large solutes, however, cannot be dissolved without breaking large numbers of hydrogen bonds, so a liquid-vapor-like interface forms around them. The surface tension of this interface decreases with temperature, rendering the solutes more soluble. While this explanation is qualitatively satisfying, the distinction between "small" and "large" is not transparent for many interesting solutes. Most prominently, the building blocks of proteins (amino acids) are "small", the folded protein is "large", and the topographical

---

[1] Most solvents exhibit a "solvophobic" effect on solutes with which they interact only weakly [49], but water is unique in the magnitude and importance of the effect.

features on the surface of the protein, where drugs, signaling molecules and other proteins bind, are intermediate in size.

Another subtle manifestation of hydrophobicity is the effective interactions between solutes in water, such as when oil droplets on the surface of water in a pot of pasta come together and coalesce into larger and larger droplets. In biology, many proteins are driven to fold to bring many of their hydrophobic amino acids into a compact core, leaving mostly hydrophilic amino acids on their outer surface [1]. When heated, most proteins denature (unfold) because the larger entropy of extended random conformations can overcome the forces, hydrophobic ones among them, that stabilize the folded state. More puzzlingly, though, some proteins also denature when they are cooled or when subjected to high pressures [91]. This phenomenon is believed to be caused by changes to the hydrophobic effect that affect the folded and unfolded states differently [12]. For example, cold denaturation can be rationalized in terms of the small amino acids becoming more soluble while the large folded protein becomes less so, which favors exposing the hydrophobic groups in an unfolded state.

When hydrophobic solutes assemble, the water around them plays a surprisingly active role. Naïvely, since hydrophobicity is a collective effect that results from the statistical behavior of large numbers of water molecules, one might expect that the net effect of the solvent could be reduced to an effective potential on the solute. This assumes that solutes in water move very slowly compared to the water molecules, and is analogous to using the Born-Oppenheimer approximation to construct effective potentials for the nuclei of atoms in molecules. While this approach is often sensible, it fails spectacularly in important cases. Studies of model hydrophobic polymers undergoing hydrophobic collapse [78, 106] and of model hydrophobic plates coming together [6, 52] both show that the two end states are separated by large free energy barriers that are due entirely to the solvent degrees of freedom. That is, the main bottleneck to collapse or assembly in these systems is a large-scale collective rearrangement of the solvent. Such a bottleneck is invisible in a solute-centric description of the process. Similar phenomena have been observed in simulations of protein systems [64], and it is tempting to speculate that several puzzling free energy barriers observed in FRET experiments between stable states of certain chaperones [77] may be due to solvent motions and not protein rearrangements.

About a decade ago, Lum, Chandler and Weeks (LCW) presented a solid theoretical framework for understanding various aspects of hydrophobicity [68]. LCW theory attempts to reduce the behavior of water (and other solvents) to an understanding of solvent density fluctuations, from which hydrophobic behaviors can then be inferred. Solvent density fields are described in terms of a large-length-scale, smooth component and an overlying component of small-length-scale fluctuations. The physics describing fluctuations in either component individually are transparent (interfacial energetics vs. Gaussian fluctuations), and the two components are then coupled weakly. In this framework, most of the subtleties of the hydrophobic effect arise from the interplay between large-scale interfacial physics and molecular-scale fluctuations. The theory has been implemented at different levels of detail and been used to quantify many of the above phenomena, with significant success [46–49, 68, 71, 106, 107, 123, 124]. Unfortunately, because of its technical complexity, LCW theory is not yet used routinely in fields like biomolecular simulations, where many of its most interesting consequences are yet to be discovered. This thesis makes an important advance in this direction.

Before proceeding, we present an outline for the remainder of the thesis. In Chapter 2, we lay down the statistical mechanical framework for quantifying the hydrophobic effect. We summarize the theoretical aspects of small- and large-length-scale solvation, and conclude with a presentation of LCW theory in its continuum formulation. In Chapter 3, we focus on quantifying various kinds of fluctuations in atomistic models of water. We introduce computational probes for performing the relevant probability distribution measurements and systematically quantify large-length-scale fluctuations in bulk, near hydrophobic and hydrophilic surfaces and around hydrophobic protein surfaces. We further discuss tools to probe liquid-vapor interfaces and their relations to dissolved solutes. In Chapter 4, we present a coarse-grained theory of solvation on a lattice, based on LCW ideas, improving and extending previous attempts to do this. We discuss some of the pitfalls involved in discretizing the LCW equations, their consequences, and satisfactory alternative discretizations that avoid them. We also discuss the validity of many of the approximations made. A detailed comparison of the final theory is made with the detailed numerical results of Chapter 3 to demonstrate the theory's accuracy. We conclude with a discussion the future developments necessary for widespread use of LCW theory in biological and other settings. In Chapter 5, we consider how hydrophobic and hydrophilic surfaces influence the hydrophobic effect in nearby dissolved solutes. We first present detailed numerical results, and then build a minimal LCW-like theory to rationalize all the observed behavior. Our main conclusion is that hydrophobic surfaces may act as generic catalysts for unfolding proteins, an observation that may contribute to our understanding of chaperone function. Finally, in Chapter 6, we employ our tools for looking at liquid-vapor interfaces to examine the process of water evaporation. We find an interesting connection between local surface curvature and evaporation. However, we also find that in our simulations, the process appears to be barrierless, in slight disagreement with recent experiments, and we suggest a possible reason for the disagreement.

Chapter 2

Theories of Solvation

2.1 Preliminaries

At heart, hydrophobicity is a collective effect, so we need statistical mechanics to understand it. On the other hand, quantum effects play an insignificant role in most liquids at ambient conditions [41]. Hence, we work within the Born-Oppenheimer approximation and treat a system of N atoms as if the nuclei were classical point particles at positions $\mathbf{R}^N = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N\}$. These nuclei interact via a classical many-body potential $U(\mathbf{R}^N)$ that results from integrating out the electronic degrees of freedom. We assume that $U(\mathbf{R}^N)$ is well modeled by a force field of (a) pairwise interactions that capture intermolecular effects like Pauli exclusion, dispersive attractions and electrostatic attractions, and (b) harmonic or low-order Fourier expansion terms that capture chemical bonding detail, and involve between two and four atoms each. As in most other work in the field, we neglect the effects of polarization on the electron densities of molecules.[1]

For concreteness, we use the following form of $U(\mathbf{R}^N)$, common to many force fields like AMBER [16] and CHARMM [70]:

$$\begin{aligned}
U(\mathbf{R}^N) = \sum_{i,j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^{6} \right] + \sum_{i,j} \frac{q_i q_j}{4\pi\epsilon_0 R_{ij}} \\
+ \sum_{i\text{-}j\text{ bonds}} \frac{1}{2} K^{(\mathrm{b})}_{ij} (R_{ij} - R^{(0)}_{ij})^2 + \sum_{i\text{-}j\text{-}k\text{ angles}} \frac{1}{2} K^{(\mathrm{a})}_{ijk} (\theta_{ijk} - \theta^{(0)}_{ijk})^2 \\
+ \sum_{i\text{-}j\text{-}k\text{-}l\text{ torsions}} \sum_{n} \frac{1}{2} K^{(\mathrm{t})}_{ijkl,n} [1 + \cos(n\chi_{ijkl} - \delta_{ijkl,n})] \qquad (2.1)
\end{aligned}$$

---

[1] For water in particular, this is approximation is reasonable. Its static dielectric constant, 80, results primarily from molecular rearrangements, not electronic polarization. The dielectric constant at optical frequencies, arising mostly from electronic polarization, is close to 2.

The indices i through l denote atomic nuclei, and unrestricted sums are over all the nuclei in the system. Here and everywhere, the vector $\mathbf{R}_{ij}$ is defined as $\mathbf{R}_i - \mathbf{R}_j$, and its length is R_{ij}, i.e. the distance between nuclei i and j. The angle θ_{ijk} is that formed between $\mathbf{R}_{ij}$ and $\mathbf{R}_{kj}$, and the torsion angle χ_{ijkl} is the angle between the vector $\mathbf{R}_{lk}$ and the plane containing atoms i, j and k. The first term models Pauli exclusion and dispersion forces between all atoms i and j via a Lennard-Jones potential [34]. The second term captures electrostatic effects, and the charge density of molecules is modeled by point partial charges q_i centered at the atomic nuclei. All chemical bonds are taken to be fixed, and the harmonic terms are the first terms in a Taylor expansion of $U(\mathbf{R}^N)$ about the equilibrium lengths of bonds and the equilibrium angles between bonded trios of atoms. Since rotation about single bonds is a relatively soft degree of freedom, the potential energy as a function of torsion angle χ is instead modeled by a few terms in a Fourier series. All remaining symbols are parameters of the force field. The chemical identity of the system being modeled is encoded in these parameters. These are determined through various heuristics [16, 70], but are generally either fit systematically to the results of detailed quantum-mechanical calculations for a representative set of small molecules, or chosen to match a canonical set of experimental measurements like normal mode frequencies, liquid densities and vaporization enthalphies.

When modeling bulk systems in a computer, we are restricted to small numbers of atoms. To mitigate the effect of small system size, we impose periodic boundary conditions on our systems [34]. That is, we imagine that the atomic nuclei are enclosed in a cubic box of dimensions $L_x \times L_y \times L_z$ that periodically tiles space. In such an infinite system, the potential energy is infinite, but the potential energy density is finite. Hence, when calculating $U(\mathbf{R}^N)$, we restrict sums over i to be over the atoms in only one of the cubic boxes, and in terms describing atomic bonding, $\mathbf{R}_{ij}$ always refers to the vector joining atom i in one cubic box to the nearest image of atom j in any periodically-replicated box. For conciseness, we call the resulting value of $U(\mathbf{R}^N)$ the system's potential energy, although strictly speaking, it is the potential energy density multiplied by the periodic box size. Because of periodic boundary conditions, the function $U(\mathbf{R}^N)$ depends implicitly on the system volume V, i.e., $U(\mathbf{R}^N) = U(\mathbf{R}^N; V)$.

Thermodynamic quantities are derived from the potential energy function $U(\mathbf{R}^N; V)$ in the usual way. When studying solvation behavior, the most natural external conditions to impose are fixed temperature T and fixed pressure P. This corresponds to working in the isothermal-isobaric ensemble [41], where the system's partition function, $\Delta(N, P, T)$, is given by

$$\Delta(N,P,T) = C^{-1} \int \beta P \mathrm{d}V \int \prod_{i=1}^{N} \frac{\mathrm{d}\mathbf{R}_i}{\Lambda_i^3} \exp\left[-\beta P V - \beta U(\mathbf{R}^N; V)\right]. \tag{2.2}$$

Here, C^{-1} is a constant that ensures proper Boltzmann counting, $\Lambda_i = h/\sqrt{2\pi m_i k_\mathrm{B} T}$ is the thermal wavelength of particle i, h is Planck's constant, β is $1/k_\mathrm{B}T$, and k_B is Boltzmann's constant. When integrating over volume V, the unit box's shape is kept fixed. The corresponding Gibbs free energy is given by

$$G(N,P,T) = -k_\mathrm{B}T \ln \Delta(N,P,T). \tag{2.3}$$

The thermal average of an observable $A(\mathbf{R}^N;V)$, denoted by $\langle A \rangle$, is given by

$$\langle A \rangle = \frac{\int \mathrm{d}V \int \mathrm{d}\mathbf{R}^N \, e^{-\beta PV - \beta U(\mathbf{R}^N;V)} \, A(\mathbf{R}^N;V)}{\int \mathrm{d}V \int \mathrm{d}\mathbf{R}^N \, e^{-\beta PV - \beta U(\mathbf{R}^N;V)}}. \tag{2.4}$$

When the choice of potential $U(\mathbf{R}^N;V)$ is ambiguous, it is specified as a subscript, i.e., $\langle A \rangle_U$. The free energy difference between two systems with potential energy functions $U_1(\mathbf{R}^N;V)$ and $U_2(\mathbf{R}^N;V)$ is given by

$$G_2(N,P,T) - G_1(N,P,T) = -k_\mathrm{B}T \ln \frac{\int \mathrm{d}V \int \mathrm{d}\mathbf{R}^N \, e^{-\beta PV - \beta U_2(\mathbf{R}^N;V)}}{\int \mathrm{d}V \int \mathrm{d}\mathbf{R}^N \, e^{-\beta PV - \beta U_1(\mathbf{R}^N;V)}} \tag{2.5a}$$

$$= -k_\mathrm{B}T \ln \left\langle e^{-\beta(U_2 - U_1)} \right\rangle_{U_1}. \tag{2.5b}$$

The last equation is a conceptually simple way of determining a free energy difference by evaluating thermal averages only, which, unlike absolute free energies, can be estimated computationally. In practice, free energy differences are computed using more efficient methods, such as those described in Section 3.1.1.

In all subsequent sections, the distinction between the isothermal-isobaric ensemble and the more familiar canonical ensemble is not significant. Hence, for compactness, we usually write free energies and thermal averages in an ensemble-agnostic notation (as in Equation (2.5b)), and highlight the pressure dependence only where relevant.

In most cases, it is impractical and unnecessary to keep track of all the atomic nuclei in a system. It suffices to keep track of a few select degrees of freedom, which evolve approximately according to a mean field created by the remaining degrees of freedom. For example, there may be a natural separation of time scales between fast and slow degrees of freedom. Let $q^K = \{q_1(\mathbf{R}^N), \ldots, q_K(\mathbf{R}^N)\}$ denote the K slow degrees of freedom. The effective free energy surface on which the dynamics of these degrees of freedom takes place is given by

$$G(q^K) = -k_\mathrm{B}T \ln \left(\left\langle \prod_{i=1}^{K} \delta\left(q_i(\mathbf{R}^N) - q_i\right) \right\rangle \Big/ \left\langle \prod_{i=1}^{K} \delta\left(q_i(\mathbf{R}_0^N) - q_i\right) \right\rangle \right). \tag{2.6}$$

Here, $\mathbf{R}_0^N$ is an arbitrary reference configuration chosen to fix the zero of the effective free energy. In the simplest description, when there indeed is a natural separation of time scales, the system's dynamics might be described by Langevin dynamics, i.e., diffusion along the effective free energy surface, with the net effect of the fast degrees of freedom implicitly accounted for by an effective noise and friction [10]. The simplest realization of such dynamics is

$$m_i \ddot{q}_i = -\partial G / \partial q_i - \gamma \dot{q}_i + \eta_i(t), \qquad 1 \le i \le K, \tag{2.7}$$

where m_i is an effective mass for degree of freedom i, γ is an effective friction constant, and $\eta_i(t)$ is a random variable that, to be consistent with the fluctuation-dissipation theorem, satisfies

$$\langle \eta_i(t)\eta_j(t')\rangle = 2\gamma k_{\mathrm{B}} T \delta_{ij}\delta(t-t'), \tag{2.8}$$

where the average is over all possible realizations of the noise. We discuss in Section 2.2.2 which degrees of freedom of water may be treated as "slow".

2.2 Thermodynamics of solvation

The solvation behavior of a solute is characterized by its solvation free energy, ΔG_{sol}, defined as the free energy of transferring the solute from vacuum into a solution. More precisely, let $\mathbf{r}^n$ and $\mathbf{s}^m$ denote the positions of the n solute and m solvent nuclei. For the form of potential energy in Equation (2.1), we can identify terms involving only solute nuclei, those involving only solvent nuclei, and those coupling the two. These terms are grouped into potential energies U_{rr}, U_{ss} and U_{rs}, respectively, so that

$$U(\mathbf{r}^n,\mathbf{s}^m) = U_{rr}(\mathbf{r}^n) + U_{rs}(\mathbf{r}^n,\mathbf{s}^m) + U_{ss}(\mathbf{s}^m). \tag{2.9}$$

With this separation, we define ΔG_{sol} precisely as[2]

$$\Delta G_{\mathrm{sol}} = -k_{\mathrm{B}} T \ln \left\langle e^{-\beta U_{rs}} \right\rangle_{U_{rr}+U_{ss}} \tag{2.10}$$

Conceptually, we can break down the transfer into three physically distinct steps [92], shown schematically in Figure 2.1. First, a cavity is carved into the solvent in the shape of the solute, at free energy cost ΔG_{cav}. Second, the solute and the solvent are allowed to interact via short-range dispersive forces, at free energy cost ΔG_{vdW}. Finally, the solute's charge density is allowed to interact with and polarize the surrounding solvent, at free energy cost ΔG_{pol}. Operationally, these steps correspond to splitting $U_{rs}(\mathbf{r}^n,\mathbf{s}^m)$ into three terms,

$$U_{rs}(\mathbf{r}^n,\mathbf{s}^m) = U_{\mathrm{cav}}(\mathbf{r}^n,\mathbf{s}^m) + U_{\mathrm{vdW}}(\mathbf{r}^n,\mathbf{s}^m) + U_{\mathrm{pol}}(\mathbf{r}^n,\mathbf{s}^m). \tag{2.11}$$

and then, exactly as in Equation (2.10), calculating the free energy difference that arises from incorporating each successive term, yielding

$$\Delta G_{\mathrm{cav}} = -k_{\mathrm{B}} T \ln \left\langle e^{-\beta U_{\mathrm{cav}}} \right\rangle_{U_{rr}+U_{ss}}, \tag{2.12}$$

$$\Delta G_{\mathrm{vdW}} = -k_{\mathrm{B}} T \ln \left\langle e^{-\beta U_{\mathrm{vdW}}} \right\rangle_{U_{rr}+U_{ss}+U_{\mathrm{cav}}}, \tag{2.13}$$

$$\Delta G_{\mathrm{pol}} = -k_{\mathrm{B}} T \ln \left\langle e^{-\beta U_{\mathrm{pol}}} \right\rangle_{U_{rr}+U_{ss}+U_{\mathrm{cav}}+U_{\mathrm{vdW}}}. \tag{2.14}$$

---

[2] Strictly speaking, the average volume of the system depends upon whether the term $U_{rs}(\mathbf{r}^n,\mathbf{s}^m)$ is or is not included, so a small concentration correction must be added to ensure that both initial and final states represent equal concentrations of the solute. The correction is $k_{\mathrm{B}} T \ln \langle V\rangle_{U_{rr}+U_{ss}} / \langle V\rangle_{U_{rr}+U_{rs}+U_{ss}}$, and vanishes as $\langle V\rangle \to \infty$.

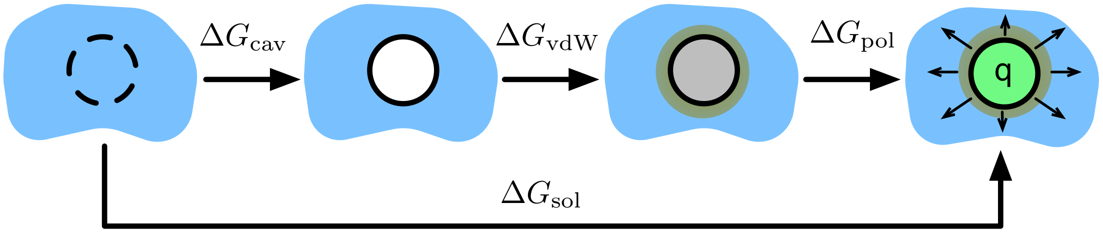


Figure 2.1: Schematic of solvation free energy $\Delta G_{\rm sol}$, broken into three physically distinct free energies: a free energy of cavitation, $\Delta G_{\rm cav}$, a free energy of dispersive solute-solvent interactions, $\Delta G_{\rm vdW}$, and a free energy of polarization, $\Delta G_{\rm pol}$.

These free energy differences are related by the equation

$$\Delta G_{\rm sol} = \Delta G_{\rm cav} + \Delta G_{\rm vdW} + \Delta G_{\rm pol}. \tag{2.15}$$

In general, $\Delta G_{\rm cav}$ is positive, whereas $\Delta G_{\rm vdW}$ and $\Delta G_{\rm pol}$ are negative. The sign of $\Delta G_{\rm sol}$ is the simplest characterization of hydrophobicity. Solutes for which $\Delta G_{\rm sol} > 0$ are hydrophobic, since it takes positive reversible work to insert one solute molecule into the solvent, whereas solutes for which $\Delta G_{\rm sol} < 0$ are hydrophilic[3]. Indeed, the experimentally measured values of $\Delta G_{\rm sol}$ of amino-acid side chains form the basis of the widely used Kyte-Doolittle hydropathy scale for measuring the hydrophobic character of the twenty naturally occurring amino acids [62].

The definitions of the various ΔG above include the difference in free energy owing to internal motions of the solute. For example, a butane molecule dissolved in water is found in a compact cis configuration slightly more often than in the gas phase [53]. We can remove this complication by defining a configuration-dependent solvation free energy $\Delta G_{\rm sol}(\mathbf{q}^n)$ to be the free energy of transferring the solute *frozen into configuration* $\mathbf{q}^n$ from vacuum into solution. Explicitly,

$$\Delta G_{\rm sol}(\mathbf{q}^n) = -k_{\rm B}T \ln \left\langle e^{-\beta U_{rs}} \prod_{i=1}^{n} \delta(\mathbf{r}_i - \mathbf{q}_i) \right\rangle_{U_{rr}+U_{ss}} \tag{2.16}$$

In general, for a large macromolecule, $\Delta G_{\rm sol}(\mathbf{q}^n)$ will have a non-trivial dependence on $\mathbf{q}^n$. Hence, $\Delta G_{\rm sol}(\mathbf{q}^n)$ acts as an effective potential on the macromolecule that describes the effect of the solvent on the molecule's configurations. For purely hydrophobic polymers, for

---

[3] These criteria apply for a solute in bulk vapor or in bulk liquid. The issue is more subtle in the presence of interfaces, as the recent work of Pat Shaffer and Phill Geissler has demonstrated (personal communication, in preparation). For example, a surface-adsorbed ion has $\Delta G_{\rm sol} < 0$, so overall, the ion is "hydrophilic" (it prefers the water to the vapor). However, from the point of view of bulk water, it appears hydrophobic, as it is expelled towards the interface.

example, $\Delta G(\mathbf{q}^n)$ is much lower when the polymer is collapsed into a globule than when it is extended, which reflects the hydrophobic polymer's tendency to adopt a compact shape when in solution. Similarly, in the context of proteins, $\Delta G_{\mathrm{sol}}(\mathbf{q}^n)$ is much lower than otherwise for configurations $\mathbf{q}^n$ in which the protein's hydrophobic residues are buried inside a globular core and only the hydrophilic residues are exposed to the surrounding water. The configurational dependence of $\Delta G_{\mathrm{sol}}(\mathbf{q}^n)$ thus captures the collective effects that lead to the hydrophobic effect. Thermodynamically, understanding the hydrophobic effect means having a tractable theory for estimating the configurational dependence of $\Delta G_{\mathrm{sol}}(\mathbf{q}^n)$.

2.2.1 Length-scale dependence of solvation free energies

We can estimate ΔG_{sol} analytically in several limits, where the essential physics of the hydrophobic effect is transparent. Before proceeding, we introduce the simplifying concept of an ideal solute, one that interacts with the solvent only by excluding solvent atoms from its interior, denoted by $V(\mathbf{r}^n)$, so that $\Delta G_{\mathrm{vdW}} = \Delta G_{\mathrm{pol}} = 0$. More precisely,

$$U_{\mathrm{cav}}(\mathbf{r}^n, \mathbf{s}^m) = \begin{cases} \infty, & \text{any } \mathbf{s}_i \text{ in } V(\mathbf{r}^n), \\ 0, & \text{otherwise,} \end{cases} \tag{2.17a}$$

$$U_{\mathrm{vdW}}(\mathbf{r}^n, \mathbf{s}^m) = 0, \tag{2.17b}$$

$$U_{\mathrm{pol}}(\mathbf{r}^n, \mathbf{s}^m) = 0. \tag{2.17c}$$

Such solutes are strongly hydrophobic, and we can understand most aspects of the hydrophobic effect by focusing on ΔG_{cav} alone.

For a macroscopic solute of volume v and surface area A, ΔG_{cav} is dominated by pressure, p, and surface tension, γ, so that

$$\Delta G_{\mathrm{cav}} \approx pv + \gamma A. \tag{2.18}$$

The surface tension is that of the liquid solvent with vacuum. For a liquid near its triple point, such as water at ambient conditions ($T = 298\,\mathrm{K}$ and $p = 1\,\mathrm{atm}$), it is essentially equal to the liquid-vapor surface tension at saturation, γ_{lv}. Its value at 298 K in various common units is [63]

$$\gamma_{lv} = 72.00\,\mathrm{mJ/m^2}, \tag{2.19a}$$

$$= 17.50\,k_{\mathrm{B}}T/\mathrm{nm}^2, \tag{2.19b}$$

$$= 43.36\,\mathrm{kJ/mol/nm^2}, \tag{2.19c}$$

$$= 103.6\,\mathrm{cal/mol/Å^2}. \tag{2.19d}$$

Microscopically, this high surface tension is due to large numbers of hydrogen-bonds being broken at a liquid-vapor interface. That is, when a large solute is embedded in water, it is impossible for the hydrogen-bonding network of water to deform and accommodate the presence of the solute without many such bonds being broken.

The surface tension of water decreases nearly linearly with increasing temperature essentially up to the very near the critical point ($T_c = 647\,\mathrm{K}$). At the critical point, it is exactly zero. Its temperature derivative at ambient conditions is

$$\frac{\partial \gamma_{lv}}{\partial T} = -0.15\,\mathrm{mJ/m^2}. \tag{2.20}$$

Hence, ΔG_{sol} decreases with increasing temperature, so that macroscopic solutes are more soluble in hot water than in cold water.

Owing to liquid water's high surface tension, and the fact that 1 atm of pressure is essentially equal to zero pressure at molecular scales, the surface tension term generally overwhelms the pressure term for all but the largest common solutes. For example, for a sphere of radius R at 1 atm pressure, the term γA exceeds pv when $R \lesssim 2\,\mu\mathrm{m}$. Moreover, the excluded volume v is nearly independent of solute configuration, so that most of the configurational dependence of $\Delta G_{\mathrm{sol}}(\mathbf{r}^n)$ results from changes in surface area. For these two reasons, we can safely ignore the pv term in subsequent discussions.

For nanoscopic solutes, ΔG_{sol} behaves differently. When v is tiny, the number of water molecules inside it, N, will almost always be 0, only every so often be 1, and almost never be any higher.[4] Let p_0 and $p_1 \approx 1 - p_0$ be the probabilities of observing 0 or 1 waters in v. The average number of waters observed in this cavity is $\rho_\ell v$, where $\rho_\ell \approx 33.3\,\mathrm{waters/nm^3}$ is the number density of water molecules at ambient conditions. Hence, $p_1 \approx \rho_\ell v$, so that $p_0 \approx 1 - \rho_\ell v$. It follows from Equation (2.12) that

$$\Delta G_{\mathrm{cav}} \approx -k_{\mathrm{B}} T \ln(1 - \rho_\ell v) \approx k_{\mathrm{B}} T \rho_\ell v, \quad (\text{tiny } v). \tag{2.21}$$

Hence, for a tiny solute, the solvation free energy scales with excluded volume and temperature, so small solutes are less soluble in hot water than in cold water. Physically, small solutes, unlike large ones, *can* be accommodated within the existing hydrogen-bonding network of water. The volume exclusion merely reduces the number of configurations available to the surrounding water, so the solvation free energy is dominated by entropic considerations.

The solutes where the above argument applies are impossibly tiny. For the argument to work, we need $\rho_\ell v \ll 1$, so that $v \ll 30\,\text{Å}^3$, the volume of a sphere of radius 1.9 Å. This is roughly the solvent-excluding volume of a hydrogen atom. If we admit using additional empirical information, we can treat somewhat larger solutes analytically, as follows.

Computer simulations of various liquids near their triple point have shown that small density fluctuations in them are Gaussian to an excellent approximation (Chapter 3 describes

---

[4] It is generally clear from context whether N is the total number of waters in the system or just the number of waters inside the probe volume v.

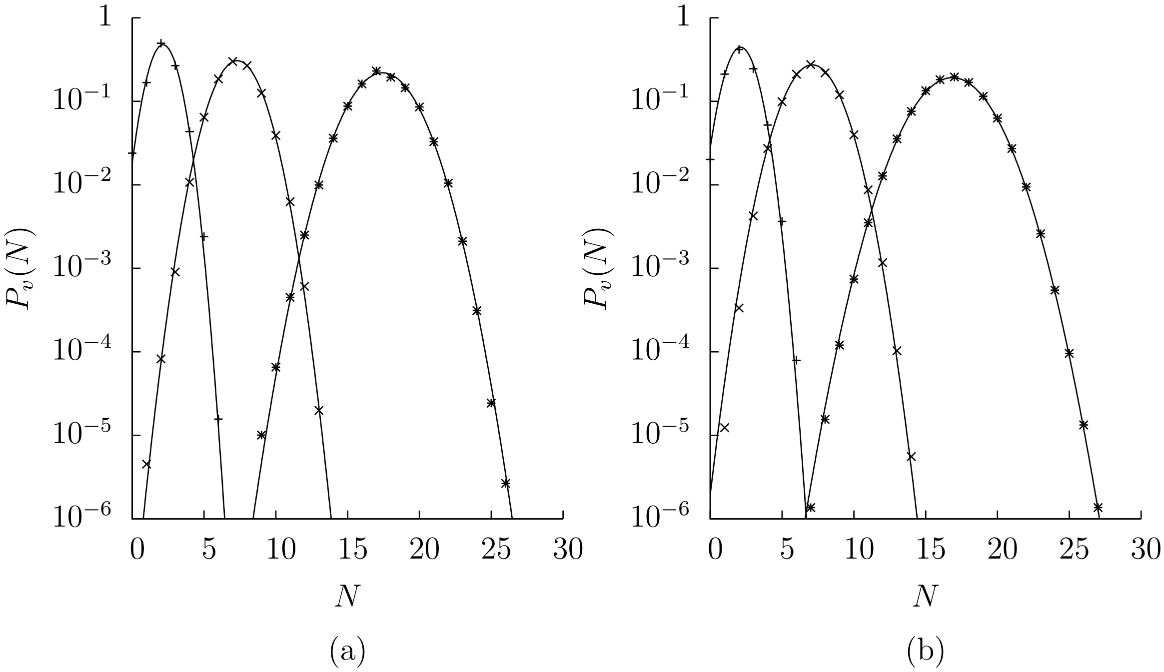


Figure 2.2: (a) Probability $P_v(N)$ of finding N water molecules inside spheres of radius 2.5 Å, 3.75 Å and 5.0 Å. After Ref. [53]. (b) Same, for a fluid of hard spheres of radius σ at number density $\rho = 0.5\sigma^{-3}$, with v a sphere of radius 1.0σ, 1.5σ and 2.0σ, respectively. Adapted from Ref. [18].

methods to probe large fluctuations, where deviations from Gaussian behavior become evident). Indeed, Chandler showed that many approximate theories of liquid structure are equivalent to assuming Gaussian statistics for small density fluctuations [11]. One way to test this approximation, introduced by Hummer et al [53], is to measure the probability of finding N water molecules inside a probe volume v, denoted by $P_v(N)$. Similar measurements have been made for a fluid of hard spheres [18] and for a fluid of Lennard-Jones particles [46]. Figure 2.2 shows a representative set of probability distributions for spherical volumes of various sizes, and corresponding Gaussian fits that confirm the accuracy of the Gaussian approximation.

The moments of $P_v(N)$ can be estimated analytically using only experimentally accessible information. The mean, $\langle N\rangle_v$, is given by

$$\langle N\rangle_v \approx \rho_\ell v, \tag{2.22}$$

and the variance, $\langle(\delta N)^2\rangle_v$ is equal to σ_v, given by

$$\sigma_v = \left\langle(\delta N)^2\right\rangle_v \approx \int_{\mathbf{r}\in v}\int_{\mathbf{r}'\in v}\langle\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\rangle\,. \tag{2.23}$$

It can be shown [41] that the integrand, denoted by $\chi(\mathbf{r},\mathbf{r}')$, is trivially related to the pair distribution function,[5] $g(r)$, by the equation

$$\chi(\mathbf{r},\mathbf{r}') = \langle\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\rangle = \rho_\ell\delta(\mathbf{r}-\mathbf{r}') + \rho_\ell^2[g(|\mathbf{r}-\mathbf{r}'|)-1]. \tag{2.24}$$

The Gaussian model predicts that $P_v(N)$ is well approximated by

$$P_v(N) \approx \frac{1}{\sqrt{2\pi\sigma_v}}\exp\left[-(N-\langle N\rangle_v)^2/2\sigma_v\right]. \tag{2.25}$$

For small enough volumes, the Gaussian approximation to $P_v(N)$ holds even at $N=0$. From Equation (2.12), we deduce that

$$\Delta G_{\text{cav}} = -k_{\text{B}}T\ln P_v(0), \tag{2.26}$$

so

$$\Delta G_{\text{cav}} \approx k_{\text{B}}T\,\langle N\rangle_v^2/2\sigma_v + k_{\text{B}}T\ln\sqrt{2\pi\sigma_v}. \tag{2.27}$$

This analytical expression for ΔG_{cav} is surprisingly powerful and accurate. For example, when the solvent is water, it correctly predicts the solvation free energies of spheres of up to about 4 Å in radius, the potential of mean force between two methyl-sized cavities (including an effective "hydrophobic bond" between the cavities stabilized by a single water molecule between them) and the bias for the cis configuration of butane with respect to the trans configuration in water [53].

We can approximate the right-hand side of Equation (2.27) further by noting that $\chi(\mathbf{r},\mathbf{r}')$ is translationally invariant and local. As a result, when v is not too small, we may approximate

$$\sigma_v \approx \int_{\mathbf{r}\in v}\int_{\mathbf{r}'}\langle\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\rangle = v\int_{\mathbf{r}}\langle\delta\rho(\mathbf{0})\delta\rho(\mathbf{r})\rangle = \rho v\chi, \tag{2.28}$$

where $\chi = \partial\rho/\partial\beta p$ is the unitless isothermal compressibility of water [41], related to the unitful isothermal compressibility, $\chi_T = -(1/V)(\partial V/\partial P)_T$, by the relation $\chi_T = \rho k_{\text{B}}T\chi$. Its value for water at ambient conditions is about 0.06. Using this approximation and neglecting the typically small log-term in Equation (2.27), we obtain

$$\Delta G_{\text{cav}} \approx k_{\text{B}}T\frac{\rho}{2\chi}v. \tag{2.29}$$

---

[5] This is strictly the pair distribution function of oxygen nuclei, which is equal, to an excellent approximation, to the pair distribution function of scattering centers measured by X-ray diffraction [81].

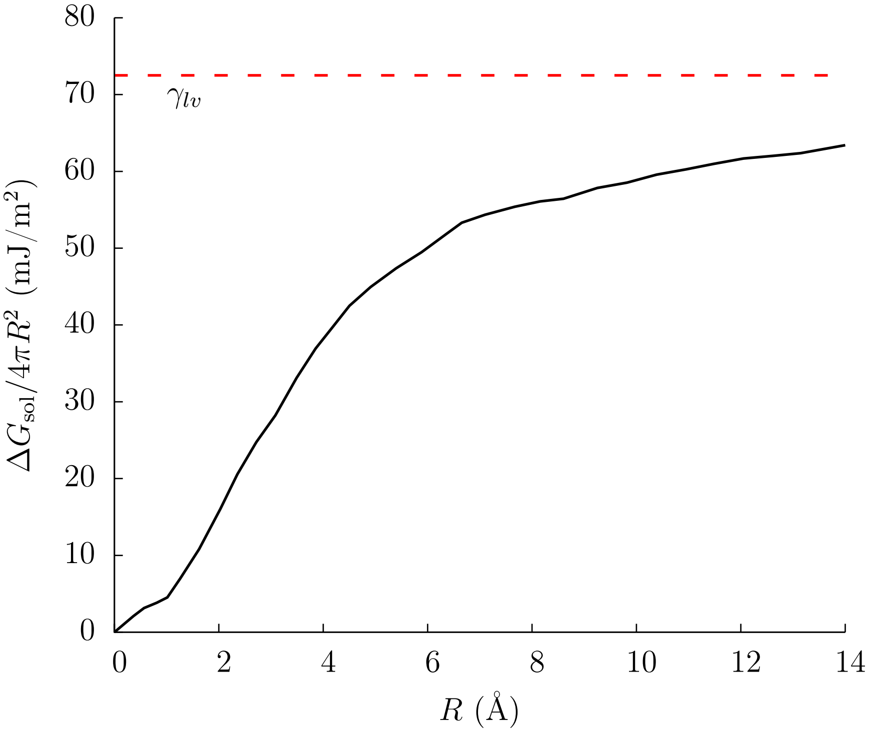


Figure 2.3: Solvation free energy per surface area for spherical cavities of radius R in SPC/E water. After Ref. [49].

Quantitatively, the approximation is not too accurate, but qualitatively, it reveals that the solvation free energies of small volumes, up to about 1 nm in size, scales with volume and temperature, and arises primarily from entropic constraints on the surrounding solvent.

Considering the microscopic and macroscopic regimes just described, it must be the case that the underlying physics of solvation depends significantly on the length scale of the solutes. The solvation free energy of small solutes is proportional to temperature and volume, while that of large solutes is proportional to surface area and surface tension (so it decreases with temperature). Solutes of intermediate size will exhibit intermediate behavior. This is indeed the case [12, 49], as is shown in Figure 2.3.

The length scale at which the solvation physics crosses over from entropic considerations to that of formation of liquid-vapor-like interfaces is about 1 nm. This is precisely the scale at which several fundamental steps of biological processes, like folding and substrate binding, take place. For instance, proteins are polymers made of amino acids, whose side chains have dimensions between a few Angstroms and about 1 nm. Similarly, many common drugs are small molecules of nanometric size that act by binding a complementary pocket on a protein surface.

The length-scale dependence of solvation free energies in Figure 2.3 is related to the effective hydrophobic force of assembly. When a set of small solutes is dispersed in solution, their total solvation free energy scales with their cumulative volume, whereas when they aggregate, the solvation free energy scales like the area of the aggregate. The temperature dependence of this force of assembly is somewhat surprising. The free energy of aggregation *increases* with increasing temperature, since the isolated monomers are less soluble, whereas the aggregate is more so. This is in contrast what would be expected if solvation free energies of solutes of all sizes were proportional to surface area and surface tension. In that case, assembly of small solutes would also be observed, since the total surface area would then be reduced, but the free energy of aggregation, being proportional to surface tension, would decrease with increasing temperature. An example where this difference is strikingly visible is in protein folding. Proteins are observed to denature when heated, since at high temperatures, disordered conformations are always favored. However, many proteins also denature when cooled. This puzzling fact can be explained by noting that the hydrophobic force of assembly decreases when the solvent is colder, and forming a hydrophobic core is a key mechanism for proteins to fold. The unusual temperature dependence of assembly can also quantitatively explain other puzzling experimental observations, such as the non-monotonic temperature dependence of the critical concentration of amphiphilic monomers above which micelles form in solution [71].

2.2.2 The role of solvent in hydrophobic collapse and assembly

When a hydrophobic polymer collapses, as well as when two hydrophobic solutes come together, the solvent plays an important dynamical role beyond simply creating an effective free energy surface for the solute's degrees of freedom. In essence, aggregation and collapse both involve a microscopic phase transition[6] in the solvent: the individual monomers are hydrated by accommodating them in cavities that form spontaneously in the solvent, but the aggregate is embedded in a vapor-like region of the solvent. The barrier to nucleating such a vapor-like region and its concomitant liquid-vapor interface can be high, and may in fact be the highest free energy barrier to collapse or assembly.

Solvent-dominated free energy barriers have been observed explicitly in the collapse of model hydrophobic polymers, as shown in the two examples of Figure 2.4. In both examples, a model hydrophobic 12-mer with high bending resistance is embedded in a solvent, which stabilizes a collapsed-globule conformation. The transition paths between the extended and collapsed states are then examined, either by explicit sampling [7] or by constructing a minimum free energy path between the states.[7] In one example, the solvent is represented by

---

[6] By "microscopic phase transition", we mean a prominent bistability in the system's free energy landscape. Since the aggregate size is fixed, we cannot consider a thermodynamic limit for this transition, in which some free energy derivative becomes discontinuous. As a result, this bistability is not a true phase transition.

[7] A minimum free energy path is a curve in a space of collective coordinates, such as the q^K coordinates of Section 2.1, that (a) connects two fixed points in this space (here, the extended and collapsed states of

a coarse-grained model based on Lum-Chandler-Weeks theory (discussed in Section 4.4.4), while in the second example, the water is modeled in atomistic detail. In both examples, a free energy surface is constructed by integrating out the small-length-scale density fluctuations of the solvent, while retaining an explicit description of the large-length-scale fluctuations. What is observed is that there are large free energy barriers along the solvent coordinates in this reduced description, which are absent when the free energy surface is further collapsed onto the solute coordinates alone. What this implies is that there is a subset of solvent degrees of freedom that evolves at least as slowly as the solute degrees of freedom. Hence, to capture the dynamics of hydrophobic collapse or assembly in these cases, qualitatively if not quantitatively, it is essential to represent these slow solvent degrees of freedom explicitly. A Langevin model for dynamics, like the one in Equation (2.7), where *all* the solvent degrees of freedom have been integrated out will instead underestimate the time scales for hydrophobic association and collapse by many orders of magnitude.

2.3 Lum-Chandler-Weeks theory

As discussed above, small-length-scale and large-length-scale solvent density fluctuations have qualitatively different thermodynamics and kinetics. One way to build a theory that describes solvation might therefore start by modeling these two kinds of density fluctuations separately, then introduce an effective coupling between the two. This is precisely what Lum-Chandler-Weeks theory (LCW) does [67, 68]. Our discussion of LCW theory here closely follows the modern presentation found in Ref. [13], instead of that found in the original LCW paper [68].

The aim of LCW theory is to describe the probability of observing any particular realization of the solvent density field, $\rho(\mathbf{r})$, in terms of an effective Hamiltonian. Orientational effects, such as hydrogen bonding, are captured implicitly in this Hamiltonian. To begin, it is useful to split the solvent density into a smooth part that varies slowly in space and time, and a field of overlying fluctuations, as follows

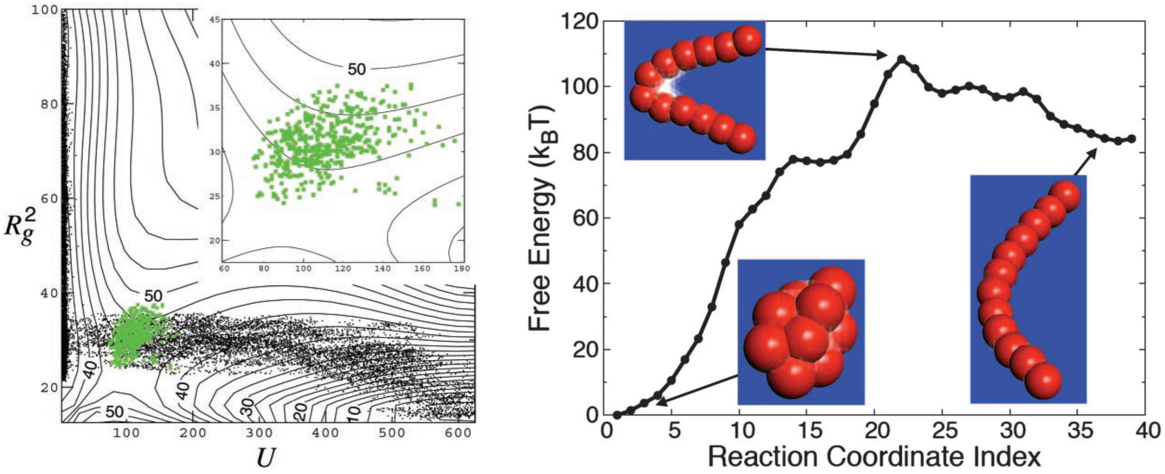
$$\rho(\mathbf{r}) = \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r}). \tag{2.30}$$

In doing so, we imagine that $n(\mathbf{r})$ captures the slow, large-length-scale features of the solvent density, and that $\delta\rho(\mathbf{r})$ time-averages to near zero on a few-picosecond time scale.

This split is not unique, so there is not a 1-1 mapping from $\rho(\mathbf{r})$ to $n(\mathbf{r})$. However, we can visualize a typical $n(\mathbf{r})$ associated with $\rho(\mathbf{r})$ by convoluting $\rho(\mathbf{r})$ with a Gaussian of width ξ comparable to a correlation length in water (about 3-4 Å at ambient conditions). Two examples of this procedure are shown in Figure 2.5. In these systems, a slab of water is prepared in coexistence with its vapor, and a hard-sphere solute is inserted in the slab's

---

the polymer) and (b) is everywhere parallel to the average direction of low-temperature diffusive motion in the reduced free energy landscape. It captures the dominant transition path between two states separated by a free energy barrier.



(a) (b)

Figure 2.4: (a) Free energy for a model hydrophobic polymer in a coarse-grained model of water, as a function of radius of gyration, R_g (Å), and size of largest contiguous vapor-like region, U (arbitrary units of volume). From Ref. [106]. The dots trace configurations observed in representative hydrophobic collapse trajectories, with the transition state of each one marked in green. The transition states involve the formation of a vapor bubble in the liquid. (b) Free energy along minimum free energy path for a collapse of the same model hydrophobic polymer, in explicit SPC/E water. The collective variables are the polymer configurations and a coarse-grained solvent density. The barrier to collapse is dominated by solvent motions, not solute motions. From Ref. [78].

center. The field $n(\mathbf{r})$ is essentially uniform in the liquid and the vapor regions, but capillary-wave fluctuations are clearly observed at the slab's edge and around the larger solute.

The $n(\mathbf{r})$ field describes the solvent on a mesoscopic scale, where a region of space is either liquid-like or vapor-like, with thin interfacial regions in between them. Such a behavior is well modeled by a Landau-Ginzburg Hamiltonian,

$$H_L[n(\mathbf{r})] = \int_{\mathbf{r}} w(n(\mathbf{r}), \mu) + \frac{m}{2} \left|\nabla n(\mathbf{r})\right|^2 . \tag{2.31}$$

Here and throughout we use an abbreviated integration notation for clarity. The function $w(\rho/\rho_\ell, \mu)$ is the grand free energy density for the liquid solvent at a given density, ρ, and chemical potential, μ, relative to that of the gas. It has two minima, at $n = 0$ and $n = 1$, corresponding to vapor-like and liquid-like phases, with a large barrier between them. The square-gradient term reflects the inability of the $n(\mathbf{r})$ field to switch instantaneously from liquid-like to vapor-like values. The value of m and the shape of the barrier in $w(n, \mu)$ determine the shape and width of the interfacial region, as well as the effective surface tension of the vapor-liquid interface [93].

In a region where $n(\mathbf{r})$ is uniformly liquid-like, the field $\delta\rho(\mathbf{r})$ has Gaussian statistics, as discussed previously. We assume that even in the non-uniform case, $\delta\rho(\mathbf{r})$ has Gaussian statistics, although the fluctuation spectrum might depend on the field $n(\mathbf{r})$. Hence, we have a model Hamiltonian for this fluctuating field

$$H_S[\delta\rho(\mathbf{r}); n(\mathbf{r})] = \frac{k_\mathrm{B}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \delta\rho(\mathbf{r}) \chi^{-1}\big(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})\big) \delta\rho(\mathbf{r}'), \tag{2.32}$$

where $\chi^{-1}\big(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})\big)$ is defined by the relation

$$\int_{\mathbf{r}'} \chi^{-1}\big(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})\big) \chi\big(\mathbf{r}', \mathbf{r}''; n(\mathbf{r})\big) = \delta(\mathbf{r} - \mathbf{r}''), \quad \text{for all } \mathbf{r},\ \mathbf{r}'', \tag{2.33}$$

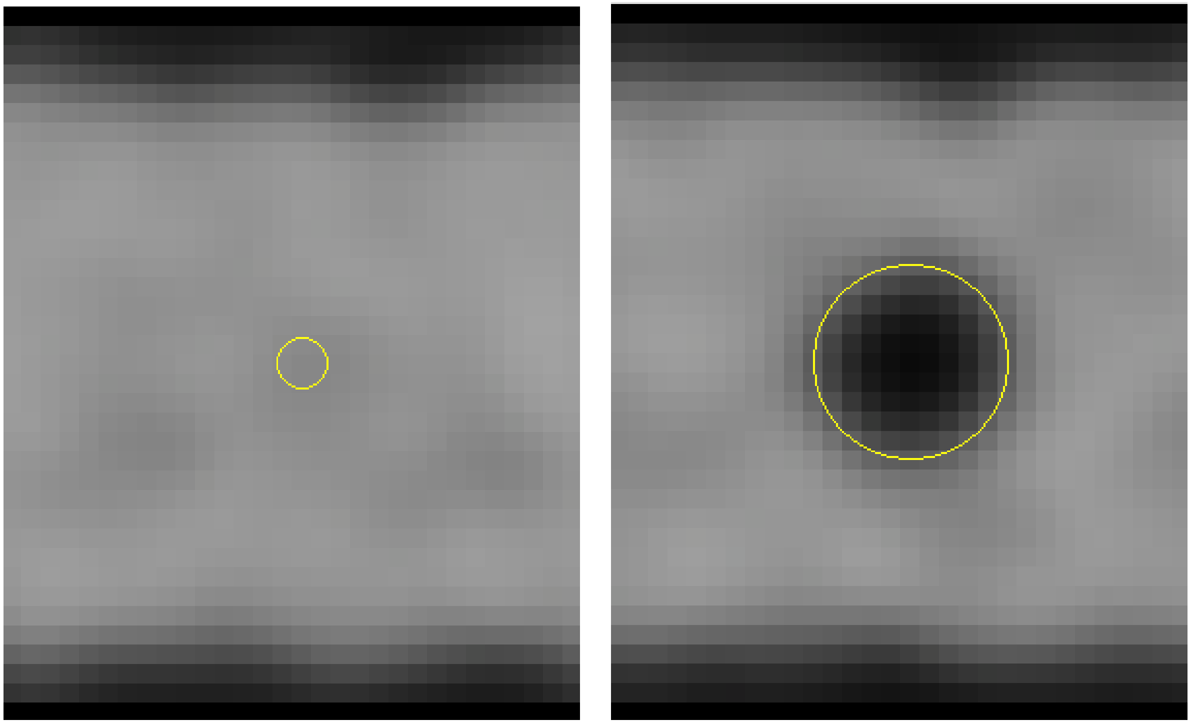
and $\chi\big(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})\big)$ is, by this construction, the fluctuation spectrum of $\delta\rho(\mathbf{r})$ in the presence of a particular configuration of the slowly-varying field,

$$\chi\big(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})\big) = \left\langle \delta\rho(\mathbf{r}) \delta\rho(\mathbf{r}') \right\rangle_{[n(\mathbf{r})]} . \tag{2.34}$$

In LCW theory, this fluctuation spectrum is assumed to be modeled well by the following formula, which is exact for uniform $n(\mathbf{r}) = 0$ and $n(\mathbf{r}) = 1$, and interpolates smoothly between these two limits:

$$\chi\big(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})\big) \approx \rho_\ell n(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') + \rho_\ell^2 n(\mathbf{r}) [g(|\mathbf{r} - \mathbf{r}'|) - 1] n(\mathbf{r}'). \tag{2.35}$$

For clarity, we henceforth omit the dependence of $\chi(\mathbf{r}, \mathbf{r}')$ on the field $n(\mathbf{r})$.



(a) (b)

Figure 2.5: Approximate $n(\mathbf{r})$-like fields derived from a full density field $\rho(\mathbf{r})$ extracted from an atomistically detailed simulation of water. A $60 \times 60 \times 60\,\text{Å}^3$ slab of SPC/E water is set up in a $60 \times 60 \times 70\,\text{Å}^3$ periodic box and evolved in time. At a fixed time, an $n(\mathbf{r})$-like field is constructed by convoluting the atomistic water density with a Gaussian of width 4 Å. Shown is a cross-section of the density field in the presence of a sphere at the center of the slab of radius (a) 2.5 Å, or (b) 10 Å (yellow outline). The bulk region of the liquid is essentially smooth, but liquid-vapor interfaces are clearly seen at the slab edges and around the 10 Å hard sphere.

Finally, when there are large-length-scale inhomogeneities in the solvent density, there arises an effective coupling between the small- and large-length-scale solvent densities. In that case, the total Hamiltonian for the density fields $n(\mathbf{r})$ and $\delta\rho(\mathbf{r})$ is

$$H[n(\mathbf{r}),\delta\rho(\mathbf{r})] = H_L[n(\mathbf{r})] + H_S[\delta\rho(\mathbf{r}); n(\mathbf{r})] + H_I[n(\mathbf{r}),\delta\rho(\mathbf{r})], \tag{2.36}$$

where

$$H_I[n(\mathbf{r}),\delta\rho(\mathbf{r})] = \int_{\mathbf{r}}\int_{\mathbf{r}'} [n(\mathbf{r})-1]u(\mathbf{r},\mathbf{r}')\delta\rho(\mathbf{r}') + H_{\mathrm{norm}}[n(\mathbf{r})]. \tag{2.37}$$

Here, $u(\mathbf{r},\mathbf{r}')$ is a short-ranged, translationally- and rotationally-invariant coupling strength, discussed below in detail. The specific form is chosen so that $H_I[n(\mathbf{r}),\delta\rho(\mathbf{r})]$ is zero for the uniform liquid $n(\mathbf{r}) = 1$, and $H_{\mathrm{norm}}[n(\mathbf{r})]$ is chosen so that

$$\int \mathcal{D}\delta\rho(\mathbf{r})\, e^{-\beta H[n(\mathbf{r}),\delta\rho(\mathbf{r})]} = e^{-\beta H_L[n(\mathbf{r})]}. \tag{2.38}$$

It can be shown that up to an irrelevant additive constant, $H_{\mathrm{norm}}[n(\mathbf{r})]$ is given by the equation

$$H_{\mathrm{norm}}[n(\mathbf{r})] = \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'}\int_{\mathbf{s}}\int_{\mathbf{s}'} [n(\mathbf{r})-1]\beta u(\mathbf{r},\mathbf{s})\chi(\mathbf{s},\mathbf{s}')\beta u(\mathbf{s}',\mathbf{r}')[n(\mathbf{r}')-1]. \tag{2.39}$$

These and other derivations are shown explicitly in Chapter 4.

The physical origin of this coupling has been worked out rigorously by Weeks and coworkers [114, 118, 120–122] for simple liquids in terms of "unbalancing forces", as explained in detail in Appendix A. Physically, solvent molecules deep in the bulk liquid are attracted equally in all directions, so they experience zero net force, but particles near a liquid-vapor interface feel an unbalanced attraction towards the bulk, as shown in Figure 2.6. When the total density is divided into slowly and quickly varying portions, the unbalanced attractions between the slowly varying portions give rise to the square-gradient term in Equation (2.31), and the remaining coupling to the quickly varying field $\delta\rho(\mathbf{r})$ gives rise to the term $H_I[n(\mathbf{r}),\delta\rho(\mathbf{r})]$. The coupling strength $u(\mathbf{r},\mathbf{r}')$ describes the short-range attractions between solvent molecules, and the results of LCW theory are only appreciably sensitive to the energy and length scale of this attraction, not its exact form. To bring out these considerations explicitly, it is useful to rewrite the coupling in terms of an unbalancing potential, $\phi(\mathbf{r})$, given by

$$\phi(\mathbf{r}) = -2a\rho_\ell[\overline{n}(\mathbf{r}) - 1]. \tag{2.40}$$

Here, the "smearing" overbar operator smooths out the field $n(\mathbf{r})$ over a few Angstroms, the range of the coupling strength $u(\mathbf{r},\mathbf{r}')$, and the parameter a characterizes the strength of the coupling. Notice that $\phi(\mathbf{r})$ depends on $n(\mathbf{r})$, but we will generally omit this dependence for clarity. With this potential, the interaction Hamiltonian can be written as

$$H_I[n(\mathbf{r}),\delta\rho(\mathbf{r})] = \int_{\mathbf{r}} \phi(\mathbf{r})\delta\rho(\mathbf{r}) + H_{\mathrm{norm}}[n(\mathbf{r})], \tag{2.41}$$

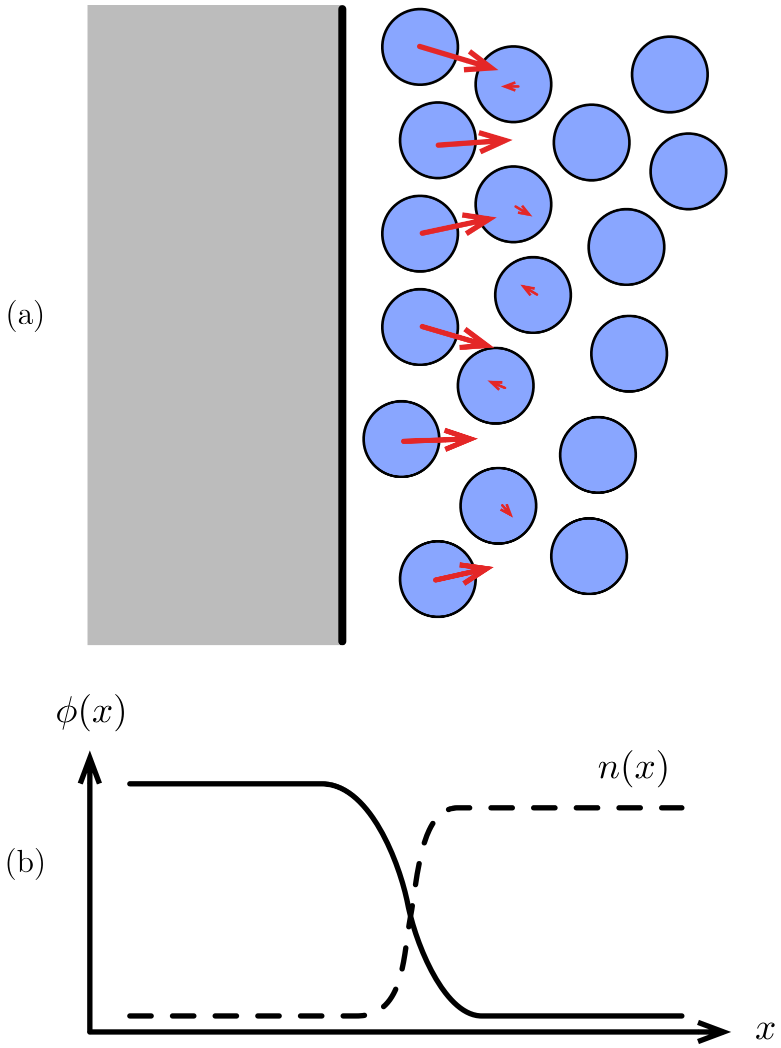


Figure 2.6: (a) Around a density inhomogeneity, such as near a hard wall (grey), solvent molecules (blue disks) experience unbalanced attractions, which result in effective forces (red arrows) towards higher density regions. (b) Such forces can be modeled by coupling an unbalancing potential $\phi(\mathbf{r})$ (solid), given by Equation (2.40), to the solvent density (dashes).

with

$$H_{\mathrm{norm}}[n(\mathbf{r})] = \frac{k_{\mathrm{B}}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \phi(\mathbf{r}) \chi(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}'). \tag{2.42}$$

Since the formal derivation of the unbalancing potential coupling doesn't extend to complex molecular liquids like water, the coupling shown here is to be regarded as a phenomenological model for the physically justified unbalancing forces that we expect are common to all cohesive liquids. The parameter a and the effect of the smearing operator $\overline{n}(\mathbf{r})$ are chosen either on physical grounds or to fit the results of a representative set of LCW calculations to results from atomistically detailed simulations.

The three energy terms discussed above, characterizing large- and small-length-scale density fluctuations and their coupling, form the basis of LCW theory. They describe the free energy cost of setting up a particular configuration of the slowly-varying density field $n(\mathbf{r})$ and the quickly-varying field $\delta\rho(\mathbf{r})$. For future reference, we group all the terms in a single equation,

$$\begin{aligned} H[n(\mathbf{r}), \delta\rho(\mathbf{r})] = &\int_{\mathbf{r}} w(n(\mathbf{r}), \mu) + \frac{m}{2} \left|\nabla n(\mathbf{r})\right|^2 && (= H_L[n(\mathbf{r})]) \\ &+ \frac{k_{\mathrm{B}}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \delta\rho(\mathbf{r}) \chi^{-1}(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})) \delta\rho(\mathbf{r}') && (= H_S[\delta\rho(\mathbf{r}); n(\mathbf{r})]) \\ &+ \int_{\mathbf{r}} \phi(\mathbf{r}) \delta\rho(\mathbf{r}) + H_{\mathrm{norm}}[n(\mathbf{r})] && (= H_I[n(\mathbf{r}), \delta\rho(\mathbf{r})]), \end{aligned} \tag{2.43}$$

where $\phi(\mathbf{r})$ is given in Equation (2.40) and $\chi(\mathbf{r}, \mathbf{r}'; n(\mathbf{r}))$ is given by Equation (2.24). The Hamiltonian is built so that when the field $\delta\rho(\mathbf{r})$ is thermally averaged, the resulting effective Hamiltonian is $H_L[n(\mathbf{r})]$ (see Equation (2.38)).

Chandler has shown [11] that the presence of an ideal solute that excludes solvent from a region v can be effectively modeled as the constraint $\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r}) = 0$ for all points $\mathbf{r}$ in v. When the field $\delta\rho(\mathbf{r})$ is thermally averaged subject to this constraint an effective Hamiltonian $H_{\mathrm{eff}}[n(\mathbf{r})]$ results, as follows

$$e^{-\beta H_{\mathrm{eff}}[n(\mathbf{r})]} = \int \mathcal{D}\delta\rho(\mathbf{r})\, e^{-\beta H[n(\mathbf{r}), \delta\rho(\mathbf{r})]} \prod_{\mathbf{r} \in v} \delta\big(\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})\big). \tag{2.44}$$

Since the Hamiltonian $H[n(\mathbf{r}), \delta\rho(\mathbf{r})]$ is quadratic in $\delta\rho(\mathbf{r})$, the integral can be performed analytically, as shown in Chapter 4. The result is

$$\begin{aligned} H_{\mathrm{eff}}[n(\mathbf{r})] = H_L[n(\mathbf{r})] &+ \frac{k_{\mathrm{B}}T}{2} \ln\det(2\pi\chi_v) \\ &+ \frac{k_{\mathrm{B}}T}{2} \int_{\mathbf{r} \in v} \int_{\mathbf{r}' \in v} \left[\rho_\ell n(\mathbf{r}) - \int_{\mathbf{s}} \chi(\mathbf{r}, \mathbf{s}) \beta\phi(\mathbf{s})\right] \\ &\qquad \chi_v^{-1}(\mathbf{r}, \mathbf{r}') \left[\rho_\ell n(\mathbf{r}') - \int_{\mathbf{s}'} \chi(\mathbf{r}', \mathbf{s}') \beta\phi(\mathbf{s}')\right], \end{aligned} \tag{2.45}$$

where $\chi_v^{-1}(\mathbf{r},\mathbf{r}')$ is the operator that satisfies

$$\int_{\mathbf{r}''\in v} \chi_v^{-1}(\mathbf{r},\mathbf{r}'')\chi(\mathbf{r}'',\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'), \qquad \text{for all } \mathbf{r},\, \mathbf{r}' \text{ in } v. \tag{2.46}$$

Equation (2.45) should be compared to the Gaussian model of Equation (2.27). The expression in square brackets is the linear response expression for the average solvent density at a point in the presence of the external field $\phi(\mathbf{r})$, so its integral over the volume v is the average number of solvent particles in v. The last two terms of Equation (2.45) are thus revealed to be the multi-variable generalization of the Gaussian model expression of Equation (2.27).

The original LCW theory [68] is a mean field theory solution to Equation (2.45), as described in Appendix A. If the mean-field approximation is omitted, then Equation (2.45) encodes valuable information about the energetics of large-length-scale fluctuations, as first observed by Ten Wolde, Sun and Chandler [107].

Using the LCW Hamiltonian, the values of interesting observables can be immediately calculated. For example, the solvation free energy of an ideal solute that excludes solvent from volume v is given by

$$\Delta G_{\text{sol}} = -k_{\text{B}}T \ln \left\langle e^{-\beta(H_{\text{eff}}[n(\mathbf{r})]-H_L[n(\mathbf{r})])}\right\rangle_{H_L}. \tag{2.47}$$

For small solutes, where the $n(\mathbf{r}) = 1$ configuration dominates the thermal average, this formula reverts to the Gaussian model of small-solute solvation, Equation (2.27). Conversely, for large solutes, the field $n(\mathbf{r})$ is most likely 0 inside the solute and 1 outside, and the solvation free energy is dominated by the interfacial energetics, captured by the term $H_L[n(\mathbf{r})]$. The theory provides a physically sound interpolation between the small- and large-solute limits, without restrictions to simple geometries like spheres. In fact, LCW theory was the first quantitative theory of solvation to possess all of these attributes.

Another interesting observable is the average solvent density at a point $\mathbf{r}$, shown in Chapter 4 to be

$$\langle\rho(\mathbf{r})\rangle = \left\langle \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'\in v}\int_{\mathbf{s}\in v} \chi(\mathbf{r},\mathbf{r}')\chi_v^{-1}(\mathbf{r}',\mathbf{s}) \left[\rho_\ell n(\mathbf{s}) - \int_{\mathbf{s}'} \chi(\mathbf{s},\mathbf{s}')\beta\phi(\mathbf{s}')\right]\right\rangle_{H_{\text{eff}}}. \tag{2.48}$$

For large solutes, the mean density is depleted just outside the solute, as would be expected from the entropy of capillary waves on the liquid-vapor interface around such solutes. When additional solute-solvent interactions are considered, this depletion layer disappears [48].

LCW theory is flexible enough to encompass microscopic phase transitions, since there is no a priori reason why $H_{\text{eff}}[n(\mathbf{r})]$ should not have more than one prominent minimum. As such, it captures the dewetting phenomena that play a critical role in assembly of model hydrophobic plates [51, 115], and the collapse of model hydrophobic polymers [78, 106], as discussed above.

Since LCW theory captures all of these important solvation phenomena, it is an attractive candidate for modeling solvation in more complicated contexts, such as protein folding, drug discovery and self-assembly, where treating solvation too crudely is inappropriate and treating it explicitly is too expensive to be practical. A principal thrust of this thesis is to construct a tractable, efficient theory of solvation based on LCW ideas, which may eventually be used in these contexts.

Chapter 3

Molecular-scale probes for solvent density fluctuations

In this chapter, we develop tools to probe solvent density fluctuations in computer simulations of water, and discuss the implications of the results we find. In previous work, such density fluctuations have been probed in Lennard-Jones fluids and hard-sphere fluids using Monte Carlo simulations and umbrella sampling [18, 46]. When the solvent is water, simulations of equal statistical value are orders of magnitude more expensive to perform, so it's important to distribute the computational effort over many computers. In this context, the straightforward methodology used in Refs. [18, 46] becomes impractical. An important objective of this chapter is to develop practical alternatives [84, 85].

The main tool we use for probing solvent density fluctuations are water number distributions, denoted $P_v(N)$, which measure the probability of finding N water molecules in a probe volume v. These are the distributions that, for small v, were examined by Hummer et al [53] and discussed in Section 2.2.1. When v is large, the distributions are qualitatively different: for N significantly below $\langle N \rangle_v$, the distributions deviate from the Gaussian behavior that occurs around the mean value. These deviations, or "fat tails", belie the formation of a vapor bubble inside v, as it is much easier to grow this vapor bubble than it is to isotropically rarify the bulk liquid. When v is located near a hydrophobic object, the fat tails appear at N much closer to $\langle N \rangle_v$. This difference reflects the presence of a preformed liquid-vapor interface around the hydrophobic object, which can simply be enlarged to encompass v and evacuate the water molecules it contains. The consequences of this observation are apparent when two hydrophobic objects are brought close together. Water number distributions between two model hydrophobic plates are bistable, as anticipated by LCW theory. One stable state is a "wet" state, where the interplate region is filled with water, and the other is a corresponding "dry" state. Such bistabilities are present in more realistic systems, such as nearby dimers of the protein melittin, the main component of bee venom. Melittin dimers have a large hydrophobic patch that mediates assembly into tetramers, as has been observed in dynamical simulations [64]. Mutations to this protein affect the dewetting transition in observable ways.

The remainder of this chapter is devoted to extensions of the methods we develop to probe density fluctuations. First, to further test whether small fluctuations in water and other liquids are truly Gaussian, we develop tools to probe the distribution functions for the Fourier components of the solvent density field. These tools are generally applicable in other contexts. Second, we develop tools to study fluctuations of the liquid-vapor interface. Our efforts are motivated by a recent development in identifying instantaneous liquid-vapor interfaces [125]. In Chapter 6, we use the tools we develop here to compute the potential of mean force for progressively pulling one water molecule out of the bulk liquid, an important object in water evaporation. These tools are also generally applicable, for instance, to study the driving forces for ions to move towards or away from liquid-vapor interfaces [59, 82, 87].

3.1 Umbrella sampling

In studying the thermodynamics of a system, many quantities of interest, including probability density functions, may be expressed as thermal averages of an observable. These averages may be estimated through sampling, as follows. A set of M statistically independent configurations of the system are first generated with a probability given by their Boltzmann weight [34], with the m^{th} such configuration denoted by $\mathbf{R}^N_m$. The thermal average of an observable $A(\mathbf{R}^N)$ is then estimated as

$$\langle A \rangle \approx \frac{1}{M} \sum_{m=1}^{M} A(\mathbf{R}^N_m). \tag{3.1}$$

The statistical error associated with this estimate is small if the configurations that dominate the Boltzmann ensemble also have values of A that are close to $\langle A \rangle$, but is much larger if the important configurations are badly sampled.

One common observable that is hard to estimate in this manner is the distribution function for A, $P(A)$, defined as

$$P(A) = \left\langle \delta(A(\mathbf{R}^N) - A) \right\rangle . \tag{3.2}$$

Practically, it is more convenient to bin this probability distribution, i.e., measure its integral over small, consecutive ranges of values of width ΔA. When doing this, the observable being estimated is

$$P(A)\Delta A \approx \left\langle h_{\Delta A}(A(\mathbf{R}^N) - A) \right\rangle , \tag{3.3}$$

where

$$h_{\Delta A}(x) = \begin{cases} 1, & -\Delta A/2 < x < +\Delta A/2, \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

When $P(A)\Delta A$ is exponentially small, the straightforward sampling estimate is extremely unreliable, because a tiny number is being estimated as the average of a large number of 0s and an exponentially small number of 1s.

Umbrella sampling, known as importance sampling in the general context of Monte Carlo integration, is a technique to avoid the problems of rare but important configurations. The key idea is to draw samples from a distribution that is not the Boltzmann distribution, but where the important configurations are well-represented, and then weighing any averages to recover Boltzmann statistics. Formally, this procedure corresponds to the trivial identity

$$\langle A\rangle_U = \frac{\int \mathrm{d}\mathbf{R}^N\, e^{-\beta U(\mathbf{R}^N)} A(\mathbf{R}^N)}{\int \mathrm{d}\mathbf{R}^N\, e^{-\beta U(\mathbf{R}^N)}} \tag{3.5a}$$

$$= \frac{\int \mathrm{d}\mathbf{R}^N\, e^{-\beta U(\mathbf{R}^N)-\beta\Delta U(\mathbf{R}^N)} A(\mathbf{R}^N) e^{\beta\Delta U(\mathbf{R}^N)}}{\int \mathrm{d}\mathbf{R}^N\, e^{-\beta U(\mathbf{R}^N)-\beta\Delta U(\mathbf{R}^N)}} \Bigg/ \frac{\int \mathrm{d}\mathbf{R}^N\, e^{-\beta U(\mathbf{R}^N)-\beta\Delta U(\mathbf{R}^N)} e^{+\beta\Delta U(\mathbf{R}^N)}}{\int \mathrm{d}\mathbf{R}^N\, e^{-\beta U(\mathbf{R}^N)-\beta\Delta U(\mathbf{R}^N)}} \tag{3.5b}$$

$$= \frac{\left\langle A e^{\beta\Delta U}\right\rangle_{U+\Delta U}}{\left\langle e^{\beta\Delta U}\right\rangle_{U+\Delta U}}. \tag{3.5c}$$

In other words, sampling is performed in the Boltzmann ensemble of a system with potential energy $U + \Delta U$ instead of U. The ratio of the thermal averages of $Ae^{\beta\Delta U}$ and $e^{\beta\Delta U}$ yields the thermal average of A in the original Boltzmann ensemble. With an appropriate choice of ΔU, the configurations that dominate the average of A can be sampled well.

The above discussion may be generalized substantially. Thermal averages may be estimated with great precision by collecting samples in many different but related ensembles, and then combining all the information in an optimal way. The resulting algorithm, the Multi-state Bennett Acceptance Ratio (MBAR) [97], is very powerful and is critical to the rest of this thesis, so we discuss it in detail below. It generalizes the closely related Weighted Histogram Analysis Method (WHAM) [32]. Both algorithms also provide an optimal estimate of the free energy difference between the sampled ensembles.

The derivation of MBAR that we present below differs substantially from the original presentation and is inspired by similar derivations of related methods [2, 60, 96]. To the author's knowledge, however, it is not available in this form in the literature.

3.1.1 Maximum-likelihood derivation of MBAR equations

We begin with a Boltzmann ensemble, called the "unbiased" ensemble, defined by a potential energy U, whose partition function is Z and related free energy is F. These quantities are given by

$$Z = \int \mathrm{d}\mathbf{R}^N\, \exp[-\beta U(\mathbf{R}^N)], \tag{3.6}$$

$$F = -k_\mathrm{B} T \ln Z. \tag{3.7}$$

We then have K biasing potentials, also called umbrella potentials, labeled V_1 through V_k. For each biasing potential V_k, we collect N_k statistically independent samples, labeled $\mathbf{R}^N_{k,n}$, that are Boltzmann-weighted with the potential $U + V_k$. The partition function, Z_k, and associated free energy, F_k, of this ensemble are given by

$$Z_k = \int \mathrm{d}\mathbf{R}^N \, \exp[-\beta U(\mathbf{R}^N) - \beta V_k(\mathbf{R}^N)], \tag{3.8}$$

$$F_k = -k_\mathrm{B}T \ln Z_k, \tag{3.9}$$

We denote by ΔF_k the free energy difference between this ensemble and the unbiased ensemble, given by

$$\Delta F_k = F_k - F = k_\mathrm{B}T \ln(Z/Z_k). \tag{3.10}$$

As in any sampling method, a model for the full probability density of the unbiased ensemble, $P(\mathbf{R}^N)$, is constructed as a sum of Dirac δ-functions centered at each of the $N_1 + \cdots + N_K$ samples observed, with unknown weights $p_{j,n}$. In other words,

$$P(\mathbf{R}^N) \approx \mathcal{Z}^{-1} \sum_{j=1}^{K} \sum_{n=1}^{N_j} p_{j,n} \delta(\mathbf{R}^N - \mathbf{R}^N_{j,n}). \tag{3.11}$$

The probability density of the k^th biased ensemble, $P_k(\mathbf{R}^N)$, is then given by

$$P_k(\mathbf{R}^N) \approx \mathcal{Z}_k^{-1} \sum_{j=1}^{K} \sum_{n=1}^{N_j} p_{j,n} e^{-\beta V_k(\mathbf{R}^N)} \delta(\mathbf{R}^N - \mathbf{R}^N_{j,n}). \tag{3.12}$$

The normalization constants $\mathcal{Z}$ and $\mathcal{Z}_k$ are given by

$$\mathcal{Z} = \sum_{j=1}^{K} \sum_{n=1}^{N_j} p_{j,n}, \tag{3.13a}$$

$$\mathcal{Z}_k = \sum_{j=1}^{K} \sum_{n=1}^{N_j} p_{j,n} e^{-\beta V_k(\mathbf{R}^N_{j,n})}. \tag{3.13b}$$

Up to an undetermined measure factor, these normalization constants are the best estimates of the partition functions Z and Z_k corresponding to the sample weights $\{p_{j,n}\}$, so that $\mathcal{Z}/\mathcal{Z}_k \approx Z/Z_k$. The choice of weights $\{p_{j,n}\}$ parametrizes the model for the underlying probability distributions. Given the observed data, we can calculate the likelihood L of having observed that data by the relation

$$L = \prod_{j=1}^{K} \left[\prod_{n=1}^{N_k} P_j(\mathbf{R}^N_{j,n}) \right]. \tag{3.14}$$

The data has a finite probability of being observed no matter what the model is. If we regard the model as a random variable for which we have no preconceived notion, then Bayes' theorem tells us that the probability of observing the data given the model is proportional to the probability of the model being correct given the observed data. Hence, the choice of weights that maximizes the former, i.e., maximizes L, is also the most likely model that explains the data.

We thus proceed to find the choice of weights $\{p_{j,n}\}$ that maximizes L, or equivalently, that maximizes $\ln L$. We have

$$\ln L = \sum_{j=1}^{K} \sum_{n=1}^{N_k} \ln(p_{j,n} e^{-\beta V_j(\mathbf{R}^N_{j,n})}/\mathcal{Z}_j) + \text{const.} \tag{3.15}$$

Setting $\partial \ln L/\partial p_{j,n} = 0$ yields the equation

$$\frac{1}{p_{j,n}} - \sum_{k=1}^{K} N_k \frac{e^{-\beta V_k(\mathbf{R}^N_{j,n})}}{\mathcal{Z}_k} = 0. \tag{3.16}$$

Solving for $p_{j,n}$ and dividing throughout by $\mathcal{Z}$ yields

$$\frac{p_{j,n}}{\mathcal{Z}} = \frac{1}{\sum_{k=1}^{K} N_k \exp\left[\beta \Delta F_k - \beta V_k(\mathbf{R}^N_{j,n})\right]}. \tag{3.17}$$

Hence, an estimate of the free energy difference between the biased ensembles and the unbiased ensemble is enough to estimate the probability of *every* observed sample in the unbiased ensemble. Using Equation (3.13), we find the following set of K equations that relate these free energies to each other,

$$\Delta F_i = -k_{\mathrm{B}} T \ln \sum_{j=1}^{K} \sum_{n=1}^{N_j} \frac{\exp[-\beta V_i(\mathbf{R}^N_{j,n})]}{\sum_{k=1}^{K} N_k \exp\left[\beta \Delta F_k - \beta V_k(\mathbf{R}^N_{j,n})\right]}. \tag{3.18}$$

These relations are known as the MBAR equations [97], and encompass the WHAM equations of Ref. [32]. They can be solved by self-consistent iteration, and are guaranteed to converge regardless of what starting guess is used for $\{\Delta F_k\}$. The solution is defined up to an additive constant, which can be fixed in many different ways. A clean, symmetric choice is to demand that $\mathcal{Z}$ be 1.

Once the free energy differences $\{\Delta F_k\}$ have been determined, the probability of observing each sample in the unbiased ensemble can be determined using Equation (3.17). The MBAR estimate of $\langle A \rangle$ is then given by

$$\langle A \rangle \approx \mathcal{Z}^{-1} \sum_{j=1}^{K} \sum_{n=1}^{N_j} p_{j,n} A(\mathbf{R}^N_{j,n}). \tag{3.19}$$

This is the procedure we use for estimating all mean values and observable distributions in the rest of the thesis. The umbrellas we use are typically of the form

$$V_k(\mathbf{R}^N) = \frac{\kappa}{2}(A(\mathbf{R}^N) - A_k)^2. \tag{3.20}$$

With such an umbrella potential, we can sample well configurations with values of A clustered within about $1/\sqrt{\beta\kappa}$ of a mean value that is close to A_k.

Since MBAR is a maximum-likelihood estimator for free energy differences and thermal averages, it has a number of advantages over most other estimators. In particular, it is the asymptotically unbiased estimator with smallest statistical uncertainty [97]. For this reason, it is preferable to use MBAR to any other method (such as Equation (2.5b)) when estimating free energy differences between many ensembles. Indeed, when only two umbrellas are used, MBAR reduces to the Bennett Acceptance Ratio (BAR) estimator for the free energy difference between the two biased ensembles, for which the optimality proof is transparent [4]. The performance of MBAR depends on the amount of "overlap" between the K ensembles. Two ensembles A and B "overlap" if the probability of observing a configuration in A is not too different from that of observing the configuration in B. The free energy difference between A and B is only well estimated if there is a sequence of ensemble between A and B, each of which overlaps with the previous and subsequent ensembles.

We now turn to two important details: sample correlation and error bar estimation. In general, it is not known how to draw statistically independent samples from a Boltzmann distribution. Instead, computer simulations use either Monte Carlo procedures, such as Metropolis sampling [76], or molecular dynamics, i.e., Newtonian evolution of the system's equations of motion [34]. In both cases, consecutive samples are highly correlated. The correlation can be quantified by a correlation time τ_k, with units of samples, of some representative observable, such as the value of the umbrella potential $V_k(\mathbf{R}^N)$. Accounting for this correlation results in a simple modification of the MBAR equations [32], namely

$$\frac{p_{j,n}}{Z} = \frac{1/g_j}{\sum_{k=1}^{K} N_k/g_k \exp\left[\beta\Delta F_k - \beta V_k(\mathbf{R}^N_{j,n})\right]}, \tag{3.21a}$$

$$\Delta F_i = -k_\mathrm{B}T \ln \sum_{j=1}^{K} \frac{1}{N_j} \sum_{n=1}^{N_j} \frac{N_j/g_j \exp[-\beta V_i(\mathbf{R}^N_{j,n})]}{\sum_{k=1}^{K} N_k/g_k \exp\left[\beta\Delta F_k - \beta V_k(\mathbf{R}^N_{j,n})\right]}, \tag{3.21b}$$

where g_k is the so-called statistical inefficiency of the sampling in ensemble k, defined as

$$g_k = 1 + 2\tau_k. \tag{3.22}$$

The precise value of g_k depends on the observable used, but ratios of these statistical inefficiencies, the only quantities that appear in Equation (3.21), are more robust.

In principle, an alternative strategy would be to subsample the collected data in each umbrella k by keeping only one of every g_k samples. These samples are then independent, so Equations (3.17) and (3.18) can be used directly. Moreover, in the large-sample limit, the likelihood L (Equation (3.14)) is approximately Gaussian, so error bars may be estimated directly by computing the Hessian of L with respect to the weights $\{p_{j,n}\}$. This is the approach taken in Ref. [97]. We have found this approach unsatisfactory because of the difficulty of estimating g_k reliably, which results in underestimation of the error bars. When MBAR is applied independently in this way to two halves of the sampled data, the estimates for thermal averages in each half differ by much more than would be expected from the size of the error bar estimates. Instead, we have chosen a more straightforward and conservative method for estimating errors. We divide our sample data into a number of blocks (usually 6 to 8), each of which is reasonably independent, then apply MBAR independently on each set, and then calculating the mean and standard error of the independent MBAR estimates of thermal averages. Unlike the direct error bar estimate, this procedure does yield consistent error estimates when the data is split into two halves and analyzed independently.

3.2 Probing density fluctuations in real space

In this section, we describe a method for computing water number distributions, i.e., $P_v(N)$. For concreteness, we define the position of a water molecule to be the position of its oxygen atom, which nearly coincides with its center of mass and its center of electron density. Let O be the subset of system atoms that are water oxygen atoms. Given a configuration $\mathbf{R}^N$, we define the observable $N(\mathbf{R}^N)$ by the relation

$$N(\mathbf{R}^N) = \sum_{i\in O} \int_{\mathbf{r}\in v} \delta(\mathbf{R}_i - \mathbf{r}). \tag{3.23}$$

It is useful to recast this expression in terms of an indicator function $h(\mathbf{r})$ defined as

$$h(\mathbf{r}) = \int_{\mathbf{r}'\in v} \delta(\mathbf{r}' - \mathbf{r}). \tag{3.24}$$

This function takes the value 1 when $\mathbf{r}$ is in v, and 0 otherwise. With this notation, we have

$$N(\mathbf{R}^N) = \sum_{i\in O} h(\mathbf{r}_i). \tag{3.25}$$

In principle, we could use MBAR to calculate $P_v(N)$, using umbrellas of the form

$$V_k(\mathbf{R}^N) = \frac{\kappa}{2}(N(\mathbf{R}^N) - N_k)^2. \tag{3.26}$$

This procedure is straightforward to implement in a Monte Carlo simulation, where only $V_k(\mathbf{R}^N)$ needs to be computed, and is essentially the method used in Refs. [18, 46]. However, Monte Carlo simulations are hard to parallelize, since the Metropolis algorithm that underpins them is inherently serial.[1] An enormous community effort, on the other hand, has gone into implementing parallelized molecular dynamics (MD) codes, of which NAMD [88], LAMMPS [89] and GROMACS [44] are three of the most widely used.

In MD, system configurations are generated by propagating configurations according to the forces implied by the system's potential energy. We thus need to calculate $\nabla_j V_k(\mathbf{R}^N)$ for all atoms j in the system. This quantity is proportional to $\nabla_j N(\mathbf{R}^N)$, given by the following integral over the surface, ∂v, of the probe volume,

$$\nabla_j N(\mathbf{R}^N) = \begin{cases} \int_{\partial v} -\delta(\mathbf{R}_j - \mathbf{r}) \mathrm{d}\mathbf{A}, & j \in O, \\ 0, & \text{otherwise}. \end{cases} \tag{3.27}$$

The forces implied by the umbrella potential in Equation (3.26) are thus impulsive, and so cannot be treated by most existing MD codes. We now describe an alternative method that circumvents this issue.

3.2.1 Basic method: cubic cavities

We first construct an observable $\tilde{N}(\mathbf{R}^N)$ that correlates closely with $N(\mathbf{R}^N)$ but is a smooth, continuous function of particle positions. To this end, we replace the Dirac δ-function in Equation (3.24) by a finite-width approximation, denoted by $\Phi(\mathbf{r})$, resulting in a smooth indicator function, $\tilde{h}(\mathbf{r})$. We obtain

$$\tilde{h}(\mathbf{r}) = \int_{\mathbf{r}' \in v} \Phi(\mathbf{r}' - \mathbf{r}), \tag{3.28}$$

$$\tilde{N}(\mathbf{r}^N) = \sum_{i \in O} \tilde{h}(\mathbf{r}_i). \tag{3.29}$$

To be useful, the smearing function $\Phi(\mathbf{r})$ must satisfy several conditions. It must be normalized, so $\int_{\mathbf{r}} \Phi(\mathbf{r}) = 1$. For prescribed volumes v, we must be able to quickly and analytically calculate the integrals in Equation (3.29), as well as their gradients with respect to atomic positions. In order to correlate strongly with $N(\mathbf{R}^N)$, $\Phi(\mathbf{r})$ must be sharply localized. Finally, for the computational cost of calculating $\tilde{N}(\mathbf{R}^N)$ to scale with the size of v and not with the size of the entire system, $\Phi(\mathbf{r})$ should have finite support.

For the simplest case where v is a cube, all of the above conditions are satisfied when $\Phi(\mathbf{r})$ is a product of truncated Gaussians along each coordinate. Explicitly, our choice is

$$\Phi(\mathbf{r}) = \phi(x)\phi(y)\phi(z), \tag{3.30}$$

where $\mathbf{r} = (x, y, z)$ and

$$\phi(\alpha) = k^{-1}\big[e^{-\alpha^2/2\sigma^2} - e^{-\alpha_\mathrm{c}^2/2\sigma^2}\big]\Theta(\alpha_\mathrm{c} - |\alpha|). \tag{3.31}$$

---

[1] Parallelization is possible and beneficial in certain restricted cases, e.g., Ref. [33].

Here, σ is the width of the truncated Gaussian, α_{c} is distance from the mean at which the Gaussian is truncated, Θ is the Heaviside step function, and the normalization constant k has the value $\sqrt{2\pi\sigma^2}\,\mathrm{erf}(\alpha_{\mathrm{c}}/\sqrt{2\sigma^2}) - 2\alpha_{\mathrm{c}}\exp(-\alpha_{\mathrm{c}}^2/2\sigma^2)$. The generic label α denotes one of the coordinate axes, x, y or z. We generalize to different coordinate systems in Section 3.2.6.

For a cubic probe volume, we need to calculate the integral of $\phi(\alpha - \alpha_i)$ in the range $[\alpha_{\mathrm{min}}, \alpha_{\mathrm{max}}]$, where α_i is the α-coordinate of atom i. We denote this integral by $\tilde{h}_\alpha(\alpha_i)$, which is a smooth indicator function of whether the α-coordinate of particle i is within the range of α values that define the probe volume. Explicitly,

$$\tilde{h}_\alpha(\alpha_i) = \int_{\alpha_{\mathrm{min}}}^{\alpha_{\mathrm{max}}} \mathrm{d}\alpha\, \phi(\alpha - \alpha_i), \tag{3.32}$$

so

$$\begin{aligned}\tilde{h}_\alpha(\alpha_i) = &\left[k_1\,\mathrm{erf}\left(\frac{\alpha_{\mathrm{max}} - \alpha_i}{\sqrt{2\sigma^2}}\right) - k_2(\alpha_{\mathrm{max}} - \alpha_i) - \frac{1}{2}\right]\Theta(\alpha_{\mathrm{c}} - |\alpha_{\mathrm{max}} - \alpha_i|)\\ &+ \left[k_1\,\mathrm{erf}\left(\frac{\alpha_i - \alpha_{\mathrm{min}}}{\sqrt{2\sigma^2}}\right) - k_2(\alpha_i - \alpha_{\mathrm{min}}) - \frac{1}{2}\right]\Theta(\alpha_{\mathrm{c}} - |\alpha_i - \alpha_{\mathrm{min}}|)\\ &+ \Theta\left(\alpha_{\mathrm{c}} + \frac{1}{2}(\alpha_{\mathrm{max}} - \alpha_{\mathrm{min}}) - \left|\alpha_i - \frac{1}{2}(\alpha_{\mathrm{min}} + \alpha_{\mathrm{max}})\right|\right),\end{aligned} \tag{3.33}$$

where $k_1 = k^{-1}\sqrt{\pi\sigma^2/2}$ and $k_2 = k^{-1}\exp(-\alpha_{\mathrm{c}}^2/2\sigma^2)$. The derivative of $\tilde{h}_\alpha(\alpha_i)$ with respect to α_i is particularly simple, namely,

$$\tilde{h}'_\alpha(\alpha_i) = -[\phi(\alpha_{\mathrm{max}} - \alpha_i) - \phi(\alpha_{\mathrm{min}} - \alpha_i)]. \tag{3.34}$$

With these functions, the full indicator function $\tilde{h}(\mathbf{r})$ is given by

$$\tilde{h}(\mathbf{r}) = \tilde{h}_x(x)\tilde{h}_y(y)\tilde{h}_z(z), \tag{3.35}$$

where the boundaries of the cube define the ranges of integration along each coordinate. The value of the observable $\tilde{N}(\mathbf{R}^N)$ follows from Equation (3.29). Its gradient with respect to the position of atom j is

$$\nabla_j \tilde{N}(\mathbf{R}^N) = \begin{cases} \nabla\tilde{h}(\mathbf{r}_j), & j \in O,\\ 0, & \text{otherwise,}\end{cases} \tag{3.36}$$

where

$$\nabla\tilde{h}(\mathbf{r}) = \begin{pmatrix} \tilde{h}'_x(x)\tilde{h}_y(y)\tilde{h}_z(z)\\ \tilde{h}_x(x)\tilde{h}'_y(y)\tilde{h}_z(z)\\ \tilde{h}_x(x)\tilde{h}_y(y)\tilde{h}'_z(z)\end{pmatrix}. \tag{3.37}$$

This gradient is a finite and smooth function of particle coordinates, unlike the gradient in Equation (3.27). As a result, $\tilde{N}(\mathbf{R}^N)$ can be used to define an umbrella potential in MD simulations.

To determine $P_v(N)$ in a particular context, we employ umbrella potentials of the form

$$V_k(\mathbf{R}^N) = \frac{\kappa}{2}(\tilde{N}(\mathbf{R}^N) - \tilde{N}_k)^2, \tag{3.38}$$

which are incorporated into MD simulations as an additional force on each atom j given by

$$-\nabla_j V_k(\mathbf{R}^N) = -\kappa(\tilde{N}(\mathbf{R}^N) - \tilde{N}_k)\nabla_j \tilde{N}(\mathbf{R}^N). \tag{3.39}$$

In constant pressure simulations, these forces also contribute to the calculation of the virial [85].

Arbitrarily-shaped probe volumes v can be approximated by a union of small, disjoint cubic cavities. We make use of this simple extension below when looking a density fluctuations in melittin. The details of this and other extensions to cavities of general shape are discussed in Section 3.2.6.

We have implemented the calculation of $\tilde{N}(\mathbf{R}^N)$ in LAMMPS, NAMD and GROMACS. The evaluation of the right-hand side of Equation (3.29) is well-suited to parallelization. In particular, we exploit a design decision of all MD codes we have used, whereby each computational node "owns" a certain subset of the system's atoms. At all times, the positions of those atoms are available to the owning node without inter-node communication. Thus, in each node, we compute a partial sum of indicator functions $\tilde{h}(\mathbf{r}_i)$ for particles i that are both in O and owned by the node. Then, using a single global reduce operation, we add up these partial sums and distribute the result, the value of $\tilde{N}(\mathbf{R}^N)$, to all the nodes. The calculation of umbrella forces on each atom then proceeds independently in each node. Figure 3.1 shows that our NAMD code to calculate $\tilde{N}(\mathbf{R}^N)$ scales nearly perfectly with the number of computational nodes. Our LAMMPS and GROMACS codes perform similarly.

Typically, we prepare an initial set of simulations with umbrellas at $\tilde{N}_k = 0, 4, 8, \ldots$ up to about $\tilde{N}_k \approx 1.2\rho v$, with $\kappa = k_\mathrm{B}T/2^2$, so that there is sufficient overlap between consecutive umbrellas. The MBAR equations (3.18) are then solved self-consistently. Finally, the distribution $P_v(N)$, recast in the form

$$P_v(N) = \left\langle \delta_{N,N(\mathbf{R}^N)} \right\rangle, \tag{3.40}$$

is estimated from all the sampled data using Equation (3.19). If it becomes clear that there are insufficient statistics for a given N, additional umbrellas are employed. Empirically, we have found that setting $\alpha_c = 0.3\,\text{Å}$ and $\sigma = 0.1\,\text{Å}$ results in forces small enough to be handled by standard MD codes, and a correlation between N and $\tilde{N}$ that is strong enough for the umbrella runs to effectively sample configurations with all values of N. For some of the results below, the procedure we used was sometimes slightly different, but always in inconsequential ways (e.g., using $\alpha_c = 0.2\,\text{Å}$). Since these discrepancies don't influence the results, we don't state them explicitly.

One important consideration in all of the following computations regards the ensemble that is used. In order to allow the probe volume v to be fully evacuated, there must be

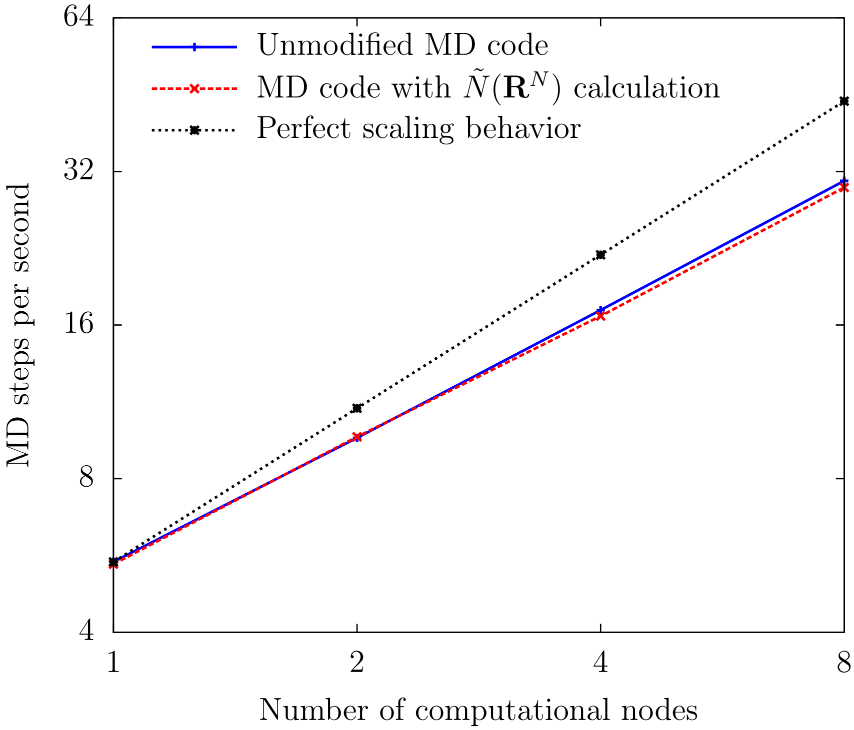


Figure 3.1: Performance of $\tilde{N}(\mathbf{R}^N)$ umbrella code, as implemented in NAMD. The test system is has nearly 4000 water molecules (modeled with the TIP3P potential) forming a $50 \times 50 \times 50\,\text{Å}^3$ slab in a $50 \times 50 \times 80\,\text{Å}^3$ periodic box. The probe volume v is a $10 \times 10 \times 10\,\text{Å}^3$ cube in the center of the water slab. Each MD step evolves the system dynamics over 2 fs. A typical sampling run consists of about 1 ns of data, or $500,000$ MD steps. Perfect scaling would be achieved if the overhead of nodes communicating with each other were negligible. NAMD is built to scale close to perfectly, and our code to calculate $\tilde{N}(\mathbf{R}^N)$ leaves this scaling behavior nearly intact.

room in the rest of the simulation box to accommodate the expelled waters. In a constant pressure simulation, the simulation box will automatically grow (at free energy cost $P\Delta V$) to accommodate the expelled waters. Implementing a constant pressure MD simulation, however, is technically challenging[2] and rife with subtle complications (for example, see Refs. [54,56,66,74,126]). In preferring simplicity, we instead perform most of our calculations in a fixed-volume simulation, with the important addition of a thin vapor layer in the box. This vapor layer ensures that the system's effective pressure is equal to the vapor pressure of water at the simulated temperature. Like atmospheric pressure, the vapor pressure is negligible at the molecular scale. The vapor layer, however, is large enough to accommodate any expelled water molecules, ensuring that the system remains at its vapor pressure.

Our usual simulation setup, shown in Figure 3.2, resembles a thick slab of water, parallel to the xy-plane, that is periodically replicated in space. Between the slabs, we place a fixed wall of purely repulsive particles, to prevent the slabs from drifting in the z direction. We always place our probe volumes and other solutes at least 10 Å away from the liquid-vapor interface to ensure that we probe bulk water behavior. shows a typical water slab setup. Below, we show that results derived with this setup are indistinguishable from those obtained in constant-pressure simulations.

3.2.2 Water number distributions in bulk

Using the tools developed in the previous section, we have calculated the water number distribution for a $12\times12\times12$ Å^3 cube in water. Our results, computed in both the constant-volume setup of Figure 3.2 and in a constant-pressure bulk water system, are shown in Figure 3.3. For $N \ll \langle N\rangle_v$, the value of $P_v(N)$ is much higher than would be predicted from Gaussian statistics alone (Equation (2.25)). We can rationalize this behavior by supposing that the most probable mechanism for evacuating $\delta N = \langle N\rangle_v - N$ water molecules from v when δN is large is for a spherical vapor bubble of volume $\delta N/\rho_\ell$ to form inside the probe volume v. The free energy cost of the liquid-vapor interface around this bubble scales as $\gamma(\delta N)^{2/3}$, and we make the crude assumption that geometric factors and other entropic factors contribute a constant amount to the total free energy. This vapor bubble model, shown in red dashes in Figure 3.3, captures the majority of the deviation of $P_v(N)$ from the Gaussian model. A more detailed treatment that accounts for the freedom of the vapor bubble to move inside the volume v, applied to a Lennard-Jones solvent in Ref. [46], rationalizes the downturn near $N = 0$ as an entropic cost to localizing the center of the vapor bubble to the center of v.

At higher pressure, the cost of creating a vapor bubble increases, and the free energy cost of emptying v can, in fact, exceed the cost predicted by the Gaussian model. The water number distributions of the same cube calculated at pressures of up to 10 kbar are

---

[2] To wit, when first computing the $P_v(N)$ distributions in Figure 3.4, we discovered a bug in the implementation of one of the constant pressure MD algorithms in GROMACS.

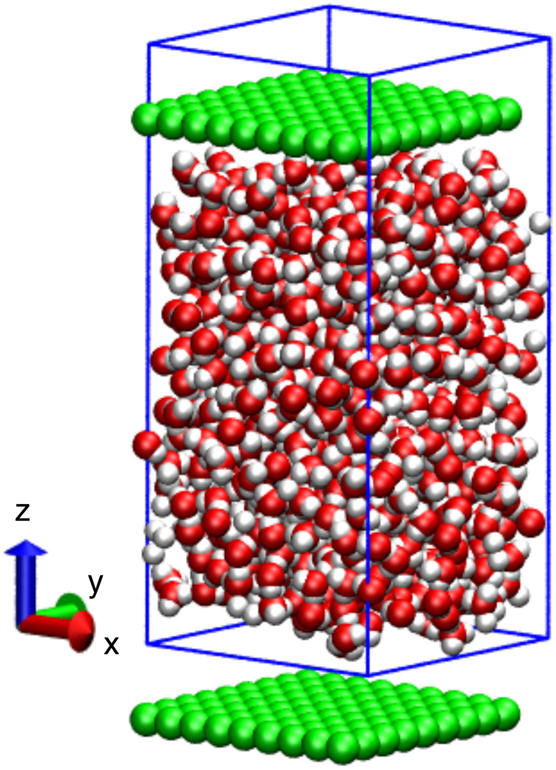


Figure 3.2: Typical water slab setup for calculating $P_v(N)$ distributions at water-vapor coexistence conditions. The blue outline depicts the boundaries of the periodic simulation box. The green particles are purely repulsive hard spheres that keep the intervening water slab (red and white particles) from drifting in the z-direction.

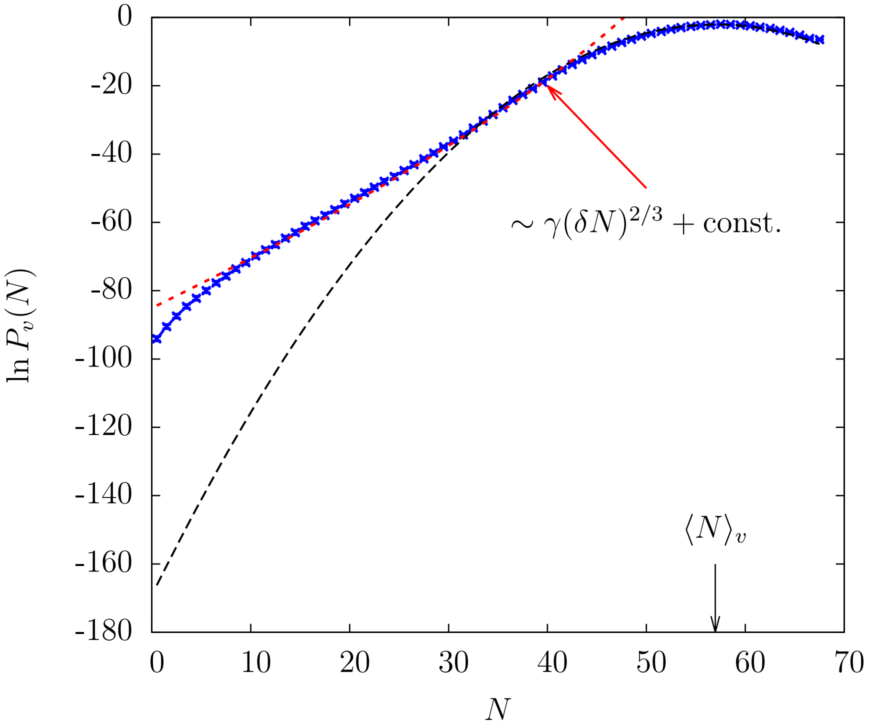


Figure 3.3: Water number distribution for a cubic probe volume of side 12 Å, in bulk water at 298 K. Black: computed in a fixed-volume simulation with a vapor bubble. Blue: computed in a fixed-pressure simulation at 1 bar. Dashes: Gaussian distribution with equal mean and variance. Red dashes: Distribution expected from vapor bubble model (see text).

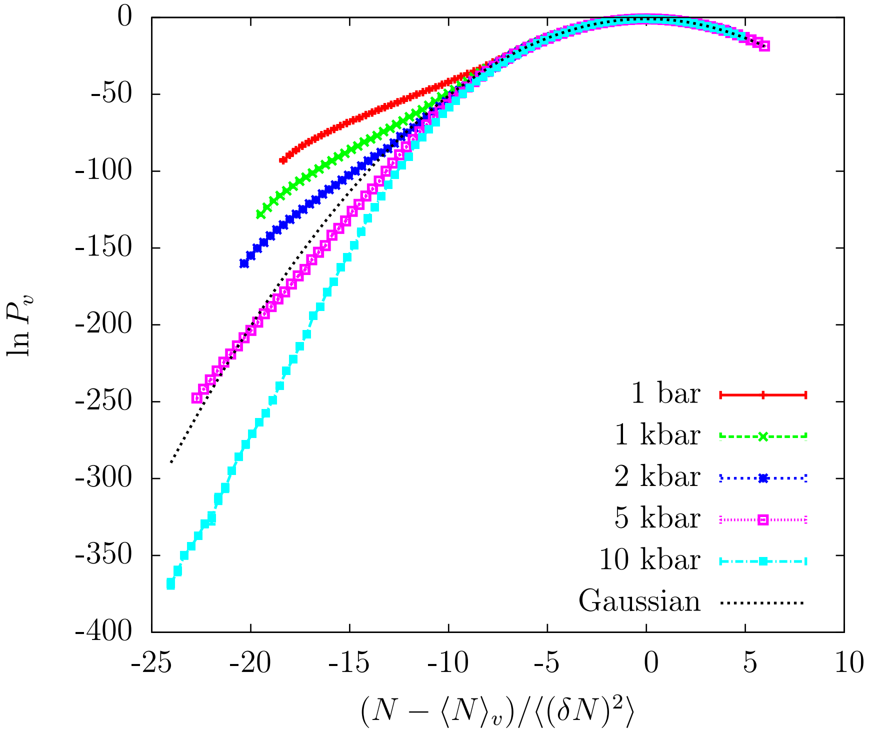


Figure 3.4: $P_v(N)$ distributions for a 12 Å-side cube, relative to the Gaussian model (Equation (2.25)) at various pressures.

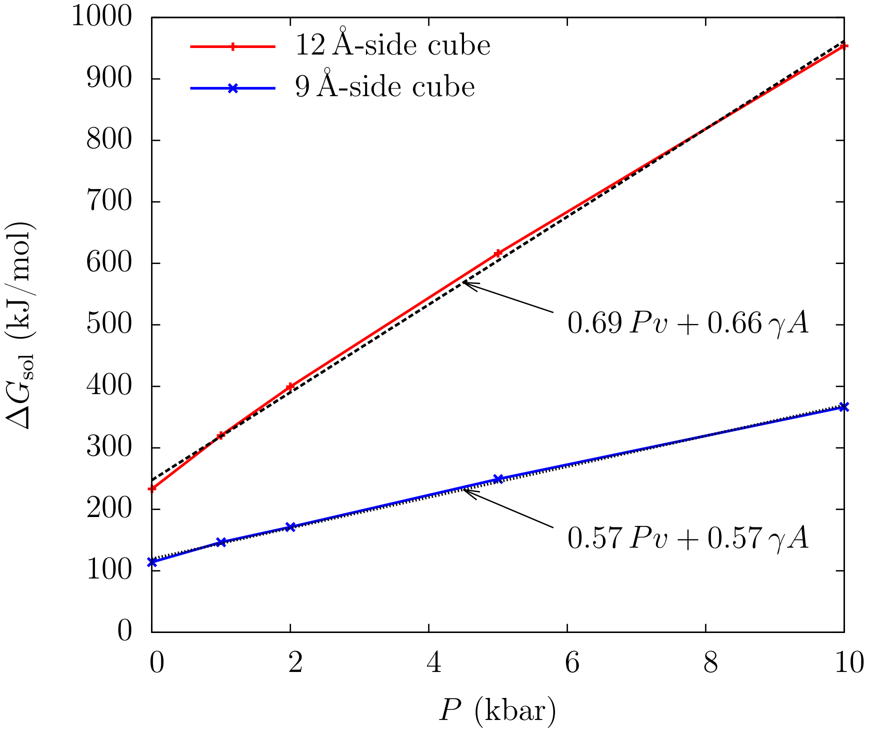


Figure 3.5: Free energy of solvation, $\Delta G_{\mathrm{sol}} = -k_{\mathrm{B}}T \ln P_v(0)$, of 12 Å-side and 9 Å-side cubes as a function of pressure. Dashed lines are fits to Equation (3.41).

shown in Figure 3.4. From the value of $P_v(0)$, we can derive the solvation free energy of the cubes (Equation (2.26)). These are linear in P, as shown in Figure 3.4 and as expected from the macroscopic limit of Equation (2.18). Quantitatively, there are important finite-size corrections. We fit the solvation free energies to the form

$$\Delta G_{\text{sol}} = c_1 P v + c_2 \gamma A, \tag{3.41}$$

and find that $c_1 \approx c_2 \approx 0.7$ for the 12 Å-side cube. This value of c_2 results from the cross-over in the dominant solvation physics from entropic to enthalpic for solutes of size about 1 nm (Figure 2.3). The value of c_1 results from density oscillations around the empty cavity v. That is, when water is expelled from v, the system's volume increases by somewhat less than v because the density of water around v is driven to be higher than the bulk density. The magnitude of both corrections is proportional to the cavity's surface area, so both c_1 and c_2 tend to 1 for macroscopic solutes. In Figure 3.5, we also show the solvation free energies of a 9 Å-side cube under various pressures to illustrate this limiting behavior. To the best of our understanding, approximate equality of c_1 and c_2 in the case of cubic solutes appears to be a geometric coincidence.

3.2.3 Water number distributions near a model plate

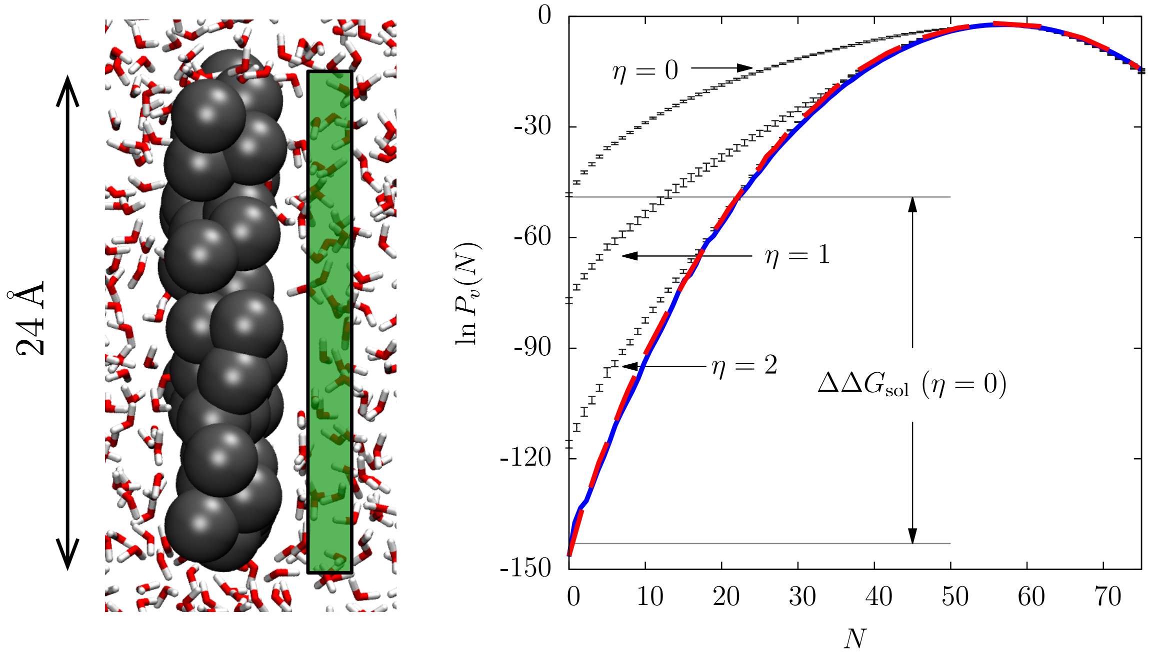
In this section, we build a set of model hydrophobic and hydrophilic solutes that we will use throughout the rest of the thesis. We use them here to explore how water number distributions are modified when the probe volume is next to a solute.

To build the model solutes, we first prepare and equilibrate a bulk water system at 298 K in the presence of a vapor bubble. At an arbitrary time, we record the positions and orientations of all the water molecules whose oxygen atoms are inside a $3 \times 24 \times 24$ Å^3 cuboid. These molecules, frozen and treated as one solute, are our model hydrophilic solute. To build a model hydrophobic solute, we replace all water molecules by WCA particles centered at the original oxygen atom positions. The energy and size parameters of the particles are as those of an OPLS/UA sp^3-hybridized CH_2 group, so that the model solute appears "oily" to the surrounding water.[3] The van der Waals radius of these particles is about 2 Å, and the effective hard-sphere radii of their solvent-excluded volume is 3.37 Å, so the plate's complete solvent-excluding volume is approximately that of a $10 \times 31 \times 31$ Å^3 cuboid. The atomic coordinates of the model hydrophobic solute are recorded in the Supplementary Information of Ref. [111].

To tune the hydrophobicity of the model solute, we scale the WCA attractive tail by a parameter η. When η is 0, the model solute is essentially ideal in the sense of Equation (2.17). As η is increased, a gradually more hydrophilic solute results.

---

[3] The exact parameters are $\epsilon = 0.118$ kcal/mol and $\sigma = 3.905$ Å, taken from Ref. [58]. Lorentz-Berthelot mixing rules are used to obtain the water-solute interaction parameters.



(a) (b)

Figure 3.6: Effect of a large solute on the density fluctuations of the surrounding solvent. (a) The model hydrophobic plate and the probe volume v. (b) Water number distributions for v in bulk (solid blue), next to the hydrophilic plate (dashed red) and next to the hydrophobic plate (dark grey error bars). Error bars for v in bulk and next to the hydrophilic plate are comparable to those for v next to the hydrophobic plate, and are omitted for clarity.

To probe the effect of a solute on the surrounding solvent, we calculate $P_v(N)$ for a probe volume of size $3 \times 24 \times 24\,\text{Å}^3$ adjacent to the model hydrophilic or hydrophobic solute, and compare it to $P_v(N)$ in the absence of the solute. Figure 3.6 shows that the distributions in bulk and next to the hydrophilic solute are essentially indistinguishable, but when the probe volume is placed next to the model hydrophobic solute (with η one of 0, 1 or 2), a fat tail appears. In other words, it is orders of magnitude easier to empty the probe volume when it is next to the hydrophobic solute than when it is next to the hydrophilic solute or in bulk. Physically, the model hydrophobic solute is large enough to disrupt the hydrogen-bonding network of the surrounding water. As such, a soft liquid-vapor interface forms around the solute. To evacuate v, this interface need simply be deformed, which costs much less free energy than to evacuate v in bulk. This physics is modeled in detail in Chapter 5.

Figure 3.6 also illustrates the effective attraction that would arise between the model hydrophobic solutes and an ideal solute with solvent-excluding volume v. The difference in free energy between evacuating v in bulk and next to the model solute, denoted by $\Delta\Delta G_{\mathrm{sol}}$

in the figure, is precisely the free energy gained by bringing these two solutes together in water. In contrast, $P_v(0)$ next to the hydrophilic plate is equal to $P_v(0)$ in bulk, so the v and the model hydrophilic plate experience no effective attraction. Conversely, the hydrophobic attraction between v and the model solute is highest when $\eta = 0$, and decreases as η grows, since evacuating v next to the solute results in a loss of solute-water adhesive energy. As a practical matter, $\Delta\Delta G_{\text{sol}}$ provides a quantitative measure of the model solute's hydrophobicity.

3.2.4 Dewetting transitions: water number distributions between two model plates

In this section, we look at water number distributions in the region between two hydrophobic plates. The basic setup is shown in Figure 3.7. We place two model hydrophobic plates oriented parallel to the xy-plane, with the line joining their centers parallel to $\hat{z}$. Since the plates are about $7\,\text{Å}$ thick, placing one plate's center at $z = 0\,\text{Å}$ and the other at $z = d + 7\,\text{Å}$ creates an interplate space that is $d\,\text{Å}$ thick and has a $24 \times 24\,\text{Å}^2$ cross-section. When $d = 3\,\text{Å}$, a water molecule just fits between the plates. The interplate space, reduced by $3\,\text{Å}$ in thickness to account for the size of a water molecule, serves as our probe volume v.

Figure 3.8 shows the calculated $P_v(N)$ distributions for various values of d and η in the ranges $7 \leq d \leq 11\,\text{Å}$ and $0 \leq \eta \leq 1$. In all cases, there is an evident concavity in these distributions. When $\eta = 1$, the plates are mildly hydrophobic, and at large separations ($d = 11\,\text{Å}$), there is only one stable state of the interplate region, a "wet", high-N state. Nevertheless, it is clear that for sufficiently low N, a second "dry" state can be induced, where a vapor tube forms between the two plates. At lower separations ($d \lesssim 9\,\text{Å}$), this dry state becomes meta-stable, and eventually around $d = 7\,\text{Å}$, it becomes more stable than the wet state. This observation is congruent with the dewetting transitions observed in dynamical simulations of the hydrophobic association of paraffin plates [52]. With our water number distributions, we can glean the existence of the dry state at very large separations where it is not yet metastable. From these observations, it is possible to predict that any small perturbation to the system, such as bringing the two plates together or changing the strength η of the attractive tail, may make the dry state meta-stable or even stable. This expectation is borne out in the $P_v(N)$ distributions calculated at $\eta = 0.5$ and $\eta = 0.0$.

3.2.5 Dewetting in biology: water number distributions between melittin dimers

Having characterized dewetting transitions between model hydrophobic plates, we now explore dewetting between hydrophobic proteins. Our model protein is melittin (PDB: `2MLT`), a 22 amino acid alpha-helical protein that is the main component of bee venom [108, 109]. Its key structural feature, shown in Figure 3.9, is a long string of hydrophobic amino acids at

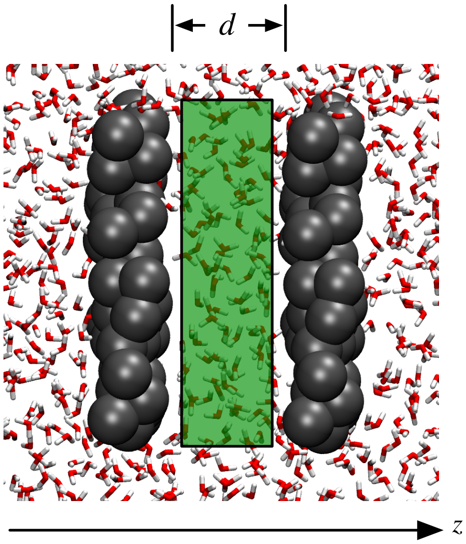


Figure 3.7: Setup for examining water fluctuations under confinement (here, $d = 8\,\text{Å}$). The model hydrophobic plates (grey particles) are placed a distance d apart: taking into account the van der Waals radii of about 2 Å of the plates' oily particles and the 3 Å thickness of each plate, the center of the first plate is placed at $z = 0\,\text{Å}$, the center of the second plate is placed at $z = d + 7\,\text{Å}$. The van der Waals radius of water (red and white sticks) being about 1.5 Å, the $24 \times 24 \times (d-3)\,\text{Å}^3$ probe volume (green) extends from $z = 5\,\text{Å}$ to $z = d + 2\,\text{Å}$. The plates are not perfectly flat, so some waters fit between the plates and the probe volume.

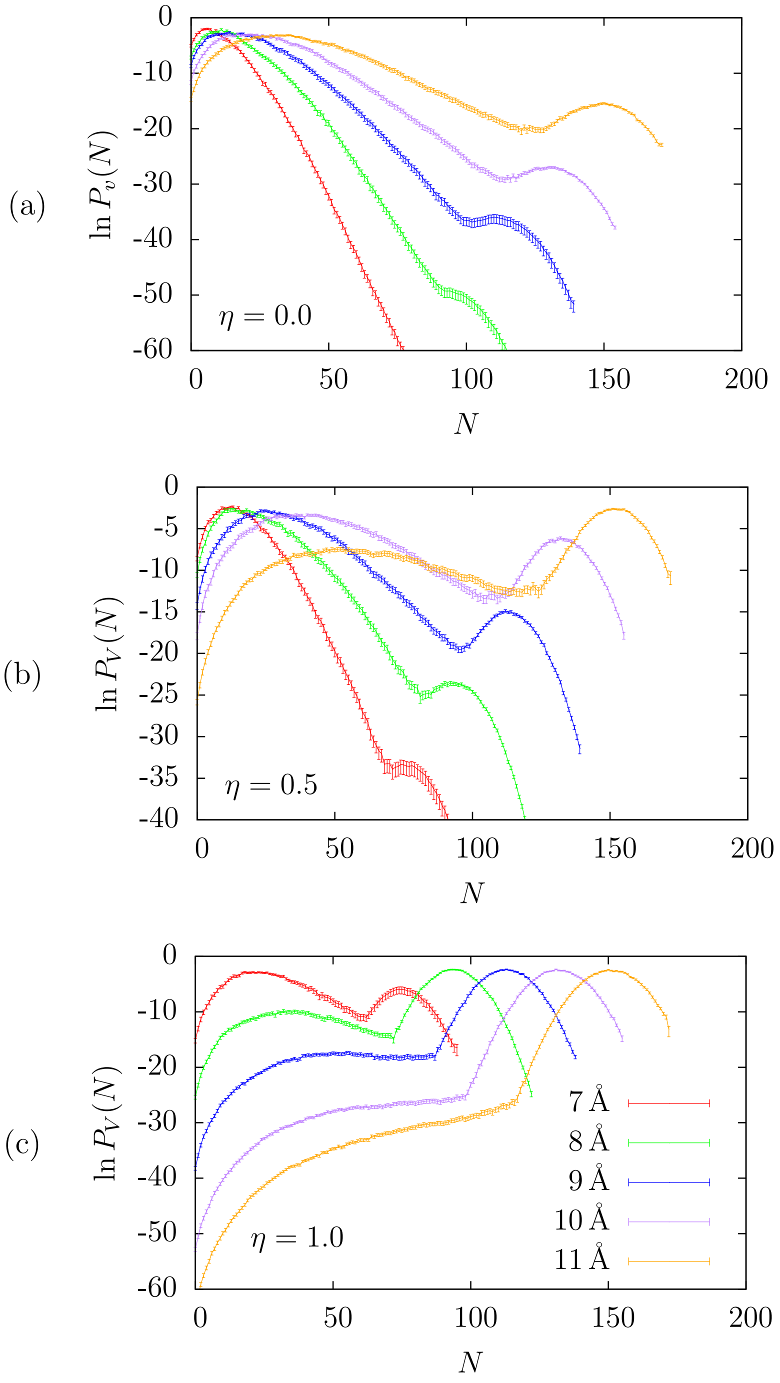


Figure 3.8: $P_v(N)$ distributions for interplate region in Figure 3.7 for η equal to (a) 0.0, (b) 0.5 and (c) 1.0. The plate separation runs from $d = 7\,\text{Å}$ (red) to $d = 11\,\text{Å}$ (orange).

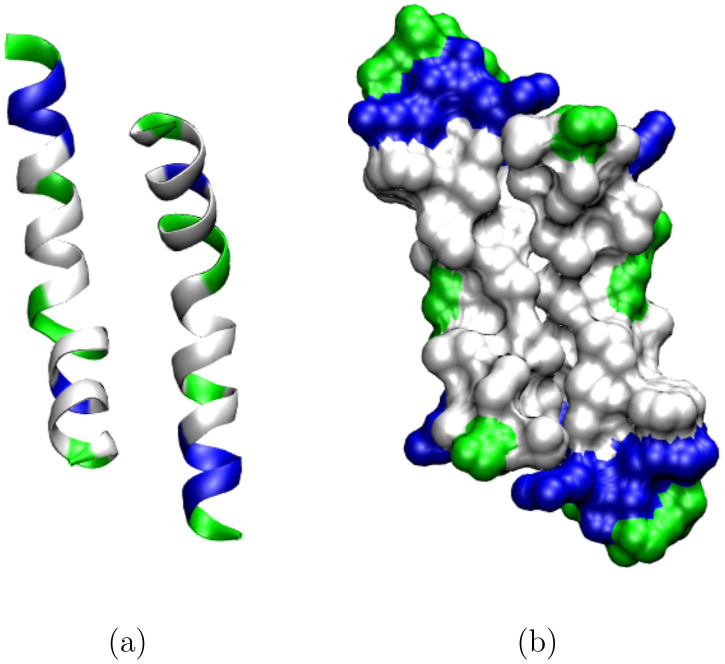
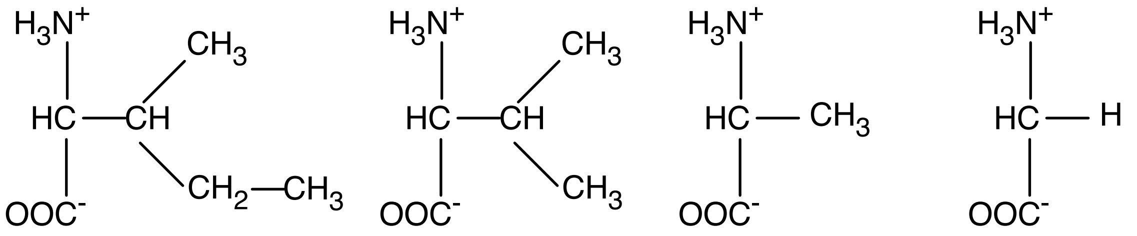


Figure 3.9: Crystal structure of melittin dimer (PDB: `2MLT`), looking down the face that mediates tetramerization in (a) cartoon and (b) solvent-accessible surface representation. Residues are colored according to whether they are non-polar (white), polar but uncharged (green) or positively charged (blue). Figures prepared with VMD. VMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign [55].

its center (mainly isoleucine, leucine, valines and alanines), capped by charged, hydrophilic amino acids (lysines and arginines). At physiological conditions, melittin remains monomeric and unfolded in solution, but at other conditions, like high ionic concentrations, it assembles into tetramers [21]. The tetramers are formed from two symmetry-related dimers, whose interaction is mediated by a large hydrophobic patch on each dimer. In vivo, melittin acts monomerically by attacking membranes and helping to create pores in them. The pores are only weakly structured, having about 4 to 7 melittin monomers each [50].

Our interest in melittin is primarily as a model hydrophobic protein, so we focus on its hydrophobic assembly in conditions where the tetramer is stable. We extend the study of Liu et al [64], which we summarize briefly now. In that study, the melittin dimers present in the crystal structure were separated by a distance d, ranging from 4 Å to 8 Å, and solvated in water at 298 K and 1 atm. The system was then evolved for up to 10 ns. When the separation was below about 7 Å, the two initially separated melittin dimers collapsed within



(a) Isoleucine (Ile) (b) Valine (Val) (c) Alanine (Ala) (d) Glycine (Gly)

Figure 3.10: Hydrophobic amino acids pertinent to the study of melittin mutations. Ile is the most hydrophobic, while Gly is the least.

less than 1 ns. On careful analysis, it was seen that in many trajectories, before the collapse occurs, most of the water between the melittin dimers is evacuated. Mutations affected this dewetting behavior. Specifically, the authors looked at point mutations of the isoleucine residues in the protein (Ile2, Ile17 and Ile20), which are very hydrophobic, into smaller, less hydrophobic residues: valine, alanine and glycine. The structures of these four amino acids are shown in Figure 3.10, and the location of the isoleucine residues in the melittin dimer is highlighted in Figure 3.11. Liu et al. found that any mutation to the Ile2 residue suppressed dewetting completely. They interpreted this finding as evidence that the Ile2 residue provides a unique topographical feature on the surface of the protein that is essential for dewetting to occur. Additionally, mutations to Ile20 slowed down the evacuation of the inter-melittin cavity, with the Ile20Gly mutation in particular resulting in large fluctuations in the number of waters in the cavity.

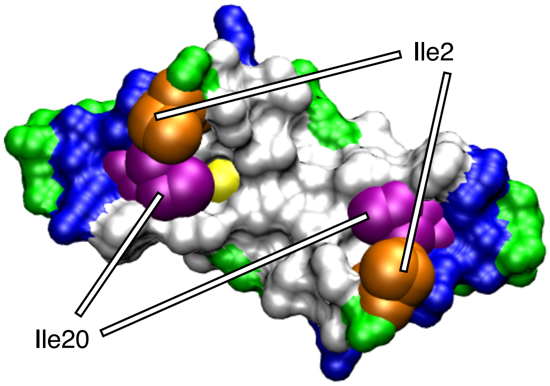
Here, we study dewetting in melittin in the language of water number distributions. In all our studies, we focus on the behavior of water, not the protein, so we have anchored the positions of all the atoms in the protein backbone. Doing so allows us to study the thermodynamics of dewetting at a fixed separation, instead of observing a dynamical collapse as reported in Ref. [64].

We begin by making a concrete definition of the inter-melittin cavity. The hydrophobic core of each monomer roughly spans residues 8 through 20. We have thus found it effective to define the inter-melittin cavity as the convex hull[4] of the α-carbons of residues 8 and 20 of each monomer (8 points in total), which we calculate using the algorithm of Ref. [20]. To focus on the behavior of water and not the protein side chains, we then remove all points from this convex hull that are within a van der Waals radius of a non-hydrogenic protein atom. Denote this volume by v'. This volume is overlaid on a 1 Å-resolution grid, and the final volume v is taken to be the union of all the grid cells whose center lies inside v'.

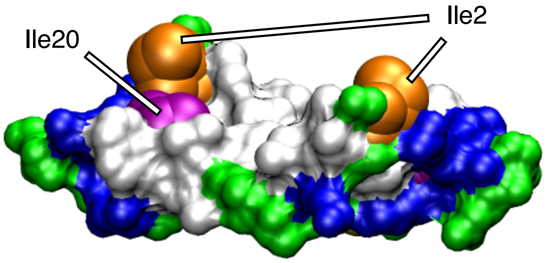
We first quantify the hydrophobicity of the melittin dimer's top face. Figure 3.12(a) shows

---

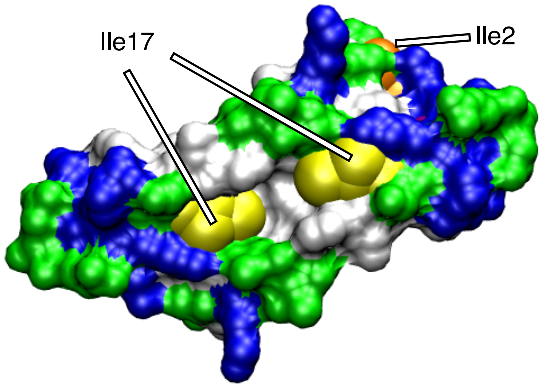
[4] The convex hull of a set of points is the smallest convex volume that contains them. It has roughly the shape of a membrane stretched tautly over these points.



(a) Top view (looking down at hydrophobic face)



(b) Side view



(c) Bottom view

Figure 3.11: Isoleucine residues in melittin dimer, mutated here and in Ref. [64] to less hydrophobic residues. Dimer shown in solvent-accessible surface representation, colored as in Figure 3.9. Highlighted residues are Ile2 (orange), Ile17 (yellow) and Ile20 (purple).

the simulation setup used. We use a probe volume defined by the convex hull procedure above, starting from the tetramer configuration at 4 Å separation, but where we remove the van der Waals interior only of the non-hydrogenic atoms of the unmutated bottom dimer. Figure 3.12(b) shows the calculated water number distributions for this volume v in bulk and next to 10 melittin mutants. As with the model hydrophobic plate system of Section 3.2.3, these distributions have a much fatter tail next to the melittin dimer than in bulk. This demonstrates that there is a liquid-vapor-like interface that spans the hydrophobic patch of the melittin dimer, which can be easily deformed to evacuate a probe volume next to it. We can also discern that the hydrophobicity of the different mutants, quantified by $\Delta\Delta G_{\text{sol}}$, is similar. The different mutations, as expected, tend to render the hydrophobic patch slightly less hydrophobic.

We now look at $P_v(N)$ distributions for a cavity between two melittin dimers held at various separations. The distributions for the unmutated protein are shown in Figure 3.13. These distributions qualitatively resemble those in the region between two model plates, shown in Figure 3.8, even though the melittin dimers are topographically heterogeneous and contain charges and polar regions. As with the model plate system, the $P_v(N)$ distribution for the largest separation in melittin (9 Å) already contains a large plateau region at low N, corresponding to a "dry" state. As the separation between dimers is reduced to 8 Å, this dry state becomes metastable, and around a separation between 6.5 Å and 7.0 Å, it becomes stable with respect to the wet state. Between the wet and dry state, there is a barrier that is several $k_{\text{B}}T$ high, which suggests that the states will be long-lived, so that the collapse dynamics of the protein will be dominated by solvent motion. These distributions provide a thermodynamics rationalization for the observations in the dynamical collapse trajectories of Ref. [64].

Calculating $P_v(N)$ helps us better understand the effects of mutations. Already in Figure 3.12, it is evident that mutations to the Ile residues of melittin make only minor changes to the hydrophobic character of its binding face. On the other hand, because the water number distributions in Figure 3.13 are close to being bistable, small changes to these distributions can result in large apparent changes in the mean behavior of the water between the melittin dimers. This sensitivity is analogous to the sensitivity of a magnet's behavior to small changes in the magnetic field when the state point is close to the phase boundary between the magnetized and unmagnetized phases.

Figures 3.14, 3.15 and 3.16 show the water number distributions in the cavity between two dimers mutated at the Ile2, Ile17 and Ile20 residues, respectively. Qualitatively, all these distributions are quite similar, with only small shifts in the relative stabilities of the dry and wet basins with respect to the behavior of the unmutated dimer. However, these small shifts result in large changes in the mean number of waters between the melittin cavities, which imply large changes in the dynamics of tetramerization. The changes that we observe, such as lack of dewetting for the Ile2Ala and Ile2Gly mutants and the flat low-N plateau in the Ile20Gly mutant, are consistent with those noted by Ref. [64], but our distributions are more informative. Our distributions clearly emphasize the delicate balance that exists between

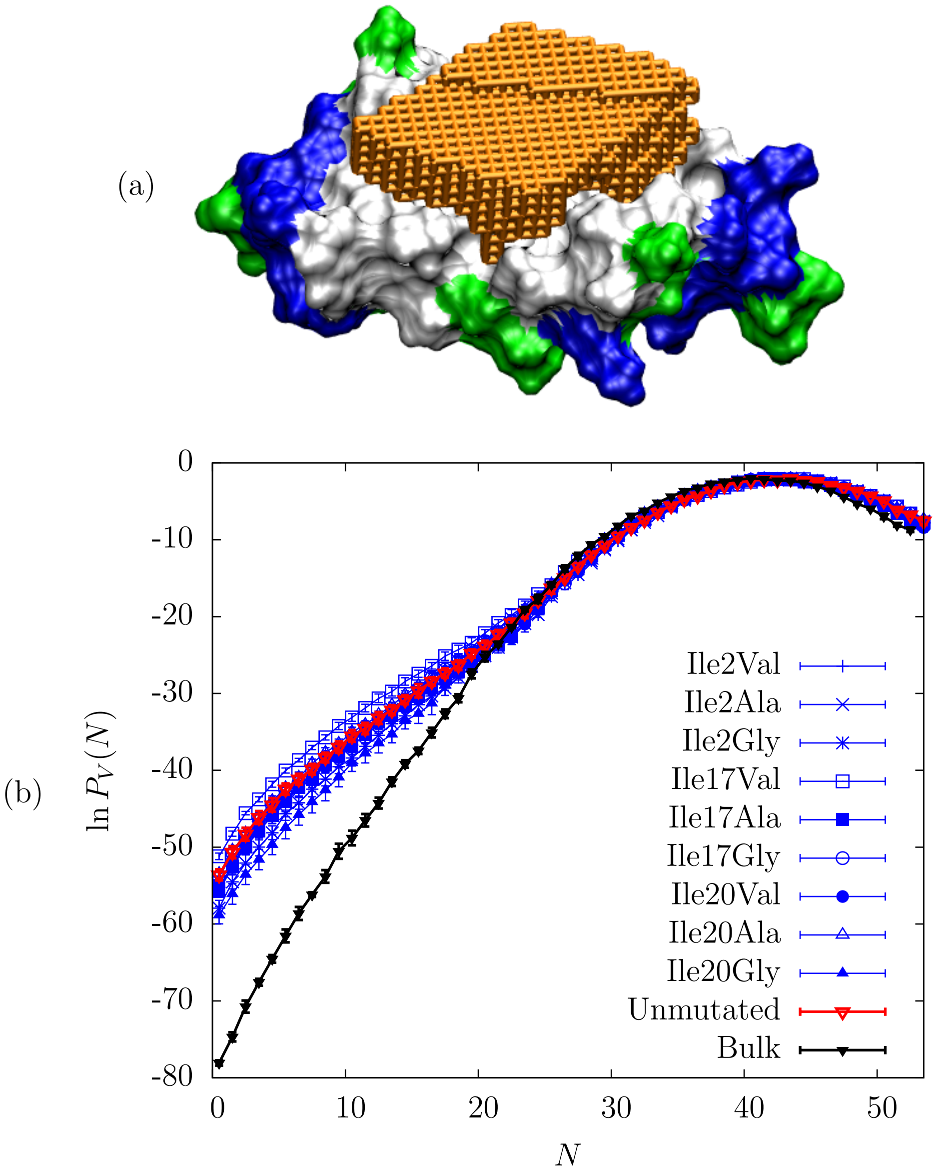


Figure 3.12: Measuring the hydrophobicity of the melittin dimer's hydrophobic face. (a) Probe volume v (orange) next to unmutated dimer (colors as in Figure 3.9). (b) $P_v(N)$ distributions of v in bulk and next to 10 melittin mutants.

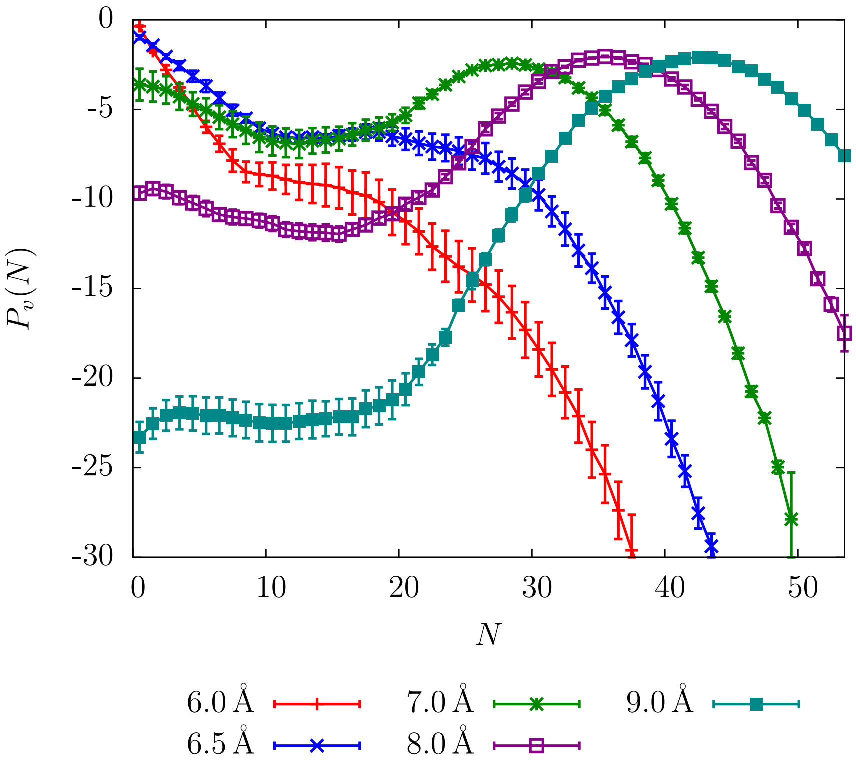


Figure 3.13: $P_v(N)$ distributions for cavity between two unmutated melittin dimers held at different separations.

the wet and dry states. Whether or not dewetting is present under particular conditions is seen to be a very delicate question, one that might even have a different answer in nature than in simulations with a particular molecular field. However, the character of these water number distributions is quite robust to small changes in the system, so we can confidently predict that any dewetting in experiments would be very sensitive to external conditions.

3.2.6 Probing cavities of general shape

The method described in Section 3.2.1 for computing $\tilde{N}(\mathbf{R}^N)$ is readily generalized to shapes beyond the union of non-overlapping cubes. Here, we describe three simple strategies in detail. All of them focus on constructing smooth indicator functions $\tilde{h}(\mathbf{r})$ for general-shaped probe volumes.

The first strategy, analogous to constructive solid geometry in computer aided design, starts by building the probe volume v by applying union, intersection and complement operators to simpler volumes. For a given volume A, we denote by $h_A(\mathbf{r})$ the indicator function and by $\tilde{h}_A(\mathbf{r})$ its smooth counterpart. When A is a cube, $h_A(\mathbf{r})$ is given by Equation (3.24) and $\tilde{h}_A(\mathbf{r})$ by Equation (3.28). The indicator function for the complement of A, denoted A', is given by

$$h_{(A')}(\mathbf{r}) = 1 - h_A(\mathbf{r}). \tag{3.42}$$

Similarly, the indicator functions for the union and intersection of two volumes, A and B, are given by

$$h_{(A\cap B)}(\mathbf{r}) = h_A(\mathbf{r})h_B(\mathbf{r}), \tag{3.43}$$

$$h_{(A\cup B)}(\mathbf{r}) = h_A(\mathbf{r}) + h_B(\mathbf{r}) - h_A(\mathbf{r})h_B(\mathbf{r}). \tag{3.44}$$

These definitions immediately suggest the following definitions for the corresponding smooth indicator functions:

$$\tilde{h}_{(A')}(\mathbf{r}) = 1 - \tilde{h}_A(\mathbf{r}), \tag{3.45}$$

$$\tilde{h}_{(A\cap B)}(\mathbf{r}) = \tilde{h}_A(\mathbf{r})\tilde{h}_B(\mathbf{r}), \tag{3.46}$$

$$\tilde{h}_{(A\cup B)}(\mathbf{r}) = \tilde{h}_A(\mathbf{r}) + \tilde{h}_B(\mathbf{r}) - \tilde{h}_A(\mathbf{r})\tilde{h}_B(\mathbf{r}). \tag{3.47}$$

These compound smoothed indicator functions can be used to define $\tilde{N}(\mathbf{R}^N)$ for the compound volume as in Equation (3.29). Practically, a union v of non-overlapping volumes $\{v_1, \ldots, v_n\}$ is an important special case, for which the smoothed indicator function takes the simple form

$$\tilde{h}_v(\mathbf{r}) = \sum_{i=1}^{n} \tilde{h}_{v_i}(\mathbf{r}). \tag{3.48}$$

A second strategy is to find volumes for which integrals of the form in Equation (3.32) can be calculated analytically. Shapes with spherical and cylindrical symmetry fall into this

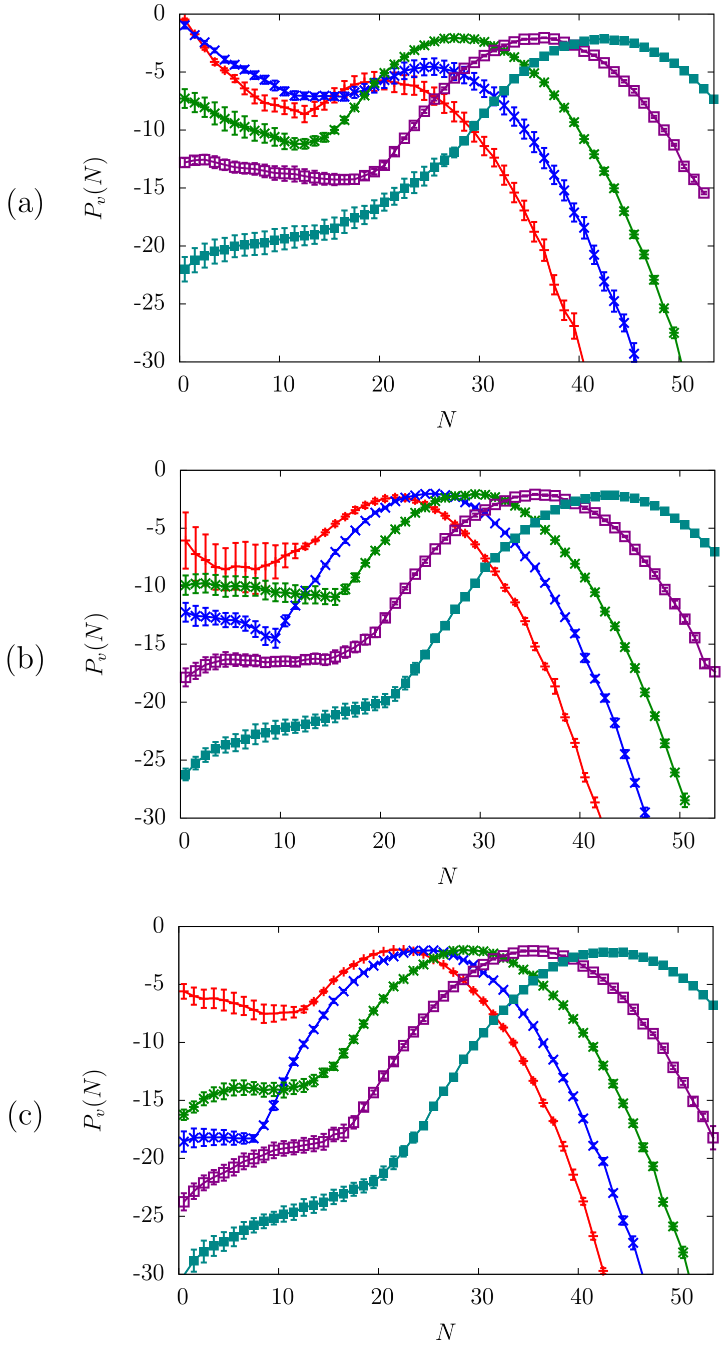


Figure 3.14: Dewetting behavior of the Ile2 series of melittin mutants. Curves as in Figure 3.13. Mutants are: (a) Ile2Val, (b) Ile2Ala and (c) Ile2Gly.

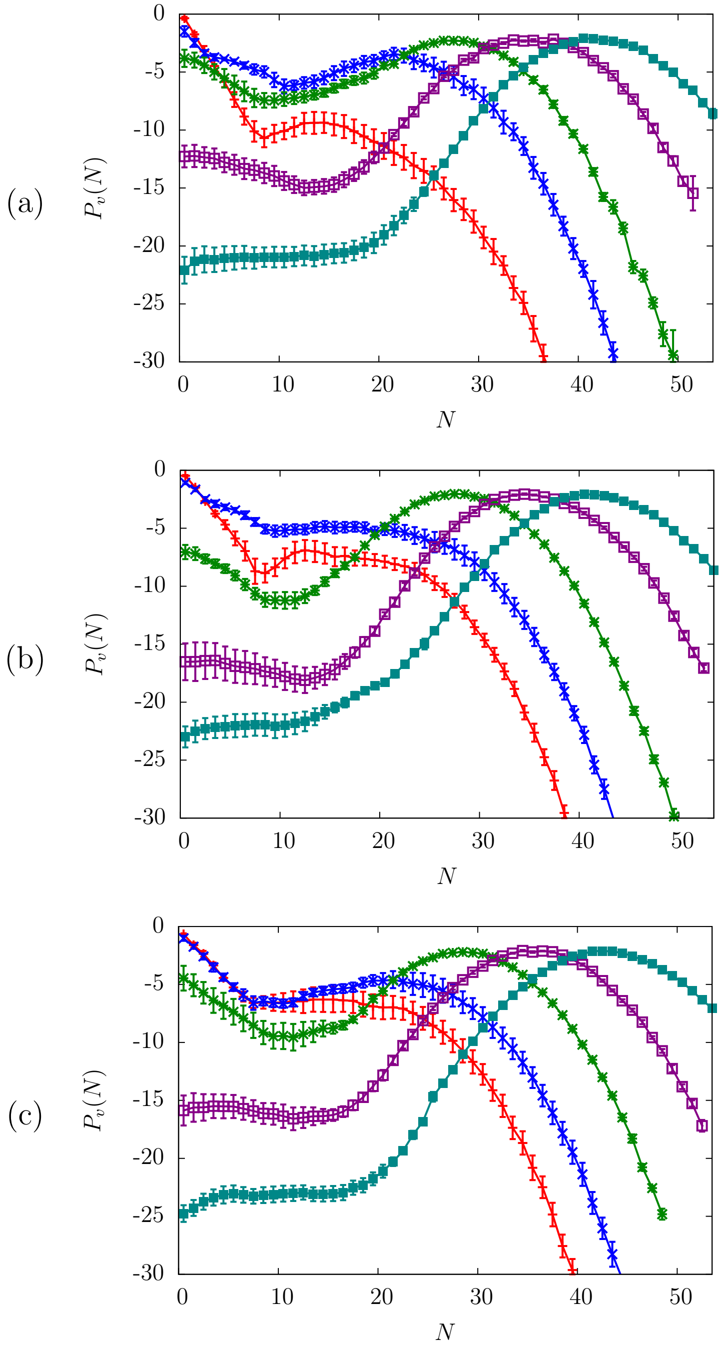


Figure 3.15: Dewetting behavior of the Ile17 series of melittin mutants. Curves as in Figure 3.13. Mutants are: (a) Ile17Val, (b) Ile17Ala and (c) Ile17Gly.

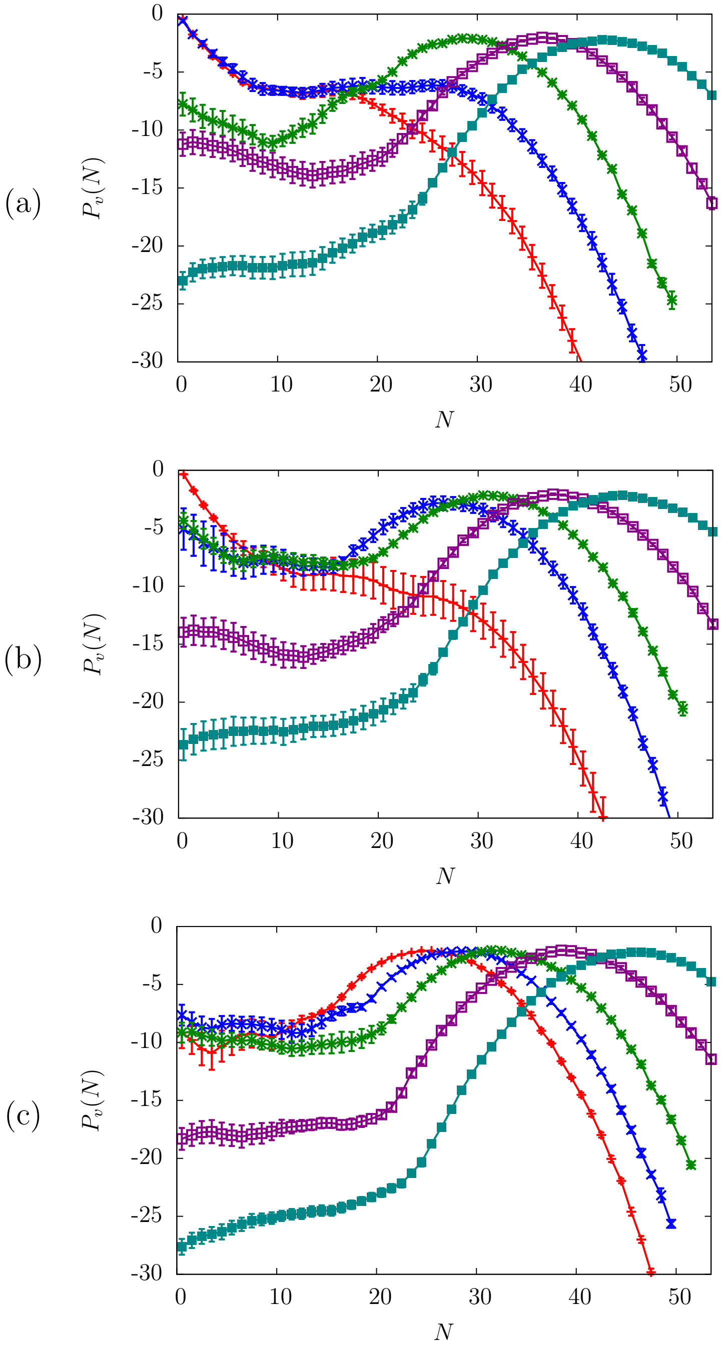


Figure 3.16: Dewetting behavior of the Ile20 series of melittin mutants. Curves as in Figure 3.13. Mutants are: (a) Ile20Val, (b) Ile20Ala and (c) Ile20Gly.

category. A probe volume can be specified by three ranges of orthogonal coordinates in a non-Cartesian coordinate system. For example, we can define a hemispherical shell in this way. If the three coordinates are denoted by α, β and γ, then the smooth indicator function for the full volume is given by

$$\tilde{h}(\mathbf{r}) = \tilde{h}_\alpha(\alpha)\tilde{h}_\beta(\beta)\tilde{h}_\gamma(\gamma). \tag{3.49}$$

In pursuing this generalization, it is useful to construct integrals of $\phi(\alpha)$ that are open at one or both ends. Explicitly, when the integral is open at its lower end, the smooth indicator function is given by

$$\begin{aligned}
\tilde{h}_\alpha(\alpha_i) &= \int_{-\infty}^{\alpha_{\max}} \mathrm{d}\alpha\, \phi(\alpha - \alpha_i), && (3.50)\\
&= \left[k_1 \operatorname{erf}\left(\frac{\alpha_{\max} - \alpha_i}{\sqrt{2\sigma^2}}\right) - k_2(\alpha_{\max} - \alpha_i) - \frac{1}{2}\right] \Theta(\alpha_{\mathrm{c}} - |\alpha_{\max} - \alpha_i|) \\
&\quad + \Theta(\alpha_{\mathrm{c}} + \alpha_{\max} - \alpha_i), && (3.51)\\
\tilde{h}'_\alpha(\alpha_i) &= -\phi(\alpha_{\max} - \alpha_i), && (3.52)
\end{aligned}$$

where k_1 is as in Equation (3.33). Similarly, when the integral is open at both sides, the expressions for $\tilde{h}_\alpha(\alpha_i)$ and its derivative are trivially given by

$$\begin{aligned}
\tilde{h}_\alpha(\alpha_i) &= \int_{-\infty}^{\infty} \mathrm{d}\alpha\, \phi(\alpha - \alpha_i) = 1, && (3.53)\\
\tilde{h}'_\alpha(\alpha_i) &= 0. && (3.54)
\end{aligned}$$

A third strategy is to define $\tilde{h}_v(\mathbf{r})$ directly instead of as an integral of $\Phi(\mathbf{r})$. To be useful, this function must have finite support, be approximately normalized, i.e. $\int_{\mathbf{r}} \tilde{h}_v(\mathbf{r}) \approx v$, closely resemble the exact indicator function, $h(\mathbf{r})$, and its value and gradient must be easy to calculate. For example, a smooth indicator function for a sphere of radius R centered at $\mathbf{r}_0$ is

$$\tilde{h}(\mathbf{r}) = \begin{cases} k_A \left[1 - \tanh\left(\frac{|\mathbf{r}-\mathbf{r}_0|-R}{\delta}\right) - k_B\right], & |\mathbf{r} - \mathbf{r}_0| < R + R_c, \\ 0, & \text{otherwise}, \end{cases} \tag{3.55}$$

where δ defines a smearing length for blurring the edge of the probe volume and R_c is a cutoff distance for ensuring that $\tilde{h}(\mathbf{r})$ has finite support. These parameters are analogous to σ and α_{c} in Equation (3.31). The other constants are $k_A = 1/\{\tanh[(R_c - R)/\delta] - \tanh(-R/\delta)\}$ and $k_B = 1 - \tanh[(R_c - R)/\delta]$. The gradient of $\tilde{h}(\mathbf{r})$ is then

$$\nabla\tilde{h}(\mathbf{r}) = \begin{cases} -\frac{k_A}{2\delta} \operatorname{sech}^2\left(\frac{|\mathbf{r}-\mathbf{r}_0|-R}{\delta}\right) \frac{\mathbf{r}-\mathbf{r}_0}{|\mathbf{r}-\mathbf{r}_0|}, & |\mathbf{r} - \mathbf{r}_0| < R + R_c, \\ \mathbf{0}. & \text{otherwise} \end{cases} \tag{3.56}$$

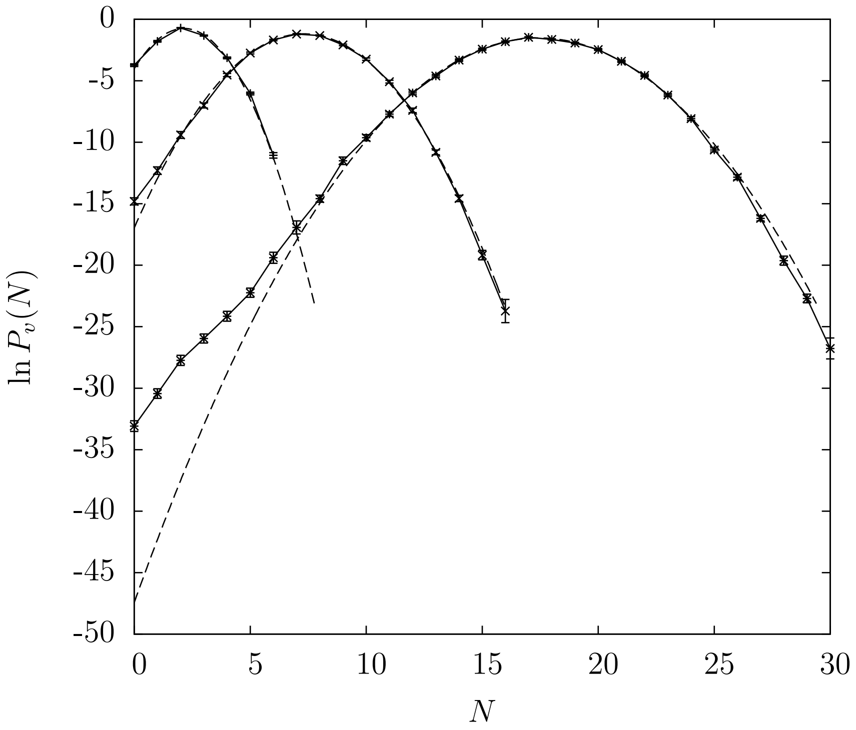


Figure 3.17: $P_v(N)$ for spheres of radius (left-to-right) 2.5 Å, 3.75 Å and 5.0 Å, in bulk water. Gaussian approximation shown in dashes.

To illustrate these generalizations, we have computed $P_v(N)$ for spheres in bulk, shown in Figure 3.17, using Equation (3.55). This data augments that shown in Figure 2.2(a) with information at low N. The distributions evidence fat tails already at radii of 5 Å, which coincides with the nanometric length scale crossover discussed in Section 2.2.1.

3.3 Probing density fluctuations in Fourier space

A fundamental assumption underlying many liquid-state theories is that small-length-scale density fluctuations are Gaussian [11], as discussed in Section 2.2.1. Traditionally, this assumption has been tested by computing water number distributions in small volumes. A more direct way to test this assumption is to calculate distributions for the Fourier components of the density field,

$$\hat{\rho}(\mathbf{k}) = \frac{1}{\sqrt{N}} \int_{\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}} \rho(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e^{-i\mathbf{k}\cdot\mathbf{R}_i}. \tag{3.57}$$

The normalization of the transform is chosen so that $\hat{\rho}(\mathbf{k})$ has a finite variance in the thermodynamic limit. If density fluctuations were perfectly Gaussian, in the sense of Equation (2.32), then $\hat{\rho}(\mathbf{k})$ should also exhibit Gaussian statistics. Moreover the values for different wavevectors $\mathbf{k}$ should be independent. That is, for any two nonzero wavevectors $\mathbf{k}_1$ and $\mathbf{k}_2$, we would have a joint distribution for the values of $\hat{\rho}(\mathbf{k}_1)$ and $\hat{\rho}(\mathbf{k}_2)$ given by

$$P[\hat{\rho}(\mathbf{k}_1), \hat{\rho}(\mathbf{k}_2)] \propto \exp\left[-|\hat{\rho}(\mathbf{k}_1)^2|/2S(\mathbf{k}_1) - |\hat{\rho}(\mathbf{k}_2)^2|/2S(\mathbf{k}_2)\right], \tag{3.58}$$

where $S(\mathbf{k}) = \langle \rho(-\mathbf{k})\rho(\mathbf{k})\rangle$ is the static structure factor of the liquid measured by scattering experiments [41].

To permit testing such assumptions, we have implemented umbrella sampling with respect to $\hat{\rho}(\mathbf{k})$ in LAMMPS. To simplify the problem, we have chosen to umbrella sample only its magnitude, $|\hat{\rho}(\mathbf{k})|$. By translational symmetry, its phase should be uniformly distributed, but the phase of $\hat{\rho}(\mathbf{k})$ at different wavevectors may, in principle, be correlated. We do not address these issues here. Were $\hat{\rho}(\mathbf{k})$ Gaussian, the probability distribution for its magnitude would resemble a Maxwell-Boltzmann distribution, i.e.

$$P[|\hat{\rho}(\mathbf{k})|] \propto |\hat{\rho}(\mathbf{k})|^2 \exp[-|\hat{\rho}(\mathbf{k})^2|/2S(\mathbf{k})]. \tag{3.59}$$

When simulations are performed in a finite, periodic box, the wavevector $\mathbf{k}$ must be of the form $\mathbf{k} = (2\pi n_x/L_x, 2\pi n_y/L_y, 2\pi n_z/L_z)$.

The umbrella potentials we have implemented are of the form

$$V_n(\mathbf{R}^N) = \frac{\kappa}{2}(|\hat{\rho}(\mathbf{k})| - \hat{\rho}_0^{(n)})^2, \tag{3.60}$$

where κ is an umbrella strength parameter, $\hat{\rho}_0^{(n)}$ defines the center of the n^{th} umbrella, and $\hat{\rho}(\mathbf{k})$ is defined by Equation (3.57). The corresponding forces have the form

$$-\nabla_i V_n(\mathbf{R}^N) = -\kappa(|\hat{\rho}(\mathbf{k})| - \hat{\rho}_0^{(n)})\frac{\Re[\hat{\rho}^*(\mathbf{k}) \cdot (-i\mathbf{k}) \exp(-i\mathbf{k} \cdot \mathbf{R}_i)]}{\sqrt{N}|\hat{\rho}(\mathbf{k})|}. \tag{3.61}$$

The computation of the instantaneous value of $\hat{\rho}(\mathbf{k})$ can be efficiently parallelized in a manner exactly analogous to $\tilde{N}(\mathbf{R}^N)$, described in Section 3.2.1, as can the computation of the umbrella forces. Our code thus allows efficient calculation of the probability distribution $P[|\hat{\rho}(\mathbf{k})|]$.

Figure 3.18 shows an example distribution computed in this way.[5] The system contains 342 molecules of the TIP4P/2005 model of water in a periodic cubic box, simulated at 298 K and at fixed density 1 g/ml. The wavevector $\mathbf{k}$ is parallel to the x-axis and its magnitude is chosen to approximately coincide with the maximum of $S(\mathbf{k})$ for this water model.

---

[5] The author would like to thank Ulf Pedersen for allowing this data from his study of density fluctuations in Fourier space to be included in this thesis, to illustrate the use of the $\hat{\rho}(\mathbf{k})$ umbrella potential code.

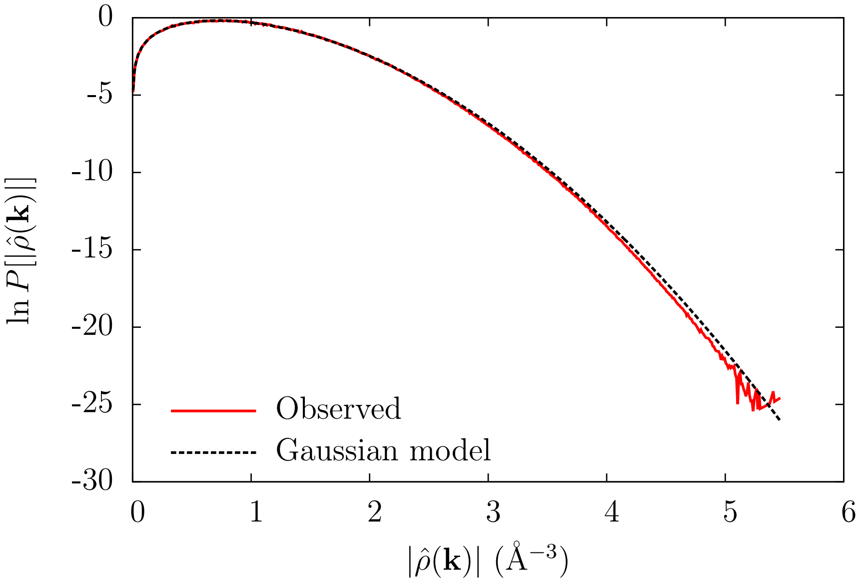


Figure 3.18: $P[|\hat{\rho}(\mathbf{k})|]$ distribution for TIP4P/2005 water, with $|\mathbf{k}| = 2.89\,\text{Å}^{-1}$. The dashed curve is the expected result when density fluctuations are Gaussian (Equation (3.59)).

3.4 Probing interfacial fluctuations

In this final section, we describe tools to characterize rare fluctuations relating to liquid-vapor interfaces. At a molecular level, these interfaces are not unambiguously defined, so in order to proceed, we need to pick one concrete definition of such an interface. We take advantage the recent work of Willard and Chandler [125], who have shown that a molecular-scale interface can be reasonably defined as the isosurface

$$\tilde{\rho}(\mathbf{r}) = (1/2)\rho_\ell \tag{3.62}$$

of a smooth density field, $\tilde{\rho}(\mathbf{r})$, given by the convolution of the exact oxygen density field, $\rho(\mathbf{r})$, with a normalized Gaussian of molecular width. That is,

$$\tilde{\rho}(\mathbf{r}) = \sum_{i \in O} \phi(\mathbf{r} - \mathbf{R}_i), \tag{3.63}$$

where

$$\phi(\mathbf{r}) = (2\pi\xi^2)^{-3/2} \exp(-r^2/2\xi^2). \tag{3.64}$$

Here, ξ is a molecular length scale. In practice, the smearing function is truncated at $|\mathbf{r}| = 3\xi$ to faciliate calculations. This change affects most properties only negligibly, but has important ramifications with respect to the curvature of the interface. We revisit this subtlety in Section 6.2.4.

Thorughout, we refer to the interface defined above as a "Willard interface". Willard and Chandler found that $\xi = 2.4\,\text{Å}$ is the smallest length scale for which the field $\tilde{\rho}(\mathbf{r})$ of a bulk configuration of water has few, if any, voids. Figure 3.19 illustrates this interface, calculated for an instantaneous configuration of SPC/E water from a water slab simulation.

The height fluctuations of the Willard interface are consistent with those of capillary-wave theory [8], which models an interface along the xy-plane as a surface $z = h(x, y)$ with the Hamiltonian $H[h(x, y)] = \int \gamma\, \mathrm{d}A = \int \gamma \sqrt{1 + |\nabla h(x, y)|^2}\, \mathrm{d}x\, \mathrm{d}y$. The Fourier components of such a membrane, $\tilde{h}(\mathbf{k})$, are approximately independent harmonic oscillators, with

$$\left\langle |\tilde{h}(\mathbf{k})|^2 \right\rangle \sim 1/\beta\gamma k^2. \tag{3.65}$$

The exact proportionality constant depends on the Fourier transform convention. Figure 3.20 shows the measured Fourier component magnitudes of the liquid-vapor interfaces at the top and bottom of a $24 \times 24 \times 3\,\text{nm}^3$ slab of SPC/E water at 298 K. The results agree with capillary-wave theory for wavevectors below about $2\pi/9\,\text{Å}$, and are independent of the ξ smearing length in that regime. In other words, liquid-vapor interfaces behave macroscopically already at length scales of about 1 nm. At smaller length scales, the measured height fluctuations are sensitive to the precise definition of the interface and to molecular detail. At this high resolution, where the contours of individual water molecules start becoming apparent, it is questionable whether any definition of a liquid-vapor interface can be physically sensible.

The Willard interface provides a convenient reference surface to characterize interfacial structure. When atomic positions are measured relative to the closest point on the Willard interface, capillary-wave fluctuations are removed, and effects such as density layering and orientational preferences near the interface are clearly revealed [125]. We have thus developed an umbrella potential to bias solute positions with respect to the position of the Willard interface immediately above or below this solute. In Chapter 6, we use this umbrella potential to map out the free energy of extracting a single water molecule from the bulk.

The umbrella potential we have chosen to implement is of the form

$$V_k(\mathbf{R}^N) = \frac{\kappa}{2}\big[z_n - h(x_n, y_n; \mathbf{R}^N) - a_k\big]^2. \tag{3.66}$$

Here, n is the index of a particle whose position we are biasing, and a_k is its target distance from the Willard interface. The value of $h(x_n, y_n; \mathbf{R}^N)$ is defined implicitly by the equation

$$\tilde{\rho}\big(x_n, y_n, h(x_n, y_n; \mathbf{R}^N)\big) = (1/2)\rho_\ell. \tag{3.67}$$

We henceforth suppress the dependence of it on $\mathbf{R}^N$. In a slab of water, there are usually two disjoint interfaces at the slab's top and bottom, so this equation has two solutions. For

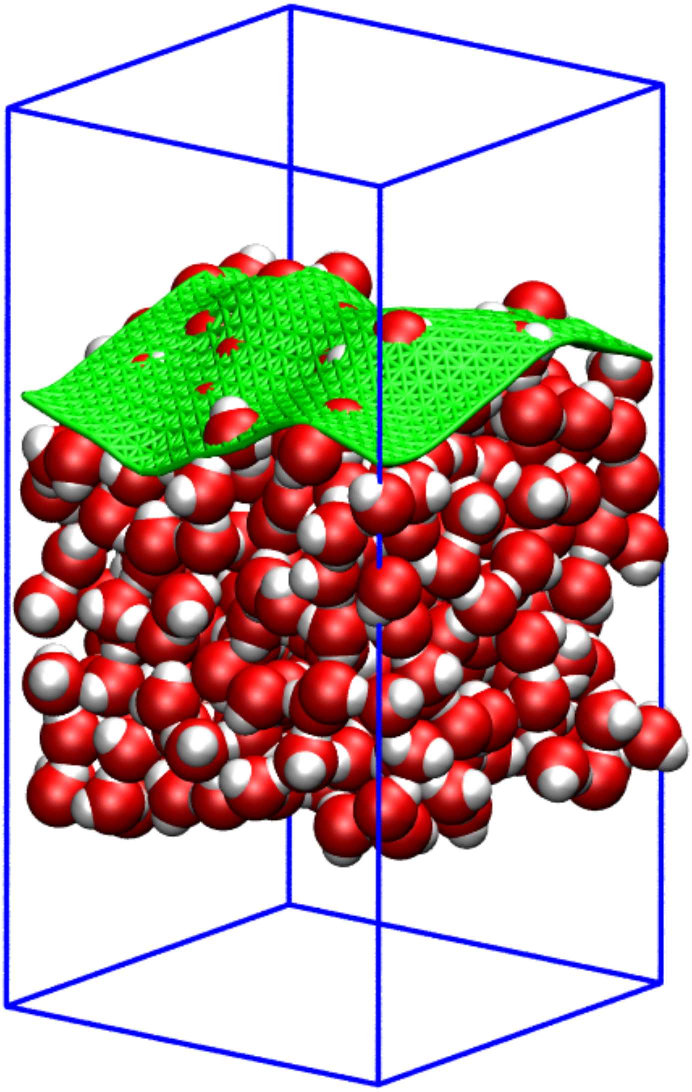


Figure 3.19: Willard interface, with $\xi = 2.4\,\text{Å}$, for an instantaneous configuration of a slab of SPC/E water at 298 K. The periodic box is $21.7 \times 21.7 \times 43.4\,\text{Å}^3$ in size and contains 343 water molecules.

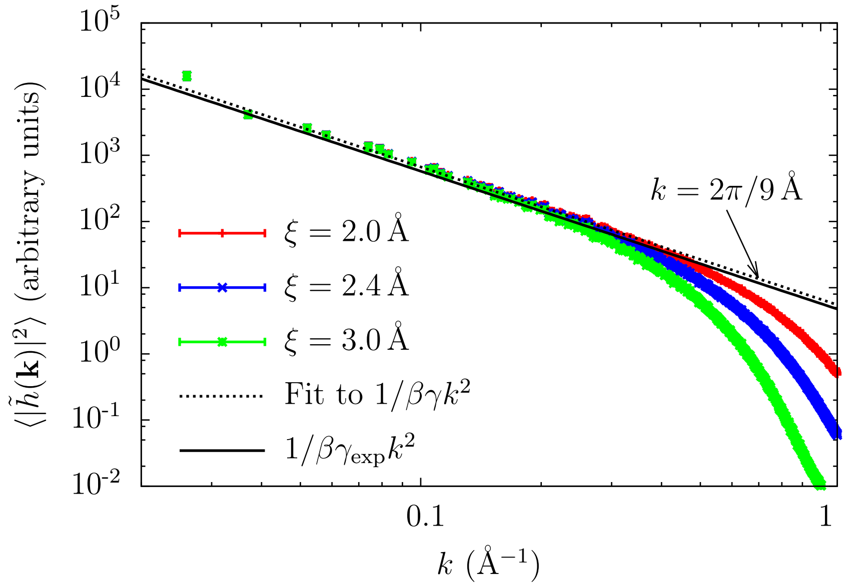


Figure 3.20: Power spectrum of liquid-vapor interface in a $24 \times 24 \times 3\,\text{nm}^3$ slab of SPC/E water at 298 K, defined as the Willard interface with $\xi = 2.0$ Å (red), 2.4 Å (blue) or 3.0 Å (green). A fit to Equation (3.65) in the range $0.01\,\text{Å}^{-1} < k < 0.3\,\text{Å}^{-1}$ (dashes) yields $\gamma = 62.0 \pm 0.5\,\text{mJ/m}^2$, in reasonable agreement with the experimental value (solid) of $72\,\text{mJ/m}^2$. Note: the error bar for the power spectrum at the lowest k value is hard to estimate reliably due to the long correlation time of $\tilde{h}(\mathbf{k})$ for low k. Thus, it is likely that the value shown here, which deviates significantly from the trend of the remaining points, has a larger statistical error than our error estimate suggests.

concreteness, we always refer to the top interface of the slab. The value of $h(x_n, y_n)$ depends on the position of the tagged particle, as well as the positions of all oxygen atoms close to the point $(x_n, y_n, h(x_n, y_n))$.

In order for this umbrella potential to be useful, it must be possible to calculate $h(x_n, y_n)$ quickly at every timestep, as well as its gradient with respect to particle positions. Both calculations must be efficiently parallelizable. The problem is simplified by observing that the value of $h(x_n, y_n)$ at one timestep is similar to its value at the next timestep. We have thus implemented a parallel Newton-Raphson solver to calculate $h(x_n, y_n)$, with the starting guess at one timestep equal to the value of $h(x_n, y_n)$ at the previous timestep. Specifically, let h_i be a guess for the value of $h(x_n, y_n)$. An improved guess, h_{i+1}, is constructed with the relation

$$h_{i+1} = h_i - \frac{\tilde{\rho}(x_n, y_n, h_i) - (1/2)\rho_\ell}{\partial_z \tilde{\rho}(x_n, y_n, h_i)}. \tag{3.68}$$

Both $\tilde{\rho}$ and $\partial_z \tilde{\rho}$ can be calculated in parallel from Equation (3.63) with a procedure analogous to that for calculating $\tilde{N}$ (Section 3.2.1). In a typical molecular dynamics simulation, the Newton-Raphson method converges after one or two iterations.

In order to calculate forces, we need to take the gradient of Equation (3.66) with respect to particle positions. In the expressions below, h and its derivatives are evaluated at (x_n, y_n), while $\tilde{\rho}$ and its derivatives are evaluated at $(x_n, y_n, h(x_n, y_n))$. To simplify the calculation, we assume that the tagged particle n is not a water oxygen, and relax this restriction below. By taking the total derivative of Equation (3.67) with respect to the position of oxygen atom i, we find that

$$\frac{\mathrm{d}(\tilde{\rho} - \rho_\ell/2)}{\mathrm{d}\mathbf{r}_i} = \frac{\partial \tilde{\rho}}{\partial z} \frac{\mathrm{d}h}{\mathrm{d}\mathbf{r}_i} + \frac{\partial \tilde{\rho}}{\partial \mathbf{r}_i} = 0. \tag{3.69}$$

Hence,

$$\frac{\mathrm{d}h}{\mathrm{d}\mathbf{r}_i} = -\frac{\partial \tilde{\rho}}{\partial \mathbf{r}_i} \bigg/ \frac{\partial \tilde{\rho}}{\partial z}. \tag{3.70}$$

The derivative with respect to the position of particle n is obtained similarly, so

$$\frac{\mathrm{d}(\tilde{\rho} - \rho_\ell/2)}{\mathrm{d}x_n} = \frac{\partial \tilde{\rho}}{\partial x} + \frac{\partial \tilde{\rho}}{\partial z} \frac{\mathrm{d}h}{\mathrm{d}x_n} = 0, \tag{3.71a}$$

$$\frac{\mathrm{d}(\tilde{\rho} - \rho_\ell/2)}{\mathrm{d}y_n} = \frac{\partial \tilde{\rho}}{\partial y} + \frac{\partial \tilde{\rho}}{\partial z} \frac{\mathrm{d}h}{\mathrm{d}y_n} = 0, \tag{3.71b}$$

$$\frac{\mathrm{d}(\tilde{\rho} - \rho_\ell/2)}{\mathrm{d}z_n} = 0. \tag{3.71c}$$

Hence,

$$\frac{\mathrm{d}h}{\mathrm{d}x_n} = -\frac{\partial\tilde{\rho}}{\partial x}\bigg/\frac{\partial\tilde{\rho}}{\partial z}, \tag{3.72a}$$

$$\frac{\mathrm{d}h}{\mathrm{d}y_n} = -\frac{\partial\tilde{\rho}}{\partial y}\bigg/\frac{\partial\tilde{\rho}}{\partial z}, \tag{3.72b}$$

$$\frac{\mathrm{d}h}{\mathrm{d}z_n} = 0. \tag{3.72c}$$

If particle n is a water oxygen atom, then $\mathrm{d}h/\mathrm{d}\mathbf{r}_n$ is the sum of the right-hand sides of Equations (3.70) and (3.72).

Assembling all the pieces, we obtain the following umbrella force corresponding to the umbrella potential of Equation (3.66):

$$-\nabla_i V_k(\mathbf{R}^N) = \kappa\big[z_n - h - a_k\big](\nabla_i h - \delta_{in}\hat{z}). \tag{3.73}$$

We have implemented this umbrella potential in LAMMPS and GROMACS. Figure 3.21 illustrates this code in action. It shows a snapshot of a simulation where a tagged water molecule was held suspended about 5 Å above the Willard interface at all times.

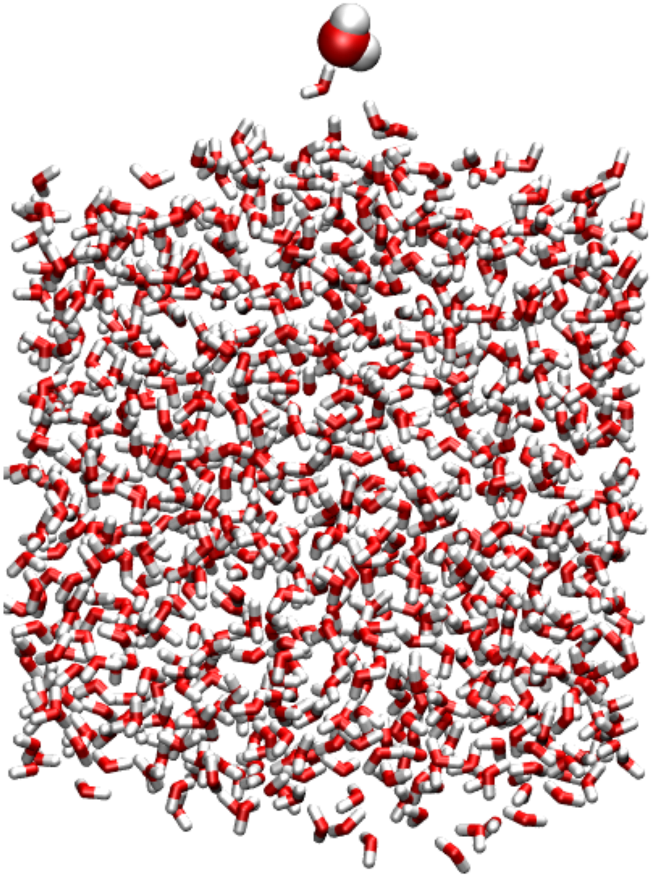


Figure 3.21: Snapshot from a simulation of SPC/E waters (licorice), with an additional umbrella potential of the form in Equation (3.66). Here, the tagged water molecule (solid) is suspended at at a distance $a_k = 5\,\text{Å}$ above the Willard interface.

Chapter 4

An improved coarse-grained model of solvation and the hydrophobic effect

In this chapter, we construct a lattice field theory based on the physical ideas of Lum-Chandler-Weeks (LCW) theory [68, 111]. As described in Section 2.3, LCW theory models solvation by constructing an effective free energy surface for the solvent density field $\rho(\mathbf{r})$. The key idea is to split this density into a slowly-varying density field, $\rho_\ell n(\mathbf{r})$, that supports interfaces, and a quickly-varying overlying fluctuation field, $\delta\rho(\mathbf{r})$, with Gaussian statistics. The energetics of these two fields, when uncoupled, are simple to model, and the two fields are coupled using Weeks and coworkers' unbalancing potential formalism, described in detail in Appendix A. Since the fluctuating field is Gaussian, we can integrate it out analytically to construct an effective field theory for the slowly-varying field alone. External solutes are incorporated as constraints on the total density and then external potentials that couple to solvent density.

We begin our discussion by reviewing constrained Gaussian field theory [11]. Its three main results are: an effective free energy of a Gaussian field in the presence of constraints, the average field configuration and the field's fluctuation spectrum. Using this field theory, we construct the effective LCW Hamiltonian for the slowly-varying field and recast it in a form that is suitable for discretization.

We then discuss how to discretize the effective LCW Hamiltonian. Using a straightforward lattice gas model, isomorphic to the Ising model of magnetism, to capture interfacial energetics, as has been done previously, is shown to suffer from undesirable degeneracies that result in large spurious entropies. We identify the source of these degeneracies and remove them by carefully discretizing a Landau-Ginzburg Hamiltonian. We then discretize the remainder of the effective Hamiltonian. The final equations, compiled in Table 4.7, constitute the main result of this chapter, a simple lattice gas-like Hamiltonian for modeling solvation.

In the remainder of the chapter, we compute solvation free energies and water number distributions in a number of settings, and compare the model results to the explicit-water simulation results obtained in Chapter 3. These comparisons demonstrate that the coarse-grained model of Table 4.7 accurately captures many of the important phenomena that we associate with hydrophobic solvation.

In the appendices to the chapter, we discuss some of the implementation details of the model. We also consider the extensions to the model that would be necessary to implement dynamics and solvent polarization, two aspects that are crucial to biological solvation.

4.1 Preliminaries

4.1.1 Gaussian integrals

We begin by reviewing Gaussian integrals, which play a central role in this chapter. Let $\mathbf{x}$ and $\mathbf{y}$ be N-component vectors, with components $\{x_i\}$ and $\{y_i\}$, and let $\mathbf{A}$ be an $N \times N$ real, symmetric matrix, with matrix elements A_{ij}. By diagonalizing $\mathbf{A}$, it can be shown that

$$\int \left(\prod_{i=1}^{N} \frac{\mathrm{d}x_i}{2\pi} \right) \exp\left[-\frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} \right] = \frac{1}{\sqrt{\det(2\pi \mathbf{A})}}. \tag{4.1}$$

If $\mathbf{A}$ is invertible, we can obtain the more general result,

$$\begin{aligned} &\int \left(\prod_{i=1}^{N} \frac{\mathrm{d}x_i}{2\pi} \right) \exp\left[-\frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{y} \right] \\ &\quad = \int \left(\prod_{i=1}^{N} \frac{\mathrm{d}x_i}{2\pi} \right) \exp\left[-\frac{1}{2} (\mathbf{x} + \mathbf{A}^{-1}\mathbf{y})^T \mathbf{A} (\mathbf{x} + \mathbf{A}^{-1}\mathbf{y}) + \frac{1}{2} \mathbf{y}^T \mathbf{A}^{-1} \mathbf{y} \right] \\ &\quad = \exp\left[\frac{1}{2} \mathbf{y}^T \mathbf{A}^{-1} \mathbf{y} \right] \frac{1}{\sqrt{\det(2\pi \mathbf{A})}}. \end{aligned} \tag{4.2}$$

We can regard the components of $\mathbf{x}$ as dynamical variables of a system whose Hamiltonian is $H(\mathbf{x}) = (k_{\mathrm{B}}T/2)\mathbf{x}^T \mathbf{A} \mathbf{x}$. If we regard the components of the vector $k_{\mathrm{B}}T\mathbf{y}$ as conjugate variables that are coupled to the components of $\mathbf{x}$, then Equation (4.2) is the partition function of the coupled system. By taking partial derivatives of the logarithm of this partition function, we find that

$$\langle x_i \rangle = -\sum_j (A^{-1})_{ij} y_j. \tag{4.3}$$

Moreover, letting $\delta x_i = x_i - \langle x_i \rangle$, we find that

$$\langle \delta x_i \delta x_j \rangle = (A^{-1})_{ij}. \tag{4.4}$$

These results can be generalized to Gaussian integrals over fields. Let $f(\mathbf{r})$ and $g(\mathbf{r})$ be arbitrary fields and $u(\mathbf{r},\mathbf{r}')$ be a linear operator. Then,

$$\int \mathcal{D}f(\mathbf{r})\,\exp\left[-\frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} f(\mathbf{r})u(\mathbf{r},\mathbf{r}')f(\mathbf{r}') - \int_{\mathbf{r}} f(\mathbf{r})g(\mathbf{r})\right] = \exp\left[\frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} g(\mathbf{r})u^{-1}(\mathbf{r},\mathbf{r}')g(\mathbf{r}')\right]\frac{1}{\sqrt{\det(2\pi u)}}. \tag{4.5}$$

Note that there are factors of 2π implicit in the notation $\mathcal{D}f(\mathbf{r})$. The operator $u^{-1}(\mathbf{r},\mathbf{r}')$ is defined by the relation

$$\int_{\mathbf{r}''} u^{-1}(\mathbf{r},\mathbf{r}'')u(\mathbf{r}'',\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'), \quad \forall \mathbf{r},\mathbf{r}'. \tag{4.6}$$

Equation (4.5) is derived by expanding $f(\mathbf{r})$ in a finite basis, and then applying Equation (4.2). The functional determinant of $2\pi u$ is only defined in this basis, and will have a different numerical value in a different basis. However, the ratio of two such determinants is independent of basis in the limit of a complete basis, so only ratios of integrals like those on the left-hand side of Equation (4.5) have physical meaning.

We now regard $f(\mathbf{r})$ as a dynamical field, with Hamiltonian,

$$H[f(\mathbf{r})] = (k_{\mathrm{B}}T/2)\int_{\mathbf{r}}\int_{\mathbf{r}'} f(\mathbf{r})u(\mathbf{r},\mathbf{r}')f(\mathbf{r}'), \tag{4.7}$$

and couple $f(\mathbf{r})$ to the field $k_{\mathrm{B}}Tg(\mathbf{r})$. Equation (4.5) is the partition function for this system. Generalizing the results for $\langle x_i\rangle$ and $\langle \delta x_i \delta x_j\rangle$ above, we find that

$$\langle f(\mathbf{r})\rangle = -\int_{\mathbf{r}'} u^{-1}(\mathbf{r},\mathbf{r}')g(\mathbf{r}'), \tag{4.8}$$

$$\langle \delta f(\mathbf{r})\delta f(\mathbf{r}')\rangle = u^{-1}(\mathbf{r},\mathbf{r}'). \tag{4.9}$$

4.1.2 Unconstrained Gaussian solvent density fields

Chandler has demonstrated that many successful liquid-state theories, such as Percus-Yevick theory and the mean spherical approximation [41], result from assuming Gaussian statistics for solvent density fluctuations [11]. The total density field for a non-uniform solvent is first decomposed into slowly-varying and quickly-varying components, as described in Section 2.3,

$$\rho(\mathbf{r}) = \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r}). \tag{4.10}$$

The special case of a uniform fluid is represented by setting $n(\mathbf{r}) = 1$ everywhere. The solvent density fluctuation field $\delta\rho(\mathbf{r})$ is then well-modeled by a Hamiltonian of the form

$$H[\delta\rho(\mathbf{r})] = \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \delta\rho(\mathbf{r})\chi^{-1}(\mathbf{r},\mathbf{r}')\delta\rho(\mathbf{r}'). \tag{4.11}$$

Solvent density field	$\rho(\mathbf{r}) = \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})$
External potential	$\Phi(\mathbf{r})$
Bare Hamiltonian	$H[\delta\rho(\mathbf{r})] = \frac{k_\mathrm{B}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \delta\rho(\mathbf{r}) \chi^{-1}(\mathbf{r}, \mathbf{r}') \delta\rho(\mathbf{r}')$
Partition function	$Z[\Phi(\mathbf{r})] = \int \mathcal{D}\delta\rho(\mathbf{r})\, e^{-\beta H[\delta\rho(\mathbf{r})] - \beta \int_{\mathbf{r}} \Phi(\mathbf{r})\delta\rho(\mathbf{r})}$
Free energy	$F[\Phi(\mathbf{r})] = -\frac{k_\mathrm{B}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \beta\Phi(\mathbf{r}) \chi(\mathbf{r}, \mathbf{r}') \beta\Phi(\mathbf{r}')$
Fluctuation spectrum	$\langle \delta\rho(\mathbf{r}) \delta\rho(\mathbf{r}') \rangle = \chi(\mathbf{r}, \mathbf{r}')$
Mean density	$\langle \rho(\mathbf{r}) \rangle = \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'} \chi(\mathbf{r}, \mathbf{r}') \beta\Phi(\mathbf{r}')$

Table 4.1: Main results of unconstrained Gaussian field theory for solvent density field. The free energy is measured with respect to the ensemble where $\Phi(\mathbf{r}) = 0$, so $F[\Phi(\mathbf{r})] = -k_\mathrm{B}T \ln Z[\Phi(\mathbf{r})]/Z[0]$.

Here, the coupling strength $\chi^{-1}(\mathbf{r}, \mathbf{r}')$ is the operator that satisfies the relation

$$\int_{\mathbf{r}''} \chi^{-1}(\mathbf{r}, \mathbf{r}'') \chi(\mathbf{r}'', \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \tag{4.12}$$

for all $\mathbf{r}$ and $\mathbf{r}'$. By construction, the thermal average of $\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')$ under this Hamiltonian is $\chi(\mathbf{r}, \mathbf{r}')$. If a field $\Phi(\mathbf{r})$ is coupled to the solvent density fluctuation field $\delta\rho(\mathbf{r})$, then the resulting mean solvent density and fluctuation spectrum are given by analogs to Equations (4.8) and (4.9). These results are summarized in Table 4.1, and recapitulate the linear response regime of normal liquids [41].

4.1.3 Constrained Gaussian solvent density fields

Chandler then considered the free energy cost of imposing constraints on the density fluctuations as a way to model solvent density exclusion. Following his argument, we impose the requirement $\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r}) = 0$ for all $\mathbf{r}$ in a given volume v. The solvent density in the remaining volume, denoted $\bar{v}$, is unconstrained. To facilitate later calculations, we again couple an external field $\Phi(\mathbf{r})$ to $\delta\rho(\mathbf{r})$. The partition function of the resulting ensemble is

$$Z_v[\Phi(\mathbf{r})] = \int \mathcal{D}\delta\rho(\mathbf{r})\, \exp\left[-\beta H[\delta\rho(\mathbf{r})] - \beta \int_{\mathbf{r}} \Phi(\mathbf{r})\delta\rho(\mathbf{r})\right] \prod_{\mathbf{r}\in v} \delta\big(\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})\big). \tag{4.13}$$

The integral is most easily evaluated by using the Fourier space representation of the δ-function. In particular,

$$Z_v[\Phi(\mathbf{r})] = \int \mathcal{D}\delta\rho(\mathbf{r}) \int_v \mathcal{D}\psi(\mathbf{r}) \exp\Big[-\frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \delta\rho(\mathbf{r})\chi^{-1}(\mathbf{r},\mathbf{r}')\delta\rho(\mathbf{r}') - \beta\int_{\mathbf{r}} \Phi(\mathbf{r})\delta\rho(\mathbf{r}) - \int_{\mathbf{r}} i\psi(\mathbf{r})\big(\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})\big)\Big], \quad (4.14)$$

where the subscript v in the integral over the dummy conjugate field $\psi(\mathbf{r})$ signifies that the $\psi(\mathbf{r})$ is only non-zero when $\mathbf{r}$ is in v. The integral over $\delta\rho(\mathbf{r})$ is now purely a Gaussian integral, so it can be evaluated analytically using Equation (4.5). The result is

$$Z_v[\Phi(\mathbf{r})] = \frac{1}{\sqrt{\det(2\pi\chi^{-1})}} \int_v \mathcal{D}\psi(\mathbf{r}) \exp\left\{\frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} [\beta\Phi(\mathbf{r}) + i\psi(\mathbf{r})]\chi(\mathbf{r},\mathbf{r}')[\beta\Phi(\mathbf{r}') + i\psi(\mathbf{r}')] - \int_{\mathbf{r}} i\psi(\mathbf{r})\rho_\ell n(\mathbf{r})\right\}. \quad (4.15)$$

The integral over the field $\psi(\mathbf{r})$ can be reduced to the following Gaussian integral,

$$Z_v[\Phi(\mathbf{r})] = \frac{1}{\sqrt{\det(2\pi\chi^{-1})}} \exp\left[\frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \beta\Phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}')\right] \int_v \mathcal{D}\psi(\mathbf{r}) \exp\Big\{-\frac{1}{2}\int_{\mathbf{r}\in v}\int_{\mathbf{r}'\in v} \psi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')\psi(\mathbf{r}') - i\int_{\mathbf{r}\in v} \psi(\mathbf{r})\Big[\rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'} \chi(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}')\Big]\Big\} \quad (4.16)$$

When evaluated, the result is

$$Z_v[\Phi(\mathbf{r})] = \frac{1}{\sqrt{\det(2\pi\chi^{-1})}} \exp\left[\frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \beta\Phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}')\right] \frac{1}{\sqrt{\det(2\pi\chi_v)}} \exp\Big\{-\frac{1}{2}\int_{\mathbf{r}\in v}\int_{\mathbf{r}'\in v} \Big[\rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}''} \chi(\mathbf{r},\mathbf{r}'')\beta\Phi(\mathbf{r}'')\Big] \chi_v^{-1}(\mathbf{r},\mathbf{r}')\Big[\rho_\ell n(\mathbf{r}') - \int_{\mathbf{r}'''} \chi(\mathbf{r}',\mathbf{r}''')\beta\Phi(\mathbf{r}''')\Big]\Big\}. \quad (4.17)$$

Here, $\chi_v^{-1}(\mathbf{r},\mathbf{r}')$ is the operator that satisfies

$$\int_{\mathbf{r}''\in v} \chi_v^{-1}(\mathbf{r},\mathbf{r}'')\chi(\mathbf{r}'',\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'), \qquad \forall \mathbf{r},\mathbf{r}' \in v. \quad (4.18)$$

The free energy difference between the free ensemble and the constrained ensemble in the presence of the external field $\Phi(\mathbf{r})$, defined as $-k_{\mathrm{B}}T \ln Z_v[\Phi(\mathbf{r})]/Z[0]$, is given by

$$\begin{aligned} F_v[\Phi(\mathbf{r})] = k_{\mathrm{B}}T \ln \sqrt{\det(2\pi\chi_v)} &- \frac{k_{\mathrm{B}}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \beta\Phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}') \\ &+ \frac{k_{\mathrm{B}}T}{2} \int_{\mathbf{r}\in v} \int_{\mathbf{r}'\in v} \left[\rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}''} \chi(\mathbf{r},\mathbf{r}'')\beta\Phi(\mathbf{r}'')\right] \\ &\qquad \chi_v^{-1}(\mathbf{r},\mathbf{r}')\left[\rho_\ell n(\mathbf{r}') - \int_{\mathbf{r}'''} \chi(\mathbf{r}',\mathbf{r}''')\beta\Phi(\mathbf{r}''')\right]. \end{aligned} \tag{4.19}$$

In this expression, we can recognize the terms in square brackets as the local solvent density that would result from a linear response in the absence of constraints of the mean solvent density at $\mathbf{r}$ to a field $\Phi(\mathbf{r})$ that couples to $\delta\rho(\mathbf{r})$ (see Table 4.1). The free energy $F_v[\Phi(\mathbf{r})]$ is thus a multi-variable generalization of the Gaussian model expression of Equation (2.27).

At this point, we hit an impasse: because of the term $\det(2\pi\chi_v)$, this free energy depends on the basis used to evaluate the functional integral in Equation (4.13). In Section 4.1.5, we discuss the origin of this problem and argue that a pragmatic resolution is to replace $\det(2\pi\chi_v)$ with $2\pi\sigma_v$, where σ_v is defined by Equation (2.23). For the moment, we sidestep this issue and discuss other consequences of Equation (4.19).

By taking functional derivatives of $F_v[\Phi(\mathbf{r})]$ with respect to $\Phi(\mathbf{r})$, we can calculate the mean value of various observables in the presence of constraints and external fields. We use the subscript v to denote that the thermal averages are evaluated with these constraints and external fields. In this way, we obtain two key results. First, the average solvent density is given by

$$\begin{aligned} \langle\rho(\mathbf{r})\rangle_v &= \rho_\ell n(\mathbf{r}) + \frac{\delta\beta F_v[\Phi(\mathbf{r})]}{\delta\beta\Phi(\mathbf{r})} \\ &= \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'} \chi(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}') - \int_{\mathbf{r}'\in v} \int_{\mathbf{r}''\in v} \chi(\mathbf{r},\mathbf{r}')\chi_v^{-1}(\mathbf{r}',\mathbf{r}'') \\ &\qquad \left[\rho_\ell n(\mathbf{r}'') - \int_{\mathbf{r}'''} \chi(\mathbf{r}'',\mathbf{r}''')\beta\Phi(\mathbf{r}''')\right]. \end{aligned} \tag{4.20}$$

Second, the fluctuation correlation function, denoted henceforth by $\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$, is given by

$$\begin{aligned} \chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}') = \langle\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\rangle_v &= \frac{\delta^2\beta F_v[\Phi(\mathbf{r})]}{\delta\beta\Phi(\mathbf{r})\delta\beta\Phi(\mathbf{r}')} \\ &= \chi(\mathbf{r},\mathbf{r}') - \int_{\mathbf{r}''\in v} \int_{\mathbf{r}'''\in v} \chi(\mathbf{r},\mathbf{r}'')\chi_v^{-1}(\mathbf{r}'',\mathbf{r}''')\chi(\mathbf{r}''',\mathbf{r}'). \end{aligned} \tag{4.21}$$

This modified fluctuation spectrum differs from $\chi(\mathbf{r},\mathbf{r}')$ principally in that it is 0 if either $\mathbf{r}$ or $\mathbf{r}'$ is in v, a result that follows from Equation (4.18). Since $\chi(\mathbf{r},\mathbf{r}')$ is usually short-ranged (it decays to near-zero if $\mathbf{r}$ and $\mathbf{r}'$ are more than about 10 Å from each other),

$\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$ and $\chi(\mathbf{r},\mathbf{r}')$ are equal whenever both $\mathbf{r}$ and $\mathbf{r}'$ are far from v. In the vicinity of v, $\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$ differs quantitatively but not qualitatively from $\chi(\mathbf{r},\mathbf{r}')$.

The expression for the mean solvent density, Equation (4.20), has a clean physical interpretation. To reveal it, we first construct an auxiliary field, $c(\mathbf{r})$, that satisfies,

$$-\int_{\mathbf{r}'\in v}\chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}') = -\rho_\ell n(\mathbf{r}), \qquad \mathbf{r}\in v, \tag{4.22a}$$

$$c(\mathbf{r}) = 0, \qquad \mathbf{r}\notin v. \tag{4.22b}$$

In other words, $k_{\mathrm{B}}Tc(\mathbf{r})$ is a field, zero outside of v, that induces a solvent density response $-\rho_\ell n(\mathbf{r})$ inside of v. With this field, we can rewrite Equation (4.20) as

$$\langle\rho(\mathbf{r})\rangle_v = \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'\in v}\chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}') - \int_{\mathbf{r}'\in\bar{v}}\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}'). \tag{4.23}$$

The restricted domains of the integrals only emphasize the regions in which the integrand is not trivially zero. This expression can be interpreted as follows: the effect of the constraints is identical to that of an external field $k_{\mathrm{B}}Tc(\mathbf{r})$ that couples to the solvent density in v, chosen to make the mean solvent density 0 in this region; the solvent further responds linearly to the external field $\Phi(\mathbf{r})$, but with the modified fluctuation spectrum $\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$. This expression makes evident that $\langle\rho(\mathbf{r})\rangle_v$ is zero inside v, as required by the density constraints, and that its value outside of v is independent of the value of the field $\Phi(\mathbf{r})$ in v.

By using the expressions for $c(\mathbf{r})$ and $\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$, we can rewrite the free energy of the constrained field in a way that emphasizes the physical effects of the constraints. We obtain

$$\begin{aligned} F_v[\Phi(\mathbf{r})] = k_{\mathrm{B}}T\ln\sqrt{\det(2\pi\chi_v)} + \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}\in v}\rho_\ell n(\mathbf{r})c(\mathbf{r}) - \int_{\mathbf{r}\in v}\Phi(\mathbf{r})\rho_\ell n(\mathbf{r}) \\ - \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'}\beta\Phi(\mathbf{r})\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}') - \int_{\mathbf{r}\in\bar{v}}\int_{\mathbf{r}'\in v}\Phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}'). \end{aligned} \tag{4.24}$$

The results of constrained Gaussian field theory are summarized in Table 4.2

4.1.4 Effective LCW Hamiltonian

We now consider the full LCW Hamiltonian, discussed in depth in Section 2.3. The microscopic details of this Hamiltonian, in the presence of an external potential $u(\mathbf{r})$, are summarized in Table 4.3.

Using the results of the previous section, we can derive an effective Hamiltonian $H_{\mathrm{eff}}[n(\mathbf{r})]$ for the field $n(\mathbf{r})$, which is the result of thermally averaging over the fluctuating field $\delta\rho(\mathbf{r})$ under the constraint of no solvent density in v. The Hamiltonian is built so that in the absence of constraints and external field $u(\mathbf{r})$, then $H_{\mathrm{eff}}[n(\mathbf{r})] = H_L[n(\mathbf{r})]$. This effective Hamiltonian can be decomposed into portions owing to the interfacial energetics, $H_L[n(\mathbf{r})]$,

Solvent density field	$\rho(\mathbf{r}) = \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})$
External potential	$\Phi(\mathbf{r})$
Bare Hamiltonian	$H[\delta\rho(\mathbf{r})] = \frac{k_\mathrm{B}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \delta\rho(\mathbf{r})\chi^{-1}(\mathbf{r},\mathbf{r}')\delta\rho(\mathbf{r}')$
Partition function	$Z_v[\Phi(\mathbf{r})] = \int \mathcal{D}\delta\rho(\mathbf{r})\, e^{-\beta H[\delta\rho(\mathbf{r})]-\beta\int_{\mathbf{r}}\Phi(\mathbf{r})\delta\rho(\mathbf{r})} \prod_{\mathbf{r}\in v}\delta\big(\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})\big)$
Solvent exclusion potential	$c(\mathbf{r}) = \begin{cases} \int_{\mathbf{r}'\in v}\chi_v^{-1}(\mathbf{r},\mathbf{r}')\rho_\ell n(\mathbf{r}'), & \mathbf{r}\in v \\ 0, & \text{otherwise} \end{cases}$
Free energy	$F_v[\Phi(\mathbf{r})] = k_\mathrm{B}T\ln\sqrt{\det(2\pi\chi_v)} + \frac{k_\mathrm{B}T}{2}\int_{\mathbf{r}\in v}\rho_\ell n(\mathbf{r})c(\mathbf{r})$ $- \int_{\mathbf{r}\in v}\Phi(\mathbf{r})\rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}\in\bar{v}}\int_{\mathbf{r}'\in v}\Phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}')$ $- \frac{k_\mathrm{B}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'}\beta\Phi(\mathbf{r})\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}').$
Fluctuation spectrum	$\langle\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\rangle_v = \chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$ $= \chi(\mathbf{r},\mathbf{r}') - \int_{\mathbf{r}''\in v}\int_{\mathbf{r}'''\in v}\chi(\mathbf{r},\mathbf{r}'')\chi_v^{-1}(\mathbf{r}'',\mathbf{r}''')\chi(\mathbf{r}''',\mathbf{r}')$
Mean density	$\langle\rho(\mathbf{r})\rangle_v = \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'\in v}\chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}') - \int_{\mathbf{r}\in\bar{v}}\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}')$

Table 4.2: Main results of constrained Gaussian field theory for solvent density field. The free energy is measured with respect to the unconstrained ensemble where $\Phi(\mathbf{r}) = 0$, so $F_v[\Phi(\mathbf{r})] = -k_\mathrm{B}T\ln Z_v[\Phi(\mathbf{r})]/Z[0]$.

Solvent density field	$\rho(\mathbf{r}) = \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})$
Unbalancing potential	$\phi(\mathbf{r}) = -2a\rho_\ell[\overline{n}(\mathbf{r}) - 1]$
Hamiltonian	$H[n(\mathbf{r}), \delta\rho(\mathbf{r})] = H_L[n(\mathbf{r})] + H_S[\delta\rho(\mathbf{r}); n(\mathbf{r})]$
	$+ H_I[n(\mathbf{r}), \delta\rho(\mathbf{r})] + H_u[n(\mathbf{r}), \delta\rho(\mathbf{r})]$
Interfacial energy	$H_L[n(\mathbf{r})] = \int_{\mathbf{r}} w(n(\mathbf{r}), \mu) + \frac{m}{2} \lvert\nabla n(\mathbf{r})\rvert^2$
Fluctuation energy	$H_S[\delta\rho(\mathbf{r}); n(\mathbf{r})] = \frac{k_\mathrm{B}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \delta\rho(\mathbf{r}) \chi^{-1}\big(\mathbf{r}, \mathbf{r}'; n(\mathbf{r})\big) \delta\rho(\mathbf{r}')$
Length-scale coupling	$H_I[n(\mathbf{r}), \delta\rho(\mathbf{r})] = \int_{\mathbf{r}} \phi(\mathbf{r})\delta\rho(\mathbf{r}) + H_\mathrm{norm}[n(\mathbf{r})]$
Normalization	$H_\mathrm{norm}[n(\mathbf{r})] = \frac{k_\mathrm{B}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \phi(\mathbf{r})\chi(\mathbf{r}, \mathbf{r}')\phi(\mathbf{r}')$
External coupling	$H_u[n(\mathbf{r}), \delta\rho(\mathbf{r})] = \int_{\mathbf{r}} u(\mathbf{r})[\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})]$

Table 4.3: Microscopic LCW model.

to the constrained solvent fluctuations, $H_S[n(\mathbf{r})]$, the effect of the unbalancing potential, $H_I[n(\mathbf{r})]$, and finally, the effect of the coupling to the external field, $H_u[n(\mathbf{r})]$. Expressions for all these Hamiltonians are collected in Table 4.4.

Before proceeding, we introduce two important simplifications. First, instead of solving Equation (4.22) to obtain the value of the field $c(\mathbf{r})$ in v, we replace $c(\mathbf{r})$ there by its average value, c_1, and obtain the much simpler relation

$$c_1 = \langle N \rangle_v / \sigma_v, \tag{4.25}$$

where

$$\langle N \rangle_v = \int_{\mathbf{r}\in v} \rho_\ell n(\mathbf{r}), \tag{4.26}$$

$$\sigma_v = \int_{\mathbf{r}\in v} \int_{\mathbf{r}'\in v} \chi(\mathbf{r}, \mathbf{r}'). \tag{4.27}$$

As described below, this approximation is closely connected to the effective Hamiltonian that would result if in Equation (4.13) we enforced only that $\int_{\mathbf{r}\in v} \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r}) = 0$, instead of imposing the multitude of constraints that $\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})$ be zero at every point $\mathbf{r}$ in v. We present numerical evidence below that this approximation, dubbed the "one-basis set approximation" in previous works [48, 107], does not appreciably change the results we obtain. Crucially, this approximation replaces the large (though sparse) linear system of Equation (4.22) with the trivial relation of Equation (4.25), and is therefore very advantageous computationally. With it, the term $H_S[n(\mathbf{r})]$ is given by

$$H_S[n(\mathbf{r})] \approx k_{\mathrm{B}}T[\langle N \rangle_v^2 / 2\sigma_v + C/2]. \tag{4.28}$$

The normalization constant C should strictly be equal to $\ln(2\pi\sigma_v)$, but this quantity tends to negative infinity as v becomes tiny, since in that limit, $\sigma_v \approx \langle N \rangle_v$. This deficiency arises from a breakdown of Gaussian statistics for solvent number fluctuations in sub-Angstrom volumes. Since solvent molecules are discrete entities, these statistics are instead Poissonian. The free-energy cost of evacuating such a cavity is approximately $k_{\mathrm{B}}T \langle N \rangle_v$ (Equation (2.21)). We thus define C by the relation

$$C = \begin{cases} \ln(2\pi\sigma_v), & \langle N \rangle_v > 1, \\ \max[\ln(2\pi\sigma_v), \langle N \rangle_v], & \text{otherwise.} \end{cases} \tag{4.29}$$

This is a simple, continuous way of compensating for the difference in fluctuation statistics at tiny length scales. The crossover[1] to $C = \langle N \rangle_v$ occurs at $\langle N \rangle_v \approx (2\pi - 1)^{-1} \approx 0.19$.

A second important simplification is that in all the solute geometries we have used, the sum of the last two terms of $H_I[n(\mathbf{r})]$ in Table 4.4 are, on average, opposite in sign but nearly

---

[1] In Ref. [111], there is an inconsequential typo in the value of this crossover.

Solvent density field	$\rho(\mathbf{r}) = \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})$
Unbalancing potential	$\phi(\mathbf{r}) = -2a\rho_\ell[\overline{n}(\mathbf{r}) - 1]$
Partition function	$Z_v[n(\mathbf{r})] = \int \mathcal{D}\delta\rho(\mathbf{r})\, e^{-\beta H[n(\mathbf{r}),\delta\rho(\mathbf{r})]} \prod_{\mathbf{r}\in v} \delta\big(\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})\big)$
Effective Hamiltonian	$H_{\mathrm{eff}}[n(\mathbf{r})] = H_L[n(\mathbf{r})] + H_S[n(\mathbf{r})] + H_I[n(\mathbf{r})] + H_u[n(\mathbf{r})]$
Interfacial energy	$H_L[n(\mathbf{r})] = \int_{\mathbf{r}} w(n(\mathbf{r}),\mu) + \frac{m}{2}\lvert\nabla n(\mathbf{r})\rvert^2$
Solvent exclusion potential	$c(\mathbf{r}) = \begin{cases} \int_{\mathbf{r}'\in v} \chi_v^{-1}(\mathbf{r},\mathbf{r}')\rho_\ell n(\mathbf{r}'), & \mathbf{r}\in v \\ 0, & \text{otherwise} \end{cases}$
Solvent exclusion energy	$H_S[n(\mathbf{r})] = k_{\mathrm{B}}T \ln\sqrt{\det(2\pi\chi_v)} + \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}\in v} \rho_\ell n(\mathbf{r}) c(\mathbf{r})$
Length-scale coupling	$H_I[n(\mathbf{r})] = -\int_{\mathbf{r}\in v} \phi(\mathbf{r})\rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}\in\bar{v}}\int_{\mathbf{r}'\in v} \Phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}')$ $+ \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \beta\Phi(\mathbf{r})(\chi - \chi^{(\mathrm{m})})(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}').$
External coupling	$H_u[n(\mathbf{r})] = \int_{\mathbf{r}} u(\mathbf{r})\left\langle\rho(\mathbf{r})\right\rangle_v + \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \beta u(\mathbf{r})\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta u(\mathbf{r}')$
Fluctuation spectrum	$\left\langle\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\right\rangle_v = \chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$ $= \chi(\mathbf{r},\mathbf{r}') - \int_{\mathbf{r}''\in v}\int_{\mathbf{r}'''\in v} \chi(\mathbf{r},\mathbf{r}'')\chi_v^{-1}(\mathbf{r}'',\mathbf{r}''')\chi(\mathbf{r}''',\mathbf{r}')$
Mean density	$\left\langle\rho(\mathbf{r})\right\rangle_v = \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'\in v} \chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}') - \int_{\mathbf{r}\in\bar{v}} \chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta(u+\phi)(\mathbf{r}')$

Table 4.4: Effective LCW Hamiltonian for slowly-varying field $n(\mathbf{r})$. The microscopic Hamiltonian $H[n(\mathbf{r}),\delta\rho(\mathbf{r})]$ is summarized in Table 4.3.

proportional to the much simpler remaining term involving $\phi(\mathbf{r})$ (as described in detail below in Section 4.4.3). Physically, these three terms capture the energetic bonus of driving $\delta\rho(\mathbf{r})$ to 0 inside v where ϕ is positive, the energetic cost of the consequent density enhancement just outside of v, and the small difference between (a) the entropic cost associated with ϕ modifying the solvent density in the presence of a solute and (b) that same cost in the absence of a solute. In typical configurations, the three terms are roughly proportional to the subvolume of v where $n(\mathbf{r}) = 1$, and capture how solvation free energies are modified by the microscopic curvature of v. We have found it accurate to model the effect of these three terms using *only* the first term of $H_I[n(\mathbf{r})]$, whose strength is then renormalized by a factor K. The resulting approximation for $H_I[n(\mathbf{r})]$ is

$$H_I[n(\mathbf{r})] \approx -K \int_{r\in v} \phi(\mathbf{r})\rho_\ell n(\mathbf{r}). \tag{4.30}$$

The effective LCW Hamiltonian with these simplifications is summarized in Table 4.5. In the remainder of this chapter, we propose and evaluate discretization schemes for these equations.

4.1.5 Shortcomings of the Gaussian model

Before concluding this section, we return to the basis-dependence problem in Equation (4.19). In particular, the value of $\det(2\pi\chi_v)$ depends on the basis used to evaluate the functional integral of Equation (4.16). The problem is a symptom of the Gaussian model breaking down when the particulate nature of the solvent becomes evident. That is, the Gaussian model allows the total number of particles in a given volume to be non-integer or even negative. This is the same problem that appears in calculating the solvation free energy of a tiny cavity with the Gaussian model, and is alleviated by our definition of C above.

If the functional integral over density fields in Equation (4.13) could be restricted so that the total solvent density was everywhere non-negative, then the single constraint $\int_{\mathbf{r}\in v} \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r}) = 0$ would be equivalent to the multitude of constraints imposed in Equation (4.13). When only the single constraint is imposed on the Gaussian model, the resulting free energy difference is

$$\beta F_v^{(1)}[\Phi(\mathbf{r})] = \ln\sqrt{2\pi\sigma_v} + \frac{\left[\int_{\mathbf{r}\in v} \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'} \chi(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}')\right]^2}{2\sigma_v} - \frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \beta\Phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')\beta\Phi(\mathbf{r}'), \tag{4.31}$$

with σ_v defined by Equation (2.23). In the absence of an external field $\Phi(\mathbf{r})$, this free energy is equal to $H_S[n(\mathbf{r})]$ in the one-basis approximation (Table 4.5). The free energy is also *exactly* the Gaussian model of Equation (2.27), which we know to be quantitatively accurate

Solvent density field	$\rho(\mathbf{r}) = \rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})$
Unbalancing potential	$\phi(\mathbf{r}) = -2a\rho_\ell[\overline{n}(\mathbf{r}) - 1]$
Partition function	$Z_v[n(\mathbf{r})] = \int \mathcal{D}\delta\rho(\mathbf{r})\, e^{-\beta H[n(\mathbf{r}),\delta\rho(\mathbf{r})]} \prod_{\mathbf{r}\in v} \delta\big(\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})\big)$
Effective Hamiltonian	$H_{\text{eff}}[n(\mathbf{r})] = H_L[n(\mathbf{r})] + H_S[n(\mathbf{r})] + H_I[n(\mathbf{r})] + H_u[n(\mathbf{r})]$
Interfacial energy	$H_L[n(\mathbf{r})] = \int_{\mathbf{r}} w(n(\mathbf{r}),\mu) + \frac{m}{2}\left\lvert\nabla n(\mathbf{r})\right\rvert^2$
Solvent exclusion potential	$c_1 = \langle N\rangle_v / \sigma_v$
	$\langle N\rangle_v = \int_{\mathbf{r}\in v} \rho_\ell n(\mathbf{r})$
	$\sigma_v = \int_{\mathbf{r}\in v}\int_{\mathbf{r}'\in v} \chi(\mathbf{r},\mathbf{r}')$
Solvent exclusion energy	$H_S[n(\mathbf{r})] \approx k_{\mathrm{B}}T[\langle N\rangle_v^2 / 2\sigma_v + C/2]$,
	$C = \begin{cases} \ln(2\pi\sigma_v), & \langle N\rangle_v > 1, \\ \max[\ln(2\pi\sigma_v), \langle N\rangle_v], & \text{otherwise.} \end{cases}$
Length-scale coupling	$H_I[n(\mathbf{r})] \approx -K \int_{\mathbf{r}\in v} \phi(\mathbf{r})\rho_\ell n(\mathbf{r})$
External coupling	$H_u[n(\mathbf{r})] = \int_{\mathbf{r}} u(\mathbf{r})\,\langle\rho(\mathbf{r})\rangle_v + \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \beta u(\mathbf{r})\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta u(\mathbf{r}')$
Fluctuation spectrum	$\langle\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\rangle_v = \chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$
	$= \chi(\mathbf{r},\mathbf{r}') - \int_{\mathbf{r}''\in v}\int_{\mathbf{r}'''\in v} \chi(\mathbf{r},\mathbf{r}'')\chi_v^{-1}(\mathbf{r}'',\mathbf{r}''')\chi(\mathbf{r}''',\mathbf{r}')$
Mean density	$\langle\rho(\mathbf{r})\rangle_v = \rho_\ell n(\mathbf{r}) - \int_{\mathbf{r}'\in v} \chi(\mathbf{r},\mathbf{r}')c_1 - \int_{\mathbf{r}\in\bar{v}} \chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')\beta(u+\phi)(\mathbf{r}')$

Table 4.5: Approximate effective LCW Hamiltonian for slowly-varying field $n(\mathbf{r})$. The microscopic Hamiltonian $H[n(\mathbf{r}),\delta\rho(\mathbf{r})]$ is summarized in Table 4.3.

for cavities of up to a few Angstroms in size. This correspondence suggests that the basis-independent expression $2\pi\sigma_v$ is a reasonable replacement for $\det(2\pi\chi_v)$. We emphasize that the solution we have chosen is purely pragmatic, and cannot be rigorously justified. In fact, the replacement has some undesirable properties, such as the total solvation energy of two distant cavities being $k_{\rm B}T\ln\sqrt{2\pi} \approx 1\,k_{\rm B}T$ less than the sum of the solvation energies of the individual cavities. In most cases we have examined, however, the term $k_{\rm B}T\ln\sqrt{2\pi\sigma_v}$ contributes negligibly to the total solvation free energy. We have thus chosen to accept the inadequacies of our pragmatic replacement as a shortcoming of the otherwise rather useful Gaussian model.

4.2 Discretized LCW Hamiltonian

In order to implement a simulation with the Hamiltonian in Table 4.5, we must discretize it. We begin by describing $n(\mathbf{r})$ with reference to a cubic grid of spacing λ, depicted in Figure 4.1, and we denote its value at the center of cell i by n_i. Then, $n(\mathbf{r})$ is given by

$$n(\mathbf{r}) = \sum_i n_i \Psi(\mathbf{r} - \mathbf{r}_i), \tag{4.32}$$

where $\mathbf{r}_i$ is the center of cell i and n_i is 1 or 0, and the sum is over all cells i. The function $\Psi(\mathbf{r})$ is maximal with value 1 at $\mathbf{r} = \mathbf{0}$; it is cubic symmetric about the origin; and it is zero when the magnitude of any of the Cartesian components of $\mathbf{r}$ is greater than λ. Additionally, it satisfies for all $\mathbf{r}$ the property

$$\sum_i \Psi(\mathbf{r} - \mathbf{r}_i) = 1,$$

so that when all cells have $n_i = 1$, the resulting field $n(\mathbf{r})$ is uniformly 1. The cell size λ should be no smaller than the intrinsic width of the liquid-vapor interface. Based upon the interfacial profile of the SPC/E model [5,49], we therefore pick $\lambda = 4\,\text{Å}$. This is the minimal scale over which the time-averaged solvent density can transition from liquid-like to vapor-like values. The typical size of interfacial energies between cells on this grid is $\gamma\lambda^2$, where γ is the liquid-vapor surface tension of the solvent.

4.2.1 Interfacial Energetics

Implementing an LCW-like theory on a lattice has been attempted previously [106,107,123]. In these efforts, the term $H_L[n(\mathbf{r})]$ has been approximated by an Ising Hamiltonian with nearest-neighbor coupling

$$H_L[\{n_i\}] \overset{?}{\to} \gamma\lambda^2 \sum_{\langle ij\rangle} (n_i - n_j)^2 - \mu\rho_\ell\lambda^3 \sum_i n_i.$$

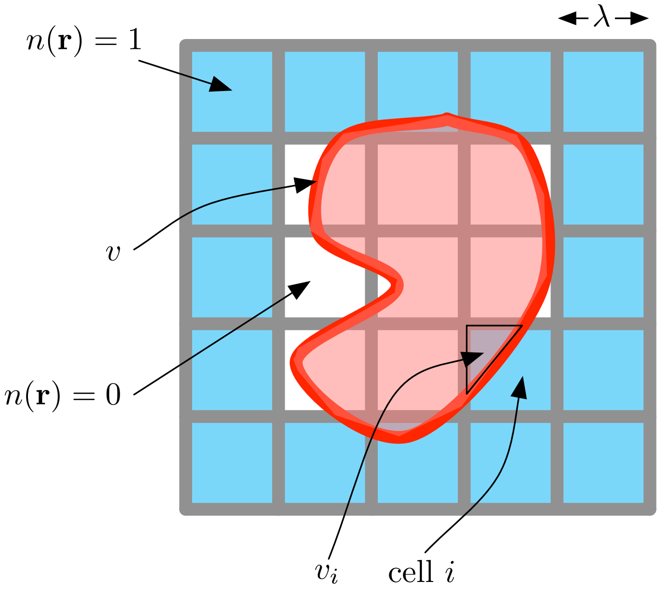


Figure 4.1: Schematic showing the solute and the large-length-scale density field on a grid

Unfortunately, the use of this Hamiltonian results in serious artifacts. Consider, for instance, the energetics of a convex vapor bubble embedded in the liquid, as represented by the field $\{n_i\}$. Many configurations of the field that are physically distinct nonetheless have equal surface area, and thus equal statistical weight. In particular, all convex shapes with equal projections onto the xy-, yz- and xz-planes have the same surface area. Hence, the use of this Hamiltonian results in an unphysical excess of entropy, as shown in detail in Section 4.4.4. Moreover, the energetic cost of common configurations of the field $\{n_i\}$ is substantially overestimated. The Ising Hamiltonian assigns a large vapor bubble of radius R an interfacial energy of about $6\pi\gamma R^2$, not $4\pi\gamma R^2$. Whereas using a renormalized γ, as in Ref. [78], can alleviate this latter problem, the problem of excess entropy is more fundamental.

Figure 4.2 illustrates the entropy problem in the Ising model. Here, a two-dimensional 27×27-cell Ising model is constructed. Filled cells ($n_i = 1$) are shown in black. The cells in the shape of a cross in the middle are constrained to be filled at all times, those at the edges to be empty at all times, and the system is initially prepared with all other cells empty. The system is then evolved using Metropolis dynamics. At temperature $T = 0$, the expectation is for the system to evolve to a state that minimizes the surface area (perimeter) of the filled region. What is instead observed is an uncontrolled meandering of the boundary between filled and unfilled cells. All the observed configurations after the initial expansion of the filled region have exactly the same energy, namely $\gamma\lambda^2 \times 4 \times 25$. At a temperature where $\gamma\lambda^2 = 5\,k_{\mathrm{B}}T$, one would expect capillary waves of appreciable amplitude to form along the boundary between filled and unfilled cells. Instead, the observed trajectory is essentially

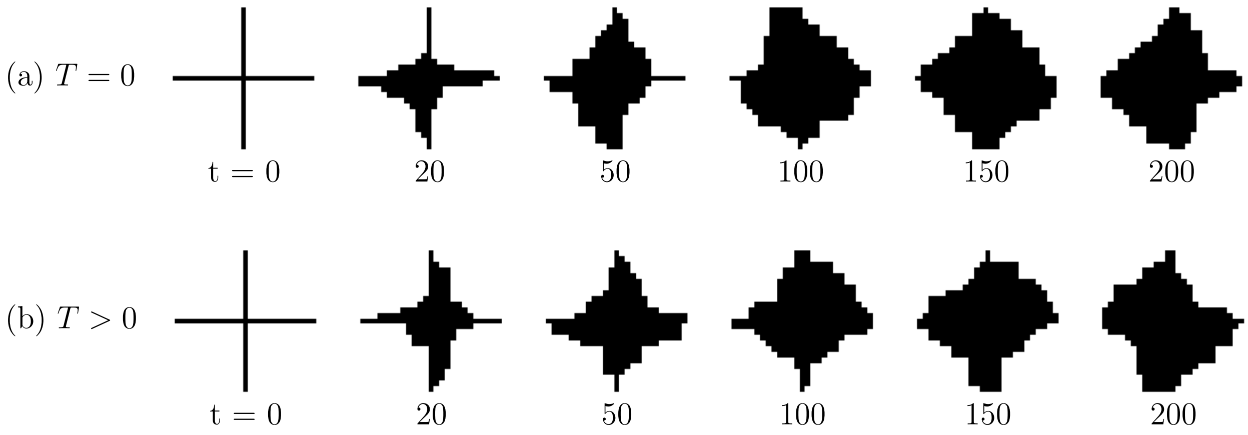


Figure 4.2: Excess entropy in the Ising model. The two-dimensional model has its central cells constrained to be filled ($n_i = 1$) and its edges to be empty ($n_i = 0$). The time t is measured in units of Metropolis sweeps through the lattice. (a) Model with $T = 0$. (b) Model with $\gamma L^2 = 5\, k_{\mathrm{B}} T$.

indistinguishable from the $T = 0$ trajectory.

Motivated by the above deficiencies of the Ising Hamiltonian, we have instead chosen to evaluate the Landau-Ginzburg integral in Equation (2.31) numerically using the representation of Equation (4.32) for $n(\mathbf{r})$. The value of the integral is conveniently expressed as the sum

$$H_L[\{n_i\}] = \gamma\lambda^2 \sum_i h_i - \mu\rho_\ell\lambda^3 \sum_i n_i, \tag{4.33}$$

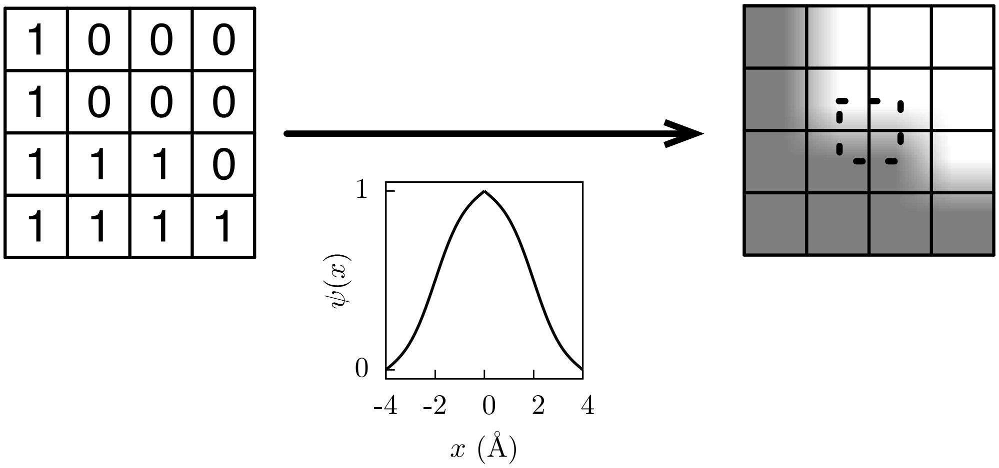
with the local integrals

$$h_i = \frac{1}{\gamma\lambda^2} \int_{x_i}^{x_i+\lambda} \mathrm{d}x \int_{y_i}^{y_i+\lambda} \mathrm{d}y \int_{z_i}^{z_i+\lambda} \mathrm{d}z \left[w(n(x,y,z),0) + \frac{m}{2} |\nabla n(x,y,z)|^2 \right]. \tag{4.34}$$

The quantity h_i depends only on the values of n_j for cells j that share one of the corners of cell i. There are only 14 distinct possible values of h_i, which can be precalculated numerically for a given free energy density $w(n,0)$ and cell size λ. To calculate h_i, we need concrete choices for $\Psi(\mathbf{r})$, $w(n,\mu)$ and m, which we describe below. The resulting values of h_i are given in Table 4.6.

Our choice for basis function $\Psi(\mathbf{r})$, depicted in Figure 4.3 for water, approximates at a local level the usual van der Waals construction (described in detail in Chapter 3 of Ref. [93]). We first construct a 1D basis function $\psi(x)$ satisfying

$$w'(\psi(x),0) - m\psi''(x) = 0, \tag{4.35}$$



1	0	0	0
1	0	0	0
1	1	1	0
1	1	1	1

Figure 4.3: Constructing $n(\mathbf{r})$ from $\{n_i\}$. The binary field specifies whether the density at the center of each lattice cell should be that of the liquid or that of the vapor. Between cell centers, the density is interpolated using the basis function $\psi(x)$ (whose form for water is shown below the arrow). The dashed lines delineate the domain of integration of the local free energy h_i given by Equation (4.34).

with boundary conditions $\psi(0) = 1$ and $\psi(\lambda) = 0$. We then extend the range of $\psi(x)$ and symmetrize it so that

$$\psi(x > \lambda) = 0, \tag{4.36}$$

and

$$\psi(x < 0) = \psi(-x). \tag{4.37}$$

Finally, the three-dimensional basis function $\Psi(\mathbf{r})$ is constructed from the one-dimensional profiles $\psi(x)$ to give

$$\Psi(x, y, z) = \psi(x)\psi(y)\psi(z).$$

The field $n(\mathbf{r})$ constructed from Equation (4.32) using this basis function has many useful properties: the value of $n(\mathbf{r})$ at the center of each cell i corresponds to the state encoded in n_i; the density interpolates smoothly between adjacent cells; and the density profile of a configuration representing an axis-aligned wall, where all n_i's are 1 on one side of a plane and 0 on the other, nearly reproduces the interface profile given by the van der Waals construction.

For water, following Ref. [48], we use the free-energy density

$$w(n, \mu) = \frac{2m}{d^2}(n-1)^2 n^2 - \mu\rho_\ell n, \tag{4.38}$$

where the parameter d captures the thickness of the liquid-vapor interface, as described shortly. This choice results in both sides of Equation (4.35) being proportional to m, so the function $\psi(x)$ is independent of m. In the free van der Waals theory, where the boundary conditions on Equation (4.35) are $\psi(-\infty) = 1$ and $\psi(+\infty) = 0$, the density profile $\psi_0(z)$ that results is

$$\psi_0(z) = [1 + \tanh(z/d)]/2,$$

which accurately describes the average density profile of an SPC/E water slab at ambient conditions. The thickness parameter d can thus be determined from simulation. A complication due to capillary waves is that d grows logarithmically with simulation box size [79,117], so different authors quote different values of d: 1.27 Å for a $19 \times 19\,\text{Å}^2$ interface in Ref. [48] and 1.54 Å for a $30 \times 30\,\text{Å}^2$ interface in Ref. [112]. We choose the smaller value because the instantaneous configuration of $n(\mathbf{r})$ should be blurred only by small-length-scale fluctuations, not by large-length-scale capillary waves, which correspond instead to different conformations of $n(\mathbf{r})$. The profile shown in Figure 4.3 corresponds to the solution of Equation (4.35) when the more restrictive boundary conditions described above are imposed, with $\lambda = 4\,\text{Å}$ and $d = 1.27\,\text{Å}$.

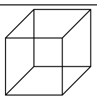
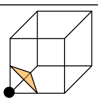
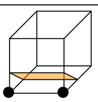
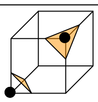
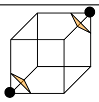
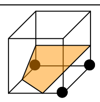
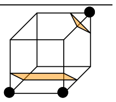
In principle, the value of m is related to the surface tension by the relation [93]

$$\gamma = \int_0^\lambda \Big[w(\psi(x), 0) + m\psi'(x)^2/2 \Big] \mathrm{d}x. \tag{4.39}$$

On a lattice, as exemplified above by the Ising Hamiltonian, this choice results in perfect interfacial energies for flat axis-aligned interfaces at the expense of more common curved interfaces. Thus, we instead choose m self-consistently such that $\psi(x)$ satisfies Equation (4.35) and the calculated interfacial energy of some reference geometry of surface area A is γA. Equation (4.39) corresponds to a cubic reference geometry. Since curved surfaces are far more common than flat one in realistic solutes, we instead use large spheres as our reference geometry.

For the specific form of $w(n, \mu)$ that we use for water, h_i is proportional to m and $\psi(x)$ is itself independent of m. The above self-consistent procedure can hence be implemented quite simply. We first calculate the h_i quantities up to a factor of m, and then pick m to obtain the correct interfacial energies.

With concrete choices of $\Psi(\mathbf{r})$, $w(n, 0)$, m and λ, the integrals h_i defined by Equation (4.34) can be evaluated. The value of h_i depends only on the values of n_j for the 8 cells j that share one of the corners of cell i. Out of the 256 possible configurations of $\{n_j\}$, only 14 are unique when one accounts for reflection, rotation and inversion symmetry. Thus, only 14 distinct integrals need to be evaluated numerically. This decomposition bears a strong resemblance to the marching cubes algorithm [65] that reconstructs interfaces in volumetric data, and is widely used in computerized tomography. The resulting values of h_i are given in Table 4.6.

Local $\{n_j\}$ configuration							
h_i (Present model)	0.000	0.387	0.676	0.725	0.754	0.851	0.965
h_i (Ising model)	0.000	0.750	1.000	1.500	1.500	1.250	1.750

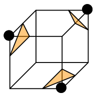
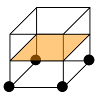
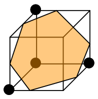
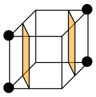
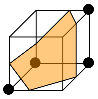
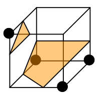
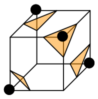
Local $\{n_j\}$ configuration							
h_i (Present model)	0.983	0.857	0.910	1.104	0.965	1.040	1.134
h_i (Ising model)	2.250	1.000	1.500	2.000	1.500	2.000	3.000

Table 4.6: Relative interfacial free energy h_i for each distinct neighboring cell configuration (diagrams after Ref. [65]). Highlighted corners denote cells j with $n_j = 1$, whereas the others refer to cells with $n_j = 0$; cell i is the lower-left corner in the back. To aid the eye, a schematic of the implied liquid-vapor interface of each configuration is shown in orange. The values of h_i are inversion-symmetric: interchanging highlighted and unhighlighted corners yields the same interface and interfacial energy. Also shown are the values of h_i that would reproduce the energetics of the standard Ising lattice gas, namely $\gamma\lambda^2 \sum_{\langle ij \rangle}(n_i - n_j)^2$.

To illustrate the effect of this improved Hamiltonian for the interfacial energetics, we consider the limit $d \to \infty$. There, the free energy density $w(n, \mu)$ is 0 for all n, and the above procedure can be implemented analytically. The resulting expression for $H_L[\{n_i\}]$ is

$$H_L[\{n_i\}] = \gamma L^2 \sum_i n_i \left[\frac{8}{3} - \frac{1}{6} \sum_j{}^{(2)} n_j - \frac{1}{12} \sum_j{}^{(3)} n_j \right] - \mu \rho_\ell \lambda^3 \sum_i n_i, \tag{4.40}$$

where both superscripted sums are over cells j that share at least one vertex with cell i: the sum $\sum^{(2)}$ is over the 12 cells that are two cells away from i, while the sum $\sum^{(3)}$ is over the 8 cells that are three cells away from i. Curiously, there is no nearest-neighbor coupling in this limit. What Equation (4.40) reveals about the above discretization of the Landau-Ginzburg Hamiltonian is that it augments the nearest-neighbor coupling of the Ising model with additional information about the local shape of the liquid-vapor interface.

In two dimensions, the analogous Hamiltonian in the $d \to \infty$ limit is

$$H_L[\{n_i\}] = \gamma L^2 \sum_i n_i \left[\frac{8}{3} - \frac{1}{3} \sum_j{}^{(1)} n_j - \frac{1}{3} \sum_j{}^{(2)} n_j \right] - \mu \rho_\ell \lambda^3 \sum_i n_i, \tag{4.41}$$

using notation analogous to that in Equation (4.40). Figure 4.4 shows how the system in Figure 4.4 behaves when the interfacial energetics are those of Equation (4.41). At zero temperature, the field quickly evolves to minimize the surface area (perimeter) between the

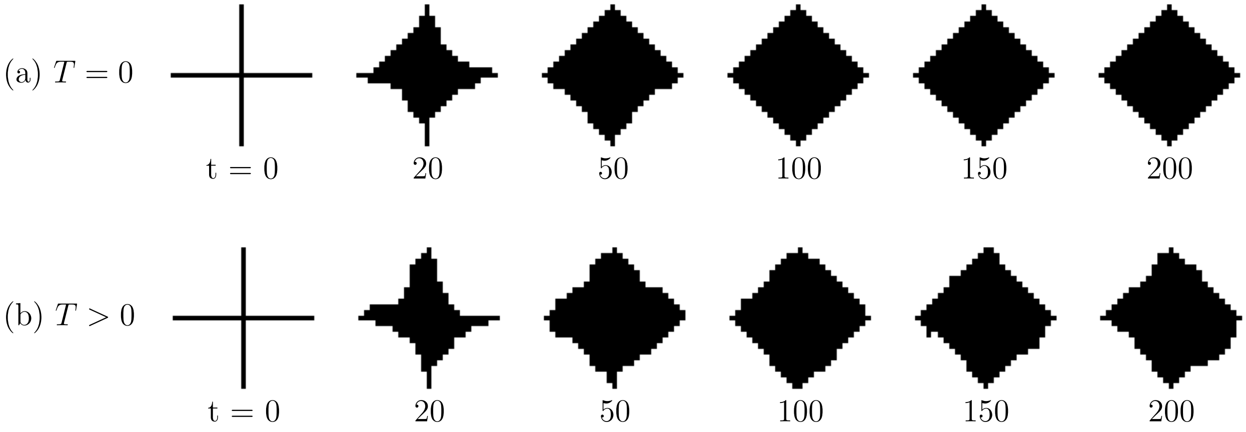


Figure 4.4: Discretized Landau-Ginzburg model of Equation (4.41). Labels as in Figure 4.2.

filled and unfilled regions, while at a moderate temperature, capillary waves of controlled magnitude appear on this interface. The energy of the stable field configuration at $T = 0$ is also quite close to its continuum value of $\gamma L^2 \times 4 \times 25/\sqrt{2}$.

4.2.2 Solvent exclusion

To construct a discretized version of $H_S[n(\mathbf{r})]$ in Table 4.5, we would ideally simply use Equation (4.32) to express $H_S[n(\mathbf{r})]$ in terms of the $\{n_i\}$ variables. The resulting equations are somewhat unwieldy computationally, so we instead proceed heuristically. We first need a discretized analog of the fluctuation spectrum $\chi(\mathbf{r}, \mathbf{r}')$. To define it, we introduce the projector $b_V(\mathbf{r})$ for a given volume V,

$$b_V(\mathbf{r}) = \begin{cases} 1, & \mathbf{r} \in V, \\ 0, & \text{otherwise.} \end{cases} \tag{4.42}$$

The complement of V is denoted by $\bar{V}$, so that $b_V(\mathbf{r}) + b_{\bar{V}}(\mathbf{r}) = 1$. We denote the overlap of v or $\bar{v}$ with cell i by v_i and $\bar{v}_i$, respectively. In the uniform fluid, the solvent number fluctuation correlations between the portions of cells i and j that overlap with two volumes V and V', respectively, form the elements of a matrix

$$\chi_{ij}(V, V') = \int_{\mathbf{r}\in i} \int_{\mathbf{r}'\in j} b_V(\mathbf{r}) \chi_0(\mathbf{r} - \mathbf{r}') b_{V'}(\mathbf{r}'), \tag{4.43}$$

where

$$\chi_0(\mathbf{r}) = \rho_\ell \delta(\mathbf{r}) + \rho_\ell^2 [g(|\mathbf{r}|) - 1]. \tag{4.44}$$

The domain of the $\mathbf{r}$ and $\mathbf{r}'$ integrals are restricted to the volume of cells i and j, as indicated. A way to estimate these elements is outlined in Section 4.4.1.

With the above notation, we can approximate the quantities $\langle N \rangle_v$ and σ_v that are referred to in $H_S[n(\mathbf{r})]$ by the expressions

$$\langle N \rangle_v = \sum_i n_i \rho_\ell v_i, \tag{4.45}$$

$$\sigma_v = \sum_{i,j} n_i \chi_{ij}(v,v) n_j. \tag{4.46}$$

4.2.3 Length-scale coupling

The integral defining $H_I[n(\mathbf{r})]$ in Table 4.5 can be approximated as a lattice sum,

$$H_I[\{n_i\}] = -K \sum_i \phi_i \rho_\ell n_i v_i, \tag{4.47}$$

where ϕ_i is the average value of $\phi(\mathbf{r})$ in cell i. To proceed, we need to choose a concrete implementation of the overbar operation that is used to define $\phi(\mathbf{r})$. Following Ref. [107], we approximate it as a weighted average involving the cell containing $\mathbf{r}$ and its nearest neighbors,[2] given by

$$\overline{n(\mathbf{r})} \approx \left[\frac{1}{2} n_i + \frac{1}{12} \sum_{j\,(\mathrm{nn}i)}' n_j\right]. \tag{4.48}$$

The primed sum over $j(\mathrm{nn}i)$ is a sum over the six cells j that are nearest neighbors to cell i. The quantity ϕ_i then follows directly,

$$\phi_i = 2a\rho_\ell \left[1 - \frac{1}{2} n_i - \frac{1}{12} \sum_{j\,(\mathrm{nn}i)}' n_j\right]. \tag{4.49}$$

4.2.4 Solute-solvent interactions

The remaining term in the effective Hamiltonian in Table 4.5, $H_u[n(\mathbf{r})]$, captures the energetic effect of a field $u(\mathbf{r})$ that couples to $\rho(\mathbf{r})$. This potential can be used to model the attractive interactions between a solute and the solvent.

To discretize $H_u[n(\mathbf{r})]$, we define a discretized analog u_i of $u(\mathbf{r})$,

$$u_i = \frac{1}{\bar{v}_i} \int_{\mathbf{r}\in i} b_{\bar{v}}(\mathbf{r}) u(\mathbf{r}). \tag{4.50}$$

---

[2] In Ref. [107], the term proportional to n_i is omitted. Since $\phi(\mathbf{r})$ only acts on cells with $n_i = 1$, this omission is inconsequential, and shows up as an extra factor of 2 in their value of a.

Notice that u_i is independent of $u(\mathbf{r})$ for values of $\mathbf{r}$ inside the solute. The apparent divergence, where v completely overlaps cell i, has no effect in the final expression. The natural discretization of $H_u[n(\mathbf{r})]$ is then

$$H_u[\{n_i\}] = \sum_i u_i n_i \Big[\rho_\ell \bar{v}_i - \sum_j \chi_{ij}(\bar{v}, v) n_j \langle N \rangle_v / \sigma_v - \sum_j \chi_{ij}(\bar{v}, \bar{v}) n_j \beta(u_j + \phi_j)\Big] + \frac{k_\mathrm{B}T}{2} \sum_{i,j} \beta u_i n_i \chi_{ij}(\bar{v}, \bar{v}) n_j \beta u_j. \quad (4.51)$$

Note the use of the volume $\bar{v}$ in the matrix elements χ_{ij}. In this expression, we have approximated $\chi^{(\mathrm{m})}(\mathbf{r}, \mathbf{r}')$ as equal to $\chi(\mathbf{r}, \mathbf{r}')$ if both $\mathbf{r}$ and $\mathbf{r}'$ are in v, and 0 otherwise.

The complete discretized version of the effective LCW Hamiltonian is summarized in Table 4.7. Although the derivation of the model is somewhat tortuous (and torturous), the model itself is quite simple.

4.2.5 Parameters of the Hamiltonian

The model of Table 4.7 has a number of parameters that specialize it to water at ambient conditions, $T = 298\,\mathrm{K}$ and 1 atm pressure, p. We have discussed our choices for cell size $\lambda = 4\,\text{Å}$ and interfacial relative free energies h_i in Section 4.2.1. In this section, we discuss appropriate choices for the remaining parameters. Further, we comment upon what changes are required for applications at different states of water.

The bulk liquid density ρ_ℓ is the experimental value [63], whereby a liquid cell contains $\rho_\ell \lambda^3 \approx 2.13$ waters on average. At ambient conditions, the experimental value for the surface tension [63] yields $\gamma\lambda^2 \approx 2.80\,k_\mathrm{B}T$. The relative chemical potential is given by $\mu \approx (p - p_\mathrm{vap})/\rho_\ell$, where p_vap is the vapor pressure at 298 K. This relationship gives $\mu \approx 7.16 \times 10^{-4}\,k_\mathrm{B}T$, which is quite small, reflecting that water at ambient conditions is nearly at coexistence with its vapor.

The matrix elements $\chi_{ij}(V, V')$ are computed from the radial distribution function, $g(r)$, and we derive this function from Narten and Levy's tabulated data [81]. It is a convenient data set because it covers a broad range of temperatures for the liquid at and near $p = 1$ atm. At one temperature, 25°C, we have checked that a different estimate of the radial distribution function, that of the SPC/E model, yields similar matrix elements, and the resulting solvation properties are essentially identical to those obtained when the $\chi_{ij}(V, V')$'s are computed from the Narten-Levy data at the same temperature (see Figure 4.12).

The only parameters that we estimate through fitting are the strength a of the unbalancing potential and the renormalization constant K. In the absence of solute-solvent attractions, only the product of a and K is relevant. Values of a and K with $Ka\rho_\ell = 2.1\,k_\mathrm{B}T$ allow us to match the solvation free energies of hard spheres in SPC/E water (see below). By comparing the average value of the computationally convenient approximate expression involving ϕ_i in Equation (4.47) with that of its complete and unrenormalized counterpart,

Effective Hamiltonian	$H_{\text{eff}}[\{n_i\}] = H_L[\{n_i\}] + H_S[\{n_i\}] + H_I[\{n_i\}] + H_u[\{n_i\}]$
Interfacial energy	$H_L[\{n_i\}] = \gamma\lambda^2 \sum_i h_i - \mu\rho_\ell\lambda^3 \sum_i n_i,$
Solvent exclusion energy	$H_S[n(\mathbf{r})] \approx k_\mathrm{B}T[\langle N\rangle_v^2 /2\sigma_v + C/2],$
	$C = \begin{cases} \ln(2\pi\sigma_v), & \langle N\rangle_v > 1, \\ \max[\ln(2\pi\sigma_v), \langle N\rangle_v], & \text{otherwise.} \end{cases}$
	$\langle N\rangle_v = \sum_i n_i\rho_\ell v_i$
	$\sigma_v = \sum_{i,j} n_i\chi_{ij}(v,v)n_j$
Unbalancing potential	$\phi_i = 2a\rho_\ell\Big[1 - \frac{1}{2}n_i - \frac{1}{12}{\sum_{j\,(\text{nn}i)}}' n_j\Big]$
Length-scale coupling	$H_I[\{n_i\}] = -K\sum_i \phi_i\rho_\ell n_i v_i$
External field	$u_i = \frac{1}{\bar{v}_i}\int_{\mathbf{r}\in i} b_{\bar{v}}(\mathbf{r})u(\mathbf{r})$
External coupling	$H_u[\{n_i\}] = \sum_i u_i n_i\Big[\rho_\ell\bar{v}_i - \sum_j \chi_{ij}(\bar{v},v)n_j\langle N\rangle_v/\sigma_v - \sum_j \chi_{ij}(\bar{v},\bar{v})n_j\beta(u_j+\phi_j)\Big] + \frac{k_\mathrm{B}T}{2}\sum_{i,j}\beta u_i n_i\chi_{ij}(\bar{v},\bar{v})n_j\beta u_j$

Table 4.7: Discretized effective LCW Hamiltonian. The values of h_i are those in Table 4.6. The matrix elements $\chi_{ij}(V,V')$ are given by Equation (4.43). The primed sum over $j(\text{nn}i)$ is a sum over the six cells j that are nearest neighbors to cell i.

as is done in Section 4.4.3, we find that K is about 1/2, so that $a\rho_\ell \approx 4.2\,k_\mathrm{B}T$. This value for a is close to the original LCW estimate [68], arrived at from a different criterion.

These values are applicable at ambient conditions. As temperature and pressure vary, only γ, μ and $g(r)$ vary appreciably, while K varies slightly. In particular, surface tension decreases roughly linearly with temperature [63] (with $\mathrm{d}\gamma/\mathrm{d}T \approx -0.15\,\mathrm{mJ/m^2{\cdot}K}$ which is $-5.8\times10^{-3}\,k_\mathrm{B}T/\lambda^2{\cdot}\mathrm{K}$ at $T = 298\,\mathrm{K}$). As noted above, μ increases roughly linearly with pressure. The pair correlation function $g(r)$ loses some structure for temperatures above 50 °C. The terms that are modeled by the renormalization constant K reflect the degree to which solvent density layers next to a solute. Since this layering reflects the structure of $g(r)$, we expect K to be slightly state-dependent, with its value increasing with temperature.

Conversely, liquid water has a nearly constant density and bulk correlation length at the temperatures and pressures where our model would be useful, so ρ_ℓ and λ can be taken as constant as well. The value of a in simple liquids (Equation (A.3)) is state-independent, so we expect that in water, a will be nearly state-independent as well [47].

4.3 Applications and Results

4.3.1 Solvation Free Energies

To test our model's ability to capture the length-scale dependence of solvation, and to parametrize the strength of the unbalancing potential, we have calculated the solvation free energy of hard spheres of different radii. Whether within our model or using explicit water simulations, we calculate the solvation free energy of a solute following the guidelines of Ref. [90]. Briefly, we first define a series of $M+1$ solutes S_0 through S_M that slowly interpolate from an empty system (S_0) to the final solute of interest (S_M). We then sequentially calculate the free energy difference between solute m and solute $m+1$ using the Bennett acceptance ratio estimator (BAR [4]), and, where necessary, the linear interpolation stratification procedure of Ref. [90]. Error estimates are calculated using BAR, and are generally smaller than 0.5%.

Our model (Table 4.7) involves only simple arithmetic, so free energies can be calculated with little computational effort. For example, calculating the solvation free energy of hard spheres of up to 14 Å in radius in increments of 0.5 Å (Figure 4.5) takes about 1 hour on a single 2 GHz machine with a code that has not been fully optimized, whereas a similar calculation in explicit SPC/E waters with GROMACS [44] would take around 600 hours on the same machine to obtain a similar statistical accuracy.

As discussed in Section 2.2.1, hard-sphere solvation free energies scale as solute volume for small spheres, and as surface area for large spheres, with a smooth crossover at intermediate sizes. Figure 4.5 illustrates this behavior and compares the results of our model to previous simulation results using SPC/E water [49]. As the model manifestly reproduces the small- and large-length-scale limits, the most significant feature illustrated in Figure 4.5

is the *gradual* crossover from volume to surface area scaling. Ignoring the unbalancing potential leads to a qualitatively correct scaling behavior. However, adjustment of the single parameter a, which determines the strength of the unbalancing potential, yields a near-exact agreement between our model and the SPC/E results for all sphere sizes. In all subsequent results, the parameter a is fixed at this value.

The model results have small lattice artifacts—results that depend upon the position of the solute relative to that of the coarse-grained lattice—as shown in the inset of Figure 4.5. When studying stationary solutes, lattice artifacts may be mitigated by performing multiple calculations, differing only by small displacements of the solutes, and then averaging the results. When studying dynamical phenomena, lattice artifacts tend to pin solutes into alignment with the coarse-grained lattice. For arbitrary molecular solutes, we expect that pinning forces acting on one portion of the molecule will generically oppose pinning forces on other parts of the molecule, so that the total pinning forces will largely cancel out. However, when treating many identical molecules, lattice artifacts can add constructively, and additional steps are needed to mitigate them [103].

Since the unbalancing potential is explicitly parametrized with the solvation free energy of hard spheres, it is useful to evaluate the accuracy of the results in other geometries. To this effect we computed the solvation free energies of a family of hexagonal plates, consisting of 37 methane-like oily sites arranged into three concentric rings. We control the size of these plates, depicted in Figure 4.6, by varying the distance d between neighboring oily sites. For our calculations with explicit SPC/E water, the sites are uncharged and interact with the solvent molecules via a standard[3] water-methane Lennard-Jones potential. To study the role of attractive interactions, we split this Lennard-Jones potential using the Weeks-Chandler-Andersen (WCA) prescription [119] into a repulsive part $u_0(r)$ and an attractive part $\Delta u(r)$. The magnitude of the attractive tail can be varied systematically with a scaling parameter η, such that

$$u_\eta(r) = u_0(r) + \eta \Delta u(r). \tag{4.52}$$

For the ideal hydrophobic plate, we set η to zero.

In the coarse-grained model, the repulsive core of the solute is represented as an excluded volume. To construct it, we replace each solute particle by a thermally-equivalent hard sphere, whose radius R_0 is estimated according to

$$R_0 = \int_0^\infty \mathrm{d}r\, [1 - \exp(-\beta u_0(r))],$$

which is a first approximation to the WCA value of this radius (Section 5.3 of Ref. [41], also see Ref. [113]) and is essentially the radius at which $u_0(r)$ is $k_\mathrm{B}T$. The excluded volume is then the union of the hard-sphere volumes of each solute site.

---

[3] The parameters of the solute-solute Lennard-Jones potential are those of Ref. [58]: $\sigma = 3.905$ Å and $\epsilon = 0.118$ kcal/mol. Lorentz-Berthelot mixing rules were used to obtain the water-solute interaction parameters. These are the same oily sites used in the model hydrophobic plate of Chapter 3

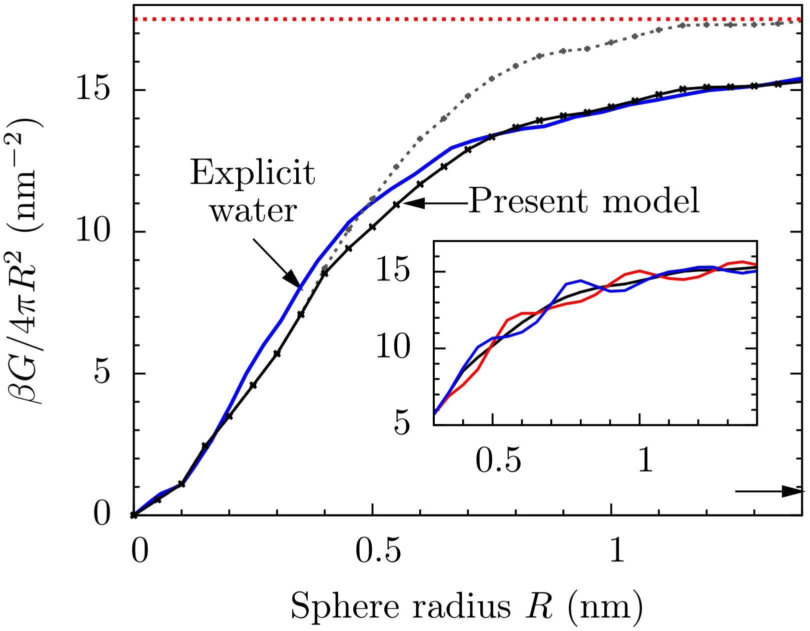


Figure 4.5: Solvation free energies G of hard spheres of increasing radii, as calculated from explicit SPC/E water simulations [49] (solid blue), from the coarse-grained model (solid black), and from one of the most common variants of GBSA [14] (arrow at bottom right). Different parametrizations of GBSA yield non-polar solvation energies that may differ by as much as a factor of ten, reflecting the crudeness of the SA portion of the model [14]. When the coarse-grained model has no unbalancing potential ($a = 0$, dashed gray), the intermediate-size regime is only qualitatively reproduced. For large spheres, the ratio of G to surface area tends to the liquid-vapor surface tension γ (horizontal red dots). Inset: Illustration of lattice artifacts. The spheres are centered at different offsets from the lattice: a generic position (0.98 Å, 0.79 Å, 1.89 Å) that breaks all rotational and mirror symmetries (black), a lattice cell corner (blue) and a lattice cell center (red). All three curves are identical for $R \leq 0.35$ nm.

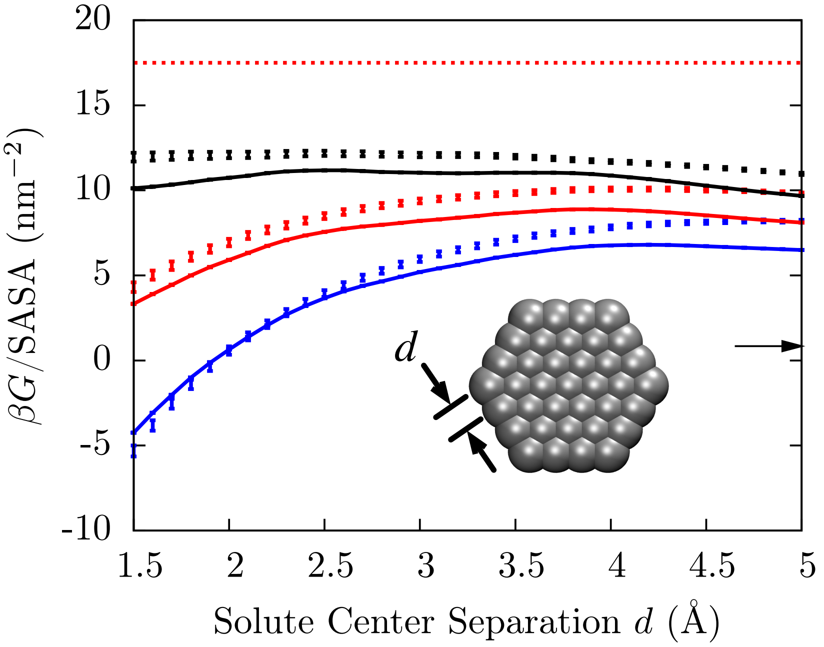


Figure 4.6: Solvation free energies G of hexagonal plates, as a function of plate size, as calculated by the coarse-grained model (solid lines), by explicit SPC/E water simulations (points), and by the same GBSA variant as in Figure 4.5 (arrow on right). Three values of the attractive interaction strength η are shown: 0.0 (black), 0.5 (red) and 1.0 (blue). Solvent-accessible surface areas (SASAs) were calculated using VMD [55], with a particle radius of 1.97 Å, a solvent radius of 1.4 Å, and 1,000,000 samples per atom. The bulk liquid-vapor surface tension of water (horizontal red dots) is shown. Inset: Detail of the hexagonal plate. The solvent-excluded volume of each oily site is a sphere of radius $R_0 = 3.37$ Å.

Figure 4.6 compares the solvation free energies for this family of solute plates computed from our atomistic simulations with those computed from the coarse-grained model with the unbalancing parameter a determined above for solvated hard spheres. Now, with this different geometry, the coarse-grained model continues to perform well. The discrepancies are primarily due to the small underestimation, shown in Figure 4.5, of the solvation free energy of small spheres. Figure 4.6 also compares the solvation free energies of plates with increasing attractions to the corresponding results from explicit-water simulations.

4.3.2 Fluctuations

In the present model, we estimate water number distributions, $P_V(N)$, by a two-step procedure. For any given solvent configuration $\{n_i\}$, the small-length-scale fluctuations of $\delta\rho(\mathbf{r})$ give rise to a Gaussian distribution in the numbers of waters, so that

$$P_V(N|\{n_i\}) \propto \exp\bigl[-(N - \langle N\rangle_V)^2/2\sigma_V\bigr], \tag{4.53}$$

where, assuming V and v are disjoint,

$$\sigma_V = \int_{\mathbf{r}\in V}\int_{\mathbf{r}'\in V} \chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}') \approx \sum_{ij} n_i \chi_{ij}(V,V) n_j, \tag{4.54}$$

and the mean number of solvent particles in the volume V, $\langle N\rangle_V$, is computed by a discrete approximation of the integral of $\langle\rho(\mathbf{r})\rangle_v$ in Table 4.5 over the volume V. That computation gives

$$\langle N\rangle_V = \sum_i n_i \Bigl[\rho_\ell V_i - \sum_j \chi_{ij}(V,v) n_j \langle N\rangle_v/\sigma_v - \sum_j \chi_{ij}(V,\bar{v}) n_j \beta(u_j + \phi_j)\Bigr]. \tag{4.55}$$

Here, V_i is the overlap of the probe volume with cell i. Notice the use of the probe volume V in the χ_{ij} matrices. Formally, we then thermally average the above result over all possible solvent configurations to obtain

$$P_V(N) \propto \sum_{\{n_i\}} P_V(N|\{n_i\}) \exp(-\beta H_{\mathrm{eff}}[\{n_i\}]).$$

In practice, we estimate this sum by sampling a lattice variable n that closely correlates with N, given by

$$n = \sum_{i\in V} n_i.$$

This procedure is analogous to obtaining the atomistic $P_V(N)$ by biasing the closely related observable $\tilde{N}(\mathbf{R}^N)$, as described in Section 3.2.1. More precisely, we divide the range of possible values of n into small overlapping windows, and sample relevant configurations at every value of n using Wang-Landau sampling [116] along n, together with replica exchange [26], to obtain good sampling and avoid kinetic traps. We then used MBAR (Section 3.1.1) to reconstruct from these runs the probability distribution for n, $P(n)$. During the umbrella sampling runs, lattice configurations with equal n are observed in proportion to their Boltzmann weight. Using the notation $\{n_i\} \in n$ to denote all observed lattice gas configurations with a particular value of n, we finally obtain

$$P_V(N) = \sum_n P(n) \sum_{\{n_i\} \in n} P_V(N|\{n_i\}). \tag{4.56}$$

To estimate the statistical errors in our procedure, we calculate $P_V(N)$ in five independent Monte Carlo runs, and estimate the standard error in the mean of $\ln P_V(N)$.

In the absence of a solute, $P_V(N)$ is sensitive only to the interfacial energetics of the lattice gas. Figure 4.7 compares the $P_V(N)$ curve obtained using the present model for a $12 \times 12 \times 12\,\text{Å}^3$ volume with the atomistic results of Figure 3.3 and with (a) a version of the coarse-grained model that lacks an unbalancing potential (a is set to zero), and (b) a version that additionally uses the naive Ising lattice gas for estimating interfacial energetics in $n(\mathbf{r})$. Our present model captures the observed deviations from Gaussian behavior better than these simpler models, which reflects its higher accuracy in estimating interfacial energetics and microscopic curvature effects.

We next evaluate our model's ability to capture how hydrophobic solutes affect water number fluctuations in nearby probe volumes. We use the same model hydrophobic plate and probe volume that was used in Section 3.2.3. Figure 4.8 shows a set of water number distributions in the probe volume adjacent to the plate, as calculated with the present model, which compare favorably to the atomistic results in Figure 3.6. With no solute-solvent attractive interactions, the probability computed from the lattice model has a clear fat tail towards lower numbers of waters in the probe volume. At higher attractive interactions, this fat tail is correspondingly depressed, but not entirely suppressed.

Figure 4.8 also evidences some of the limitations of the present model. The probe volume being less than one lattice cell thick, large lattice artifacts are inevitable. Moreover, since $P_V(N)$ distributions are much more detailed probes of solvent structure than solvation free energies, we expect more room for disagreement with simulation. Nevertheless, we emphasize that, by construction, no implicit solvation models can capture the above effects on solvent structure, which underlie the pathways of hydrophobic assembly. Other coarse-grained solvation models (for example, see Ref. [95]), on the other hand, *can* probe rare solvent fluctuations, and it would be useful to evaluate their accuracy in this respect as compared to explicit-water models and the present lattice model.

4.3.3 Confinement

To conclude, we evaluate our model's ability to reproduce the dewetting phenomenology observed in Section 3.2.4. There, we placed two model hydrophobic plates of varying attractive

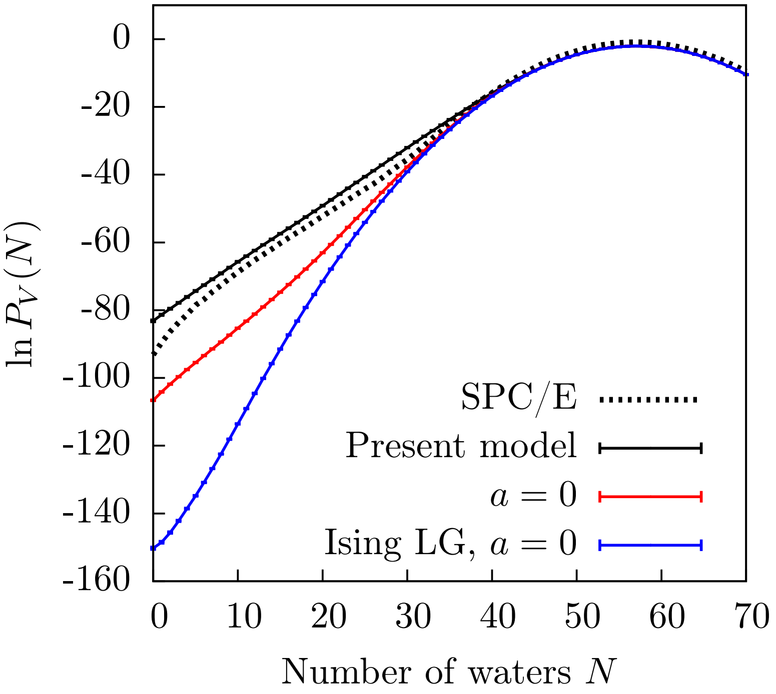


Figure 4.7: Water number distribution in a $12 \times 12 \times 12\,\text{Å}^3$, as obtained using explicit SPC/E water simulations (Figure 3.3), the present model, the present model without the unbalancing potential ($a = 0$), and a model with an Ising Lattice Gas and no unbalancing potential.

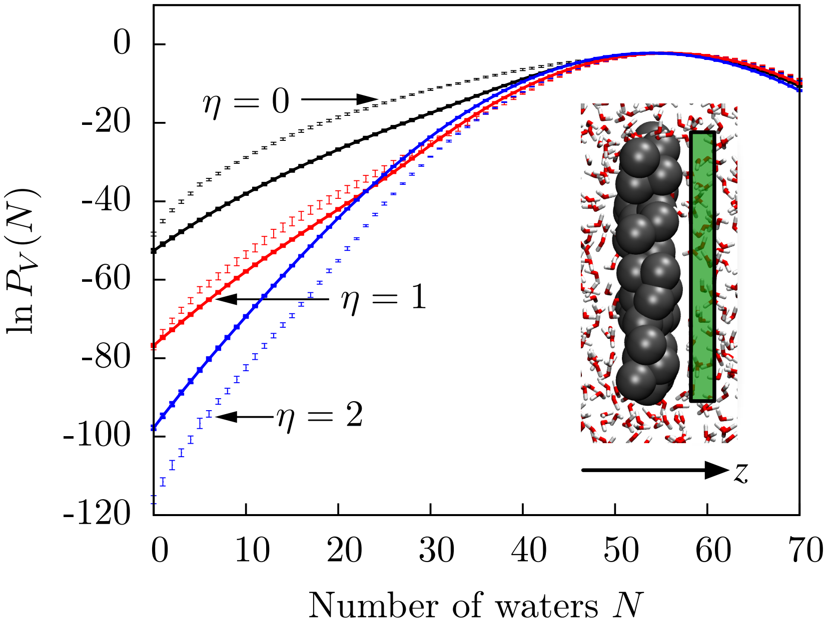


Figure 4.8: Water number distributions in a probe volume of size $24 \times 24 \times 3$ Å^3 immediately adjacent to the model plate solute (inset) with varying attractive strength η, in the coarse-grained model (solid lines) and in explicit SPC/E water (points). See Section 3.2.3 for setup details.

strength η at a distance d from each other, and looked at the water number distributions in the inter-plate volume. Figure 4.9, analogous to Figure 3.8, shows the resulting water number distributions. Using our model, we can further densely sample the space of the variables d and η to obtain more accurate information about the dewetting transition. Figure 4.10 summarizes the results in the form of a phase diagram. At small separations and low attractive strengths, the dry state (low N) is most stable, whereas high attractive strengths and large separations stabilize the wet state (high N).

The general, though not quantitative, agreement between the coarse-grained model and the SPC/E data is very encouraging: bistability is observed in the $P_V(N)$ distributions in both cases, with the barriers at the nearly equal values of N, and with barrier heights that track the SPC/E barrier heights. The phase boundary in Figure 4.10 closely tracks the phase boundary observed in explicit water, with a shift of less than 2 Å for all η. Moreover, as shown in Figure 4.11, once the general shift in the phase boundaries is accounted for, the $P_V(N)$ distributions for systems near that boundary obtained by the coarse-grained model and the SPC/E simulations agree reasonably well. Hence, the present model is better suited than implicit solvation models for studies of nanoscale self-assembly or protein-protein interactions driven by the hydrophobic effect.

4.4 Appendices

4.4.1 Estimating $\chi_{ij}(V, V')$

An essential ingredient of the model we present is the matrix $\chi_{ij}(V, V')$, given by the integral in Equation (4.43). The terms involving the delta-functions of Equation (4.44) are trivial. Owing to the rapid oscillations in $g(r)-1$, the remaining integrals are harder to estimate. We employ a two-step procedure to estimate these integrals efficiently. We begin by subdividing the $\lambda = 4$ Å-resolution grid of cells into a much finer grid of resolution $\lambda_f = 1$ Å. For clarity, below we explicitly distinguish between cells in the *coarse* grid, indexed by the letters i and j, and cells in the *fine* grid, indexed by the letters a and b. We evaluate the integrals of the non-delta-function portion of χ_0 on the fine grid without otherwise restricting the arguments to particular volumes V and V', and denote the result by χ_{ab}. Each fine cell is so small that the effect of a restriction on the integration domain can be estimated accurately with a simple interpolation formula. We then use these interpolated values in the fine grid to build up the elements of $\chi_{ij}(V, V')$ over the coarse grid.

To evaluate χ_{ab}, we use the Narten-Levy data [81] for the structure factor $S(k)$ of water. Since the $S(k)$ is unavailable for wave-numbers k higher than $16\,\text{Å}^{-1}$, we blur the domains of integration over a range of about $2\pi/16$ Å, which makes the values of the integrals practically insensitive to this missing data. Concretely, we introduce a basis function Φ, given by

$$\Phi(x, y, z) = \varphi(x)\varphi(y)\varphi(z),$$

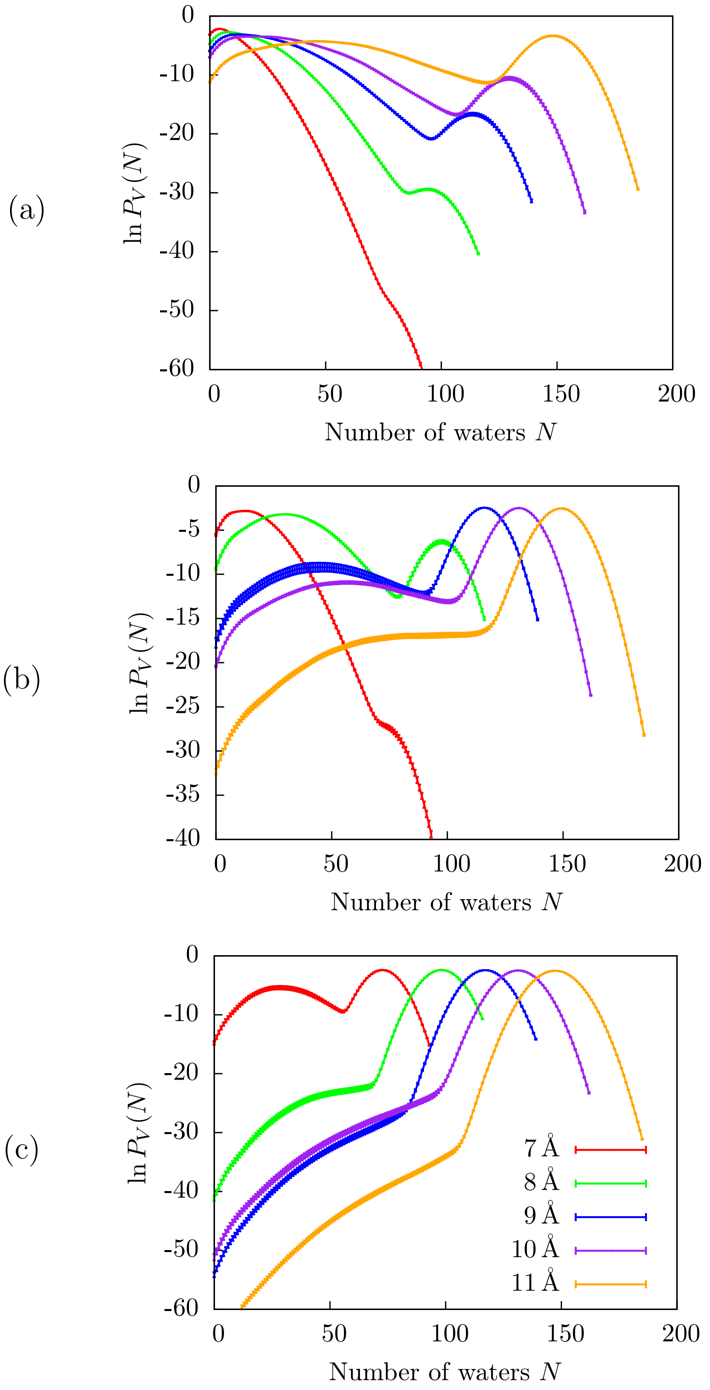


Figure 4.9: $P_V(N)$ distributions for interplate region in Figure 3.7 for η equal to (a) 0.0, (b) 0.5 and (c) 1.0. The plate separation runs from $d = 7$ Å (red) to $d = 11$ Å (orange).

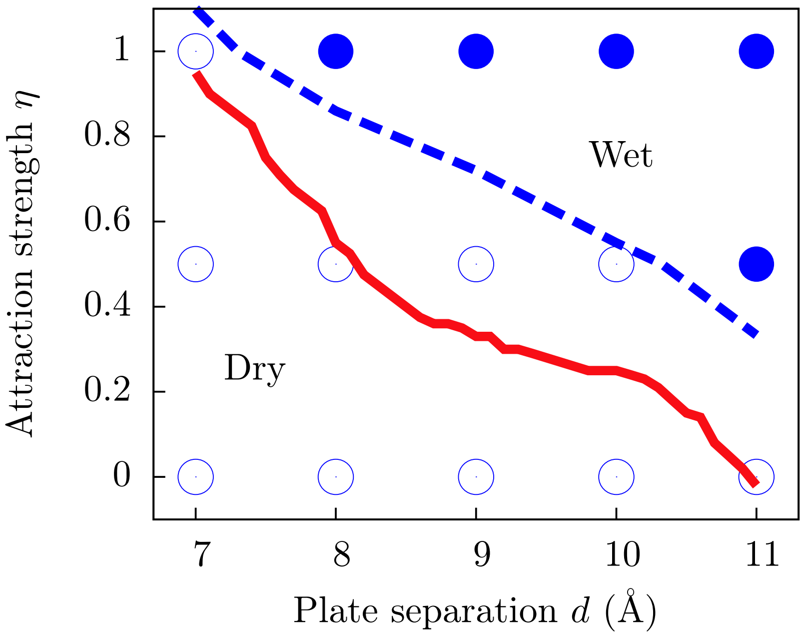


Figure 4.10: Phase diagram for the interplate region of the system depicted in Figure 3.7. For the explicit SPC/E water simulations (blue), each symbol corresponds to an individual $P_V(N)$ distribution that we have calculated, shown in Figure 3.8 (filled: wet state stable; open: dry state stable). The phase boundary (blue dashes) is estimated from a linear interpolation of the relative stability of the wet and dry states. The relative stability is determined from the relative depths of the basins in $-\ln P_V(N)$ The phase boundary for the present model (red solid line) was estimated from a dense sampling of $P_V(N)$ distributions, and is accurate to ±0.1 Å in d and ±0.1 in η.

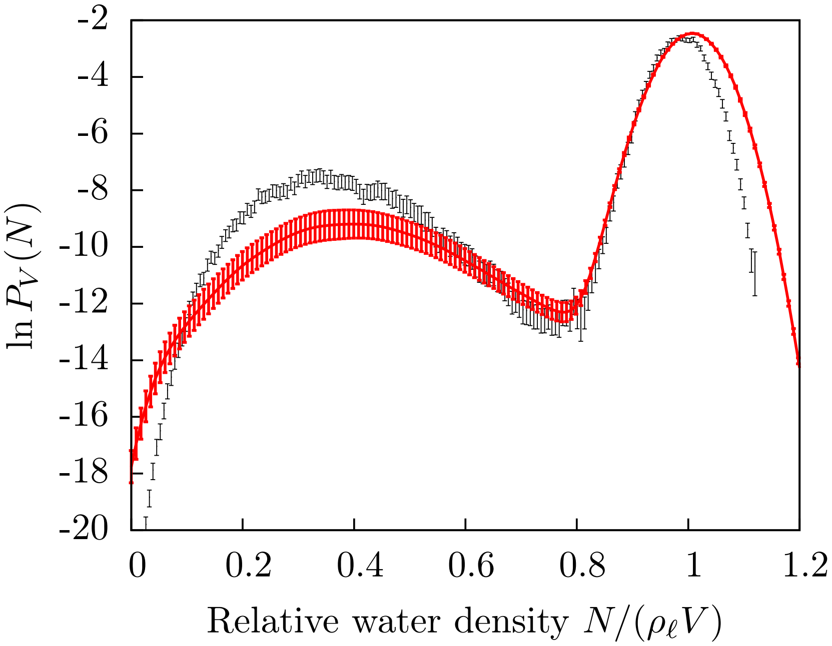


Figure 4.11: Water density distribution of confined water 1 Å from coexistence. These distributions are for the system depicted in Figure 3.7 when $\eta = 0.5$. Coexistence lines are shown in Figure 4.10. The explicit water simulation data (black) corresponds to $d = 11$ Å, while the coarse-grained model (red) results correspond to $d = 9$ Å.

with

$$\varphi(x) = \frac{1}{2}\left[\tanh\frac{x-\lambda_f/2}{\Delta} - \tanh\frac{x+\lambda_f/2}{\Delta}\right].$$

The function φ is unity around $x = 0$, and goes rapidly to zero as $|x| \gtrsim \lambda_f/2$, with Δ controlling the range of x over which this transition occurs. We have found a value of 0.1 Å for Δ to be adequate. Using the notation $\mathbf{r}_a$ to denote the center of fine cell a, the value of χ_{ab} is given by

$$\chi_{ab} = \rho_\ell^2 \int_{\mathbf{r}}\int_{\mathbf{r}'} \Phi(\mathbf{r}-\mathbf{r}_a)[g(|\mathbf{r}-\mathbf{r}'|)-1]\Phi(\mathbf{r}'-\mathbf{r}_b). \tag{4.57}$$

The integral is best evaluated in Fourier space, where the term in square brackets appears as the experimental $S(k)$ profile. We overcome the convergence problems of a rapidly oscillating integrand by using the Haselgrove-Conroy integration algorithm [15,42]. To properly account for $g(r)$ being exactly zero for $r \lesssim 2.35$ Å, we further set χ_{ab} to exactly $-\rho_\ell^2$ if all points in a are within $r_c = 2.35$ Å from all points in b. To limit the range of χ_{ab}, we also set it to zero if all points in a are more than 10 Å from all point in b. The values of χ_{ab} need only be calculated once at each state point of water, and we have spent considerable effort in compiling them at ambient conditions. Our results are tabulated in the Supplementary Information of Ref. [111].

For specific volumes V and V', we estimate the value of $\chi_{ij}(V, V')$ as a weighted average of the pertinent values of χ_{ab},

$$\chi_{ij}(V, V') \approx \rho_\ell(V \cap V') + \sum_{a\in i}\sum_{b\in j}(V_a/\lambda_f^3)\chi_{ab}(V_b'/\lambda_f^3), \tag{4.58}$$

where $(V \cap V')$ is the volume of the overlap between V and V'. This interpolation formula for χ_{ij} is manifestly linear in its arguments, so that

$$\chi_{ij}(V, V') + \chi_{ij}(V, V'') = \chi_{ij}(V, V' \cup V''),$$

whenever V' and V'' do not overlap. Most importantly, the interpolation procedure is simple, convenient, and correct for the limiting cases of where all the values of V_a are either 0 or λ_f^3.

For comparison, we have also calculated values of χ_{ab} from an explicit SPC/E water simulation in GROMACS at temperature $T = 298$ K and pressure $p = 1$ atm. The values are also included in the Supplementary Information of Ref. [111]. For the quantities we have studied in the main text, using these values for χ_{ab} instead of those derived from the Narten-Levy data yields nearly identical results. For example, Figure 4.12 shows the solvation free energies for spheres calculated using the SPC/E values for χ_{ab} instead of the Narten-Levy ones used in Figure 4.5. The agreement is representative of what we have observed for all the other observables discussed in Section 4.3.

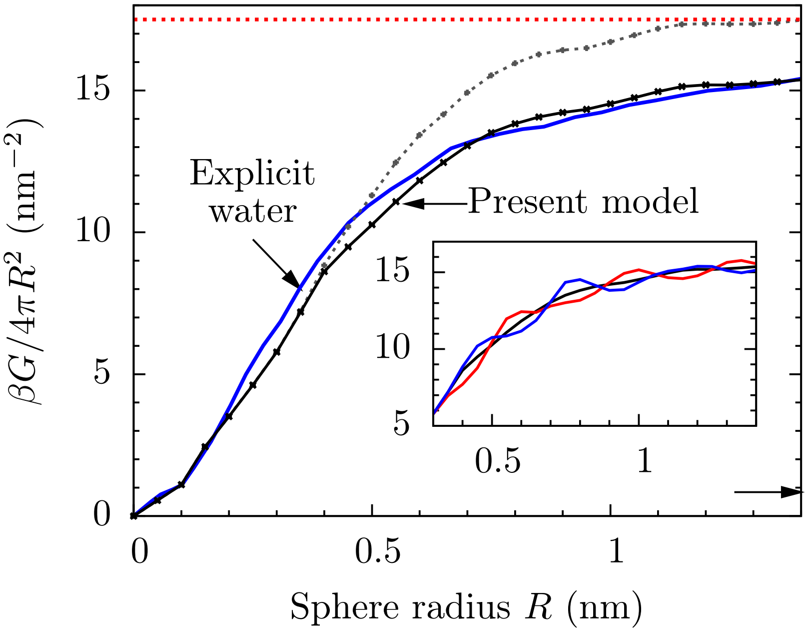


Figure 4.12: Solvation free energies G of hard spheres of increasing radii, calculated using SPC/E values for χ_{ab}. See Figure 4.5 for an explanation of the curves shown.

4.4.2 Accuracy of the one-basis approximation

In the final model of Table 4.5, we don't compute the full field $c(\mathbf{r})$ defined in Equation (4.22), but instead use the one-basis approximation of Equation (4.25). To test this approximation, we have implemented a discretized version of Equation (4.22) to calculate the values of a field c_i, the discrete analog of $c(\mathbf{r})$,

$$\sum_j n_i \chi_{ij}(v,v) n_j c_j = \rho_\ell n_i v_i. \tag{4.59}$$

This equation only determines the values of c_i for cells i that are wet ($n_i = 1$) and overlap the excluded volume ($v_i \neq 0$). This is enough information to calculate a lattice approximation to $H_S[n(\mathbf{r})]$ in Table 4.4,

$$H_S[\{n_i\}] = \frac{k_\mathrm{B}T}{2} \sum_i \rho_\ell n_i v_i c_i + C/2, \tag{4.60}$$

with C defined by Equation (4.29). Note that at every step of a Monte Carlo simulation used to simulate the model with this expression for $H_S[\{n_i\}]$, the linear system of Equation (4.59) must be solved. In our experience, this makes the model without the one-basis approximation about an order of magnitude slower to simulate than the model of Table 4.5.

Figure 4.13 shows the solvation free energy of spheres calculated with the above expression for $H_S[\{n_i\}]$, and is analogous to Figure 4.5. The unbalancing potential used here to fit the SPC/E solvation free energies, $a\rho_\ell = 5.4\,k_\mathrm{B}T$, is slightly higher than the one used in the final model. From the general agreement of the two Figures, we conclude that the one-basis approximation is a reasonably accurate one, especially given its significant computational advantage. Fortuitously, the one-basis approximation leads to lesser lattice artifacts than the purportedly more accurate model presented in this section.

4.4.3 Accuracy of the renormalization approximation to $H_I[n(\mathbf{r})]$

Above, we replaced the three terms involving $\phi(\mathbf{r})$ in the expression for $H_I[n(\mathbf{r})]$ in Table 4.4 by the simpler expression given in (4.30). We now justify this replacement.

Denote by $H_+[n(\mathbf{r})]$ the terms dropped from $H_I[n(\mathbf{r})]$ in Table 4.4. They are

$$H_+[n(\mathbf{r})] = -\int_{\mathbf{r}\in\bar{v}} \int_{\mathbf{r}'\in v} \phi(\mathbf{r})\chi(\mathbf{r},\mathbf{r}')c(\mathbf{r}') + \frac{k_\mathrm{B}T}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \beta\phi(\mathbf{r})(\chi - \chi^{(\mathrm{m})})(\mathbf{r},\mathbf{r}')\beta\phi(\mathbf{r}'). \tag{4.61}$$

Approximating $\chi^{(\mathrm{m})}(\mathbf{r},\mathbf{r}')$ as zero if $\mathbf{r}$ or $\mathbf{r}'$ are in v and $\chi(\mathbf{r},\mathbf{r}')$ otherwise, and using the one-basis set approximation of Equation (4.25), we discretize these terms to obtain a lattice

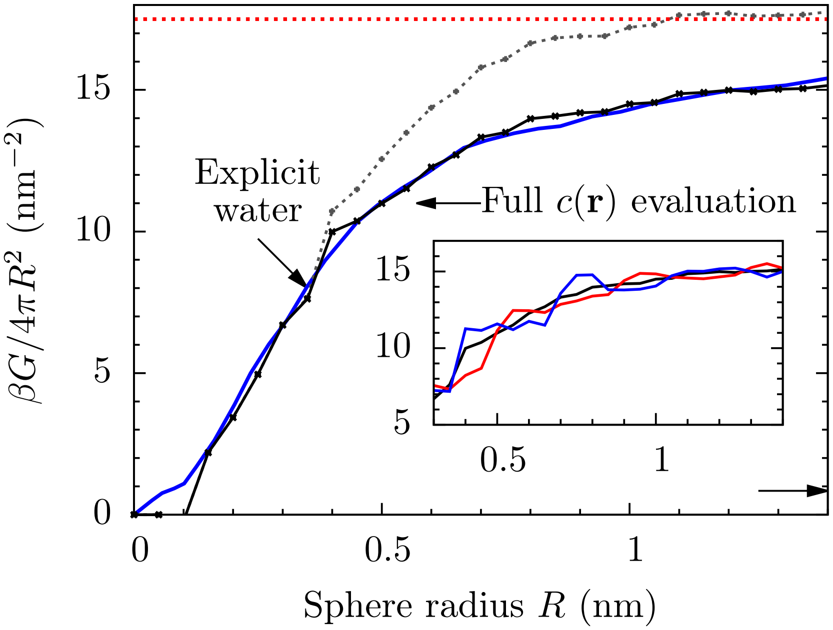


Figure 4.13: Solvation free energies G of hard spheres of increasing radii, calculated without the one-basis approximation and without the correction for Poisson statistics for small spheres (Equation (4.60), but with $C = \ln(2\pi\sigma_v)$). See Figure 4.5 for an explanation of the curves shown.

version of $H_+[n(\mathbf{r})]$,

$$H_+[\{n_i\}] = -\sum_{i,j} \phi_i n_i \chi_{ij}(\bar{v}, v) n_j \langle N \rangle_v / \sigma_v \\ + k_\mathrm{B}T \sum_{i,j} \beta\phi_i n_i [\chi_{ij}(v,v)/2 + \chi_{ij}(\bar{v},v)] n_j \beta\phi_j. \quad (4.62)$$

Because of the double sums in the formula, calculating $H_+[\{n_i\}]$ is the most computationally-demanding part of calculating $H_\mathrm{eff}[\{n_i\}]$. Since a single cell flip changes the value of ϕ_i in up to 7 cells, calculating incremental changes to $H_+[\{n_i\}]$ is also much more expensive than calculating incremental changes to $H_u[\{n_i\}]$ (Equation (4.51)), which has a similar structure.

Figure 4.14 presents the solvation free energies of hard spheres calculated when the $H_+[\{n_i\}]$ term is included and the renormalization constant K is set to 1. As can be seen, the term corresponding to $H_I[\{n_i\}]$ has a much larger absolute value, and in the region where their values are not negligible, the average values of $H_I[\{n_i\}]$ and $H_+[\{n_i\}]$ are, as claimed, essentially proportional. The renormalization procedure we implement thus seems justified, a conclusion borne out by the results in the text. For completeness, we have verified that the solvation free energies of the hexagonal plate solute (Figure 4.6) calculated when $H_+[\{n_i\}]$ is included and K is 1 are essentially identical to the ones calculated using the simpler Hamiltonian of Table 4.5.

4.4.4 Comparison to the ten Wolde-Chandler model

A particularly clean previous attempt to implement a lattice theory based on LCW ideas is that of ten Wolde and Chandler [106]. Using the notation in this chapter, the Hamiltonian of that model is

$$H[\{n_i\}] = \gamma\lambda^2 \sum_{\langle ij \rangle} (n_i - n_j)^2 - \mu\rho_\ell\lambda^3 \sum_i n_i + \frac{k_\mathrm{B}T}{2} c_0 \lambda^3 \sum_i n_i v_i, \quad (4.63)$$

where c_0 is a constant approximation to the one-basis value of c_1 (Equation (4.25)) with a value of about $120\,\mathrm{nm}^{-3}$. This value is chosen to obtain the correct solvation free energy of a 5 Å-radius sphere. The lattice cell size λ was chosen as 2.1 Å in that work. The resulting solvation free energies for spheres of radii of up to 14 Å are shown in Figure 4.15 (solid black line).

Above, we argued that the Ising Hamiltonian estimate for $H_L[n(\mathbf{r})]$ overestimates the interfacial energy of a sphere of radius R by a factor of $3/2$. However, the ten Wolde-Chandler model uses precisely this Hamiltonian, yet the solvation free energy of spheres seems to tend to the correct value as R grows. This apparent paradox is explained in the remainder of Figure 4.15, which shows how the solvation energies of spheres obtained with the ten Wolde-Chandler model differ when the lattice cell size of $\lambda = 2.1$ Å is changed to

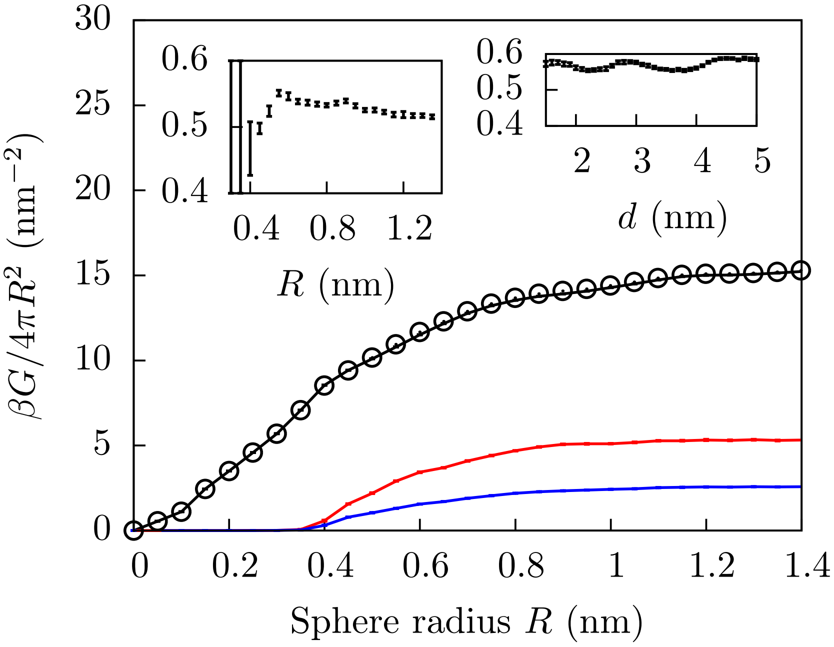


Figure 4.14: Solvation free energies G of hard spheres as a function of sphere radius, where the term $H_+[\{n_i\}]$ (Equation (4.61)) has been included and the renormalization constant K has been set to 1 (solid black), compared to the simpler model of Table 4.5 (circles). The averages of $-\langle H_{\text{int}}[\{n_i\}]\rangle$ (red) and $\langle H_+[\{n_i\}]\rangle$ (black) are nearly proportional to each other. Left Inset: implied renormalization constant K, equal to $\langle H_{\text{int}}[\{n_i\}] + H_+[\{n_i\}]\rangle \,/\, \langle H_{\text{int}}[\{n_i\}]\rangle$. Note that both the numerator and denominator take on essentially zero value for $R \lesssim 0.4\,\text{nm}$. Right Inset: Implied value of K for hexagonal plate solute (Figure 4.6) with $\eta = 1.0$. The implied value of K is similar for different η.

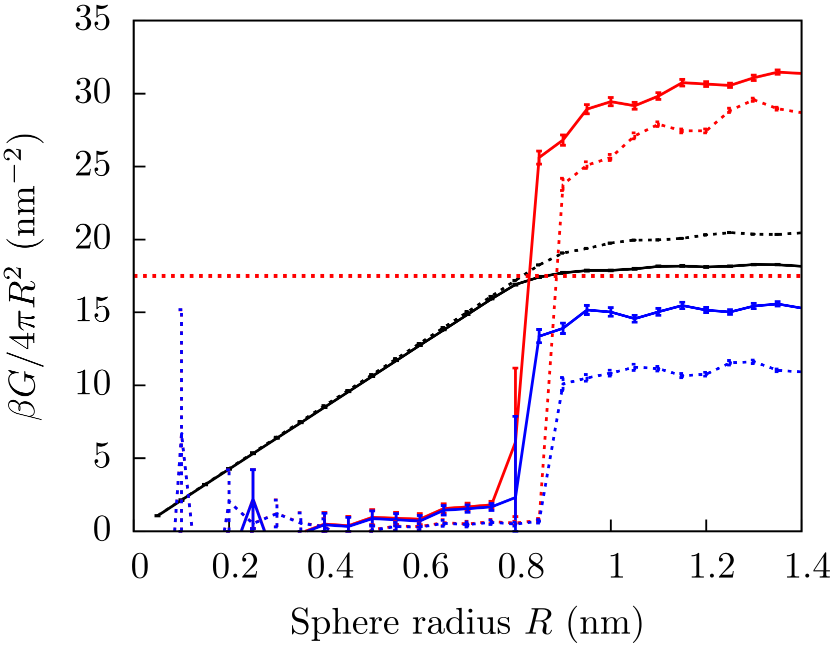


Figure 4.15: Solvation free energies G of spheres in the model of Ref. [106] (black), for cell sizes $\lambda = 2.1\,\text{Å}$ (solid) and $\lambda = 2.3\,\text{Å}$ (dashes). The use of the Ising Hamiltonian causes the average value of $H_L[\{n_i\}]$ (red) to significantly exceed the solvation free energy, but also leads to large excess entropies (blue, $TS = \langle H\rangle - G$). At $\lambda = 2.1\,\text{Å}$, but not at $\lambda = 2.3\,\text{Å}$, a fortuitous cancellation leads to correct solvation free energies.

$\lambda = 2.3\,\text{Å}$. As claimed, $\langle H_L[\{n_i\}]\rangle$ is much larger than it should be, but for $\lambda = 2.1\,\text{Å}$, the excess entropy resulting from the unphysical degeneracies of the Ising Hamiltonian exactly cancels this excess energy. This fortuitous cancellation does not occur for different cell sizes, and will not, in general, hold for solutes of different geometries.

4.4.5 Moving solutes

The solvent degrees of freedom in the model of Table 4.5, $\{n_i\}$, can be efficiently sampled using Metropolis Monte Carlo. On the other hand, sampling the solute degrees of freedom is more complex. The principal problem is to calculate the overlap volume v_a between the excluded volume v and a given fine cell a (see Section 4.4.1). The volume v is typically a union of overlapping spheres, one for each atom in the solute. For the results of this chapter, we calculated these overlap volumes for a fixed solute geometry using expensive, brute-force Monte Carlo integration. In Refs. [106] and [123], an interpolation scheme was used that only works if no point in space is within the solvent-excluding radius of three or more spheres simultaneously, and a solute geometry was chosen that avoids this situation. Neither scheme is adequate for a generic, moving solute. Moreover, if the solute is propagated through some variant of molecular dynamics, such as Langevin dynamics [106], then the gradient of $H_{\text{eff}}[\{n_i\}]$ with respect to solute positions is also needed.

Here, we discuss a partial solution to the above problem. Specifically, we present an approximate method of calculating v_a when v is, as above, a union of possibly overlapping spheres of a few different sizes.[4] The gradient of v_a with respect to the centers of these spheres is also easy to calculate. In principle, propagating these gradients to obtain gradients of $H_{\text{eff}}[\{n_i\}]$ is then simply a (non-trivial) bookkeeping exercise.

In describing our scheme, we treat cell indices a as vectors that can be added and subtracted. We denote by $\mathbf{r}_a$ the coordinates of the corner of cell a with lowest Cartesian components. For a solvent-excluding sphere of radius R centered at $\mathbf{r}_0$, we can precalculate the overlap volumes $\hat{v}_s$ of all cells s by any method, such as Monte Carlo integration. We do this once at the beginning of a simulation.

Generically, the center $\mathbf{R}$ of a solvent-excluding sphere will not coincide with a cell corner. We denote the indices of the eight corners of the cell containing $\mathbf{R}$ by $a_1, \dots, a_8$, and their positions by $\mathbf{r}_1, \dots, \mathbf{r}_8$. We construct eight non-negative weights $c_1(\mathbf{R}), \dots, c_8(\mathbf{R})$, the sum of which is 1 and whose value depends continuously on $\mathbf{R}$, such that $\mathbf{R} = \sum_{k=1}^{8} c_k(\mathbf{R})\mathbf{r}_k$. Any scheme with these characteristics, such as trilinear interpolation, can be used. The overlap volumes for cells a near $\mathbf{r}$ are then estimated by

$$\tilde{v}_a(\mathbf{R}) \approx \sum_{k=1}^{8} c_k(\mathbf{R}) \hat{v}_{a-a_k}. \tag{4.64}$$

The gradient of this volume is easy to calculate,

$$\nabla \tilde{v}_a(\mathbf{R}) \approx \sum_{k=1}^{8} \nabla c_k(\mathbf{R}) \hat{v}_{a-a_k}. \tag{4.65}$$

This interpolation scheme has the desirable property that the total volume of a sphere, given by $\sum_a \tilde{v}_a(\mathbf{R})$, is independent of $\mathbf{R}$.

When a solute is composed of multiple spheres, centered at $\mathbf{R}^N$, we simply add together the overlap volumes given by Equation (4.64) for each solute, but we cap the sum at the total volume of each fine cell, λ_f^3. In summary, we have

$$v_a \approx \min\left[\lambda_f^3, \sum_{n=1}^{N} \tilde{v}_a(\mathbf{R}_n)\right], \tag{4.66}$$

$$\nabla_n v_a \approx \begin{cases} \mathbf{0}, & v_a = \lambda_f^3, \\ \nabla \tilde{v}_a(\mathbf{R}_n), & \text{otherwise.} \end{cases} \tag{4.67}$$

---

[4] This method was developed with Tuomas Knowles.

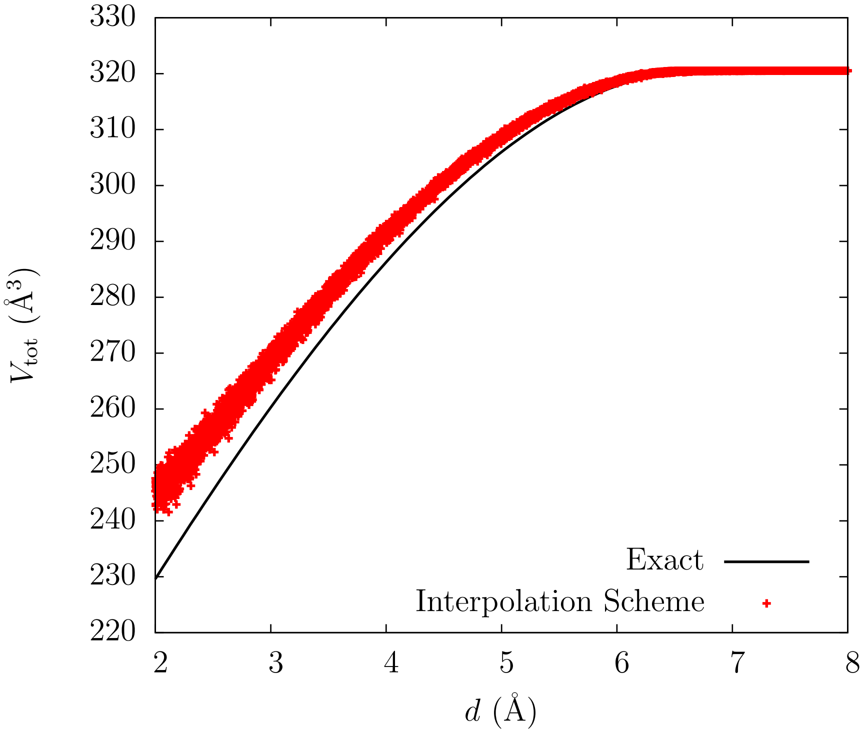


Figure 4.16: Total volume excluded by two spheres of radius $R = 3.37$ Å when a distance d apart, as calculated exactly (Equation (4.68)) and by the numerical scheme of Section 4.4.5.

This scheme is correct whenever one or more spheres overlaps cell a completely, as well as when two spheres both overlap cell a but not each other. When two or more spheres both partially overlap cell a and each other, our scheme mildly overestimates the overlap volume.

We have evaluated the precision of our scheme by calculating the total volume of two spheres of radius R whose position is varied, and the two spheres are placed at arbitrary positions with respect to the fine grid. The total volume can be calculated analytically when the two sphere centers are a distance d apart,

$$V_{\mathrm{tot}}(d) = \begin{cases} 2 \times \frac{4}{3}\pi R^3, & d > 2R, \\ 2 \times \frac{4}{3}\pi R^3 - \frac{\pi}{12}(4R + d)(2R - d)^2, & \text{otherwise.} \end{cases} \tag{4.68}$$

The results of this comparison are shown in Figure 4.16 when $R = 3.37$ Å and the fine grid resolution is 1 Å. As expected, the exact and numerical results agree closely. Of equal importance, the spread in the numerical estimate of the total volume is small, which suggests that the lattice artifacts of our overlap-volume scheme are quite modest.

4.4.6 Electrostatics

The model presented in this chapter approximates the cavity formation (ΔG_{cav}) and dispersion interaction (ΔG_{vdW}) parts of the solvation energy (ΔG_{sol} in Equation (2.15)). The remaining part, ΔG_{pol}, accounts for solvent polarization. Here, we summarize a few methods for estimating ΔG_{pol} that could, in the future, augment the model of Table 4.5. All of these methods treat the solvent as a linear dielectric continuum, and so neglect any coupling between density and polarization fluctuations. The excluded volume v and the volumes lacking solvent (where $n(\mathbf{r}) = 0$) create a non-uniform dielectric constant $\epsilon(\mathbf{r})$. The methods then estimate the work of inserting the partial charges of the solute into place in this non-uniform medium.

Conceptually, the solute charges induce charges in the dielectric medium, which can be calculated using Poisson's equation

$$\nabla \cdot \mathbf{D}(\mathbf{r}) = \rho_f(\mathbf{r}), \tag{4.69}$$

where

$$\mathbf{D}(\mathbf{r}) = \epsilon(\mathbf{r})\mathbf{E}(\mathbf{r}), \tag{4.70}$$

$$\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r}), \tag{4.71}$$

and

$$\rho_f(\mathbf{r}) = q_i\delta(\mathbf{r} - \mathbf{R}_i). \tag{4.72}$$

For the purposes of computing polarization energies, we can regard the simulation box as infinite, so the boundary conditions of the Poisson equation would be $\phi(\mathbf{r}) \to 0$ as $|\mathbf{r}| \to \infty$. The total work to create all the partial charges (including their self-energies) is then [40]

$$W = \frac{1}{2}\int_{\mathbf{r}} \mathbf{D} \cdot \mathbf{E} = \frac{1}{2}\int_{\mathbf{r}} \epsilon(\mathbf{r})|\nabla\phi(\mathbf{r})|^2. \tag{4.73}$$

The polarization energy ΔG_{pol} is then equal to this work minus the self-energy of the charges. Numerical methods exist to solve the Poisson equation (or the related Poisson-Boltzmann equation where the effect of ions is included approximately) and calculate ΔG_{sol} to arbitrary accuracy (see introduction of Ref. [72]). Most of these methods, however, are unsuitable for a dynamical treatment of solvation, either because they are too computationally demanding or suffer from serious discretization errors (i.e., ΔG_{sol} is a discontinuous function of partial charge positions).

A pragmatic alternative to solving the Poisson equation is the Generalized Born (GB) method [3, 102]. Following Ref. [3], we first consider the solvation free energy of N point charges $\{q_i\}$ at the centers of spherical cavities of radii $\{a_i\}$ inside a medium with dielectric constant ϵ_w. Charges i and j are separated by a distance r_{ij}. In the limit where all distances $\{r_{ij}\}$ are much larger than any cavity radius, we obtain the usual Born expression for polarization term of the solvation free energy,

$$\Delta G_{\text{pol}} = \sum_{i=1}^{N} \frac{q_i^2}{2a_i}(\epsilon_w^{-1} - 1) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{q_i q_j}{r_{ij}}(\epsilon_w^{-1} - 1). \tag{4.74}$$

The GB method generalizes this formula to any charge distribution in an arbitrarily shaped cavity as

$$\Delta G_{\text{pol}} = \frac{1}{2}\sum_{i,j} \frac{q_i q_j}{f_{\text{GB}}(r_{ij})}(\epsilon_w^{-1} - 1), \tag{4.75}$$

where

$$f_{\text{GB}}(r_{ij}) = \sqrt{r_{ij}^2 + R_i R_j \exp(-r_{ij}^2/4R_i R_j)}. \tag{4.76}$$

Here, the constants $\{R_i\}$ are the so-called effective Born radii of the atoms, chosen so that when all charges but charge i are turned off, Equation (4.75) is exact. In the limit of well separated spherical cavities, R_i and a_i coincide, and Equation (4.75) reduces to Equation (4.74). The particular form of $f_{\text{GB}}(r_{ij})$ in Equation (4.76) is chosen to enforce these two limits, as well as recovering the polarization energy of a small point dipole in a spherical cavity.

The effective Born radii are approximated using the so-called Coulomb field approximation. When all charges but charge i are turned off, the displacement field, $\mathbf{D}(\mathbf{r})$, is approximated by the formula

$$\mathbf{D}(\mathbf{r}) \approx \frac{q_i}{4\pi|\mathbf{r} - \mathbf{r}_i|^{3/2}}(\mathbf{r} - \mathbf{r}_i). \tag{4.77}$$

This approximation is exact when the cavity is spherical. From Equations (4.73) and (4.75), it follows that

$$R_i^{-1} \approx \int_{\mathbf{r}\ \text{out}} \frac{1}{|\mathbf{r} - \mathbf{r}_i|^4}, \tag{4.78}$$

where "out" is the region of space with dielectric constant ϵ_w. When the "out" region is a collection of possibly overlapping spherical cavities, efficient algorithms exist to estimate R_i and its gradients with respect to particle positions [43,83]. In order to integrate the Generalized Born method with our model of the hydrophobic effect, one would need to generalize these estimates to allow the solvent-deficient region (where $n(\mathbf{r}) = 0$) to be included in the "cavity" region.

Finally, a third approach would be to model the statistics of the polarization field using a Marcus-Felderhof free energy functional [30, 73], an approach that is closely related to the Gaussian field theory of Song *et al.* [100, 101]. These theories start from the following free energy functional of the polarization field $\mathbf{P}(\mathbf{r})$,

$$\begin{aligned} F_{\text{pol}}[\mathbf{P}(\mathbf{r})] = \frac{1}{2}\int_{\mathbf{r}} \mathbf{P}(\mathbf{r}) \cdot \kappa(\mathbf{r})^{-1} \cdot \mathbf{P}(\mathbf{r}) \\ - \frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \frac{(\nabla \cdot \mathbf{P}(\mathbf{r}))(\nabla \cdot \mathbf{P}(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} - \int_{\mathbf{r}} \big(\nabla \cdot \mathbf{P}(\mathbf{r})\big)\phi_0(\mathbf{r}), \end{aligned} \tag{4.79}$$

where $\phi_0(\mathbf{r})$ is the potential energy set up by the charges in vacuum,

$$\phi_0(\mathbf{r}) = \sum_{i=1}^{N} \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|}, \tag{4.80}$$

and

$$\kappa(\mathbf{r}) = \epsilon(\mathbf{r}) - 1. \tag{4.81}$$

The free energy of Equation (4.79) encodes the free energy cost of establishing the polarization field (as determined by the dielectric constant), and the interaction of the bound charge density $\nabla \cdot \mathbf{P}(\mathbf{r})$ with itself and with the free charges. Minimizing $F_{\text{pol}}[\mathbf{P}(\mathbf{r})]$ with respect to variations in $\mathbf{P}(\mathbf{r})$ yields the usual equations for electric fields in non-uniform media. However, the idea of minimizing a free energy can be approximated in a dynamical simulation via a Car-Parrinello-like extended Lagrangian [9, 34]. A clean realization of this proposal is the model of Marchi *et al.* [72].

The free energy functional in Equation (4.79) contains information beyond macroscopic dielectric continuum theory, since it allows for thermal fluctuations in the polarization field. Recasting this free energy as a Gaussian field theory for $\mathbf{P}(\mathbf{r})$ yields

$$F_{\text{pol}}[\mathbf{P}(\mathbf{r})] = \frac{1}{2}\int_{\mathbf{r}} \mathbf{P}(\mathbf{r}) \cdot \kappa(\mathbf{r})^{-1} \cdot \mathbf{P}(\mathbf{r}) - \frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \mathbf{P}(\mathbf{r}) \cdot \nabla\nabla' \frac{1}{|\mathbf{r}-\mathbf{r}'|} \cdot \mathbf{P}(\mathbf{r}') + \int_{\mathbf{r}} \mathbf{P}(\mathbf{r}) \cdot \mathbf{E}_0(\mathbf{r}), \tag{4.82}$$

which follows from integration by parts. Here, $\mathbf{E}_0(\mathbf{r}) = -\nabla\phi_0(\mathbf{r})$. In a language closer to that of Gaussian fields, we have

$$F_{\text{pol}}[\mathbf{P}(\mathbf{r})] = \frac{k_{\text{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \mathbf{P}(\mathbf{r}) \cdot \chi^{-1}(\mathbf{r},\mathbf{r}') \cdot \mathbf{P}(\mathbf{r}') + \int_{\mathbf{r}} \mathbf{P}(\mathbf{r}) \cdot \mathbf{E}_0(\mathbf{r}), \tag{4.83}$$

where

$$k_{\text{B}}T\chi^{-1}(\mathbf{r},\mathbf{r}') = \frac{1}{\epsilon(\mathbf{r})-1}\delta(\mathbf{r}-\mathbf{r}') - \nabla\nabla'\frac{1}{|\mathbf{r}-\mathbf{r}'|}. \tag{4.84}$$

Note that $\chi(\mathbf{r},\mathbf{r}')$ here is a tensor. Since Equation (4.83) is a Gaussian free energy functional, we can immediately obtain the free energy of polarization from the results in Table 4.1,

$$\Delta G_{\text{sol}} = -\frac{k_{\text{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \beta\mathbf{E}_0(\mathbf{r}) \cdot \chi(\mathbf{r},\mathbf{r}') \cdot \beta\mathbf{E}_0(\mathbf{r}') + \sum_{i\neq j}\frac{q_i q_j}{r_{ij}}. \tag{4.85}$$

The last term corresponds to the interaction of the partial charges in vacuum. In practice, this equation is not useful in the case of a non-uniform dielectric constant $\epsilon(\mathbf{r})$, because calculating $\chi(\mathbf{r},\mathbf{r}')$ in that case amounts to obtaining the Green's function of the relevant Poisson equation. Song *et al.*'s contribution [100, 101] is to recast the non-uniform dielectric as a zero-fluctuation constraint inside a cavity volume c of an otherwise uniform dielectric. The function $\chi(\mathbf{r}, \mathbf{r}')$ for a uniform dielectric with dielectric constant ϵ is known [30],

$$\beta\chi(\mathbf{r}, \mathbf{r}') = \frac{\epsilon - 1}{\epsilon}\left(\delta(\mathbf{r} - \mathbf{r}') + (\epsilon - 1)\nabla\nabla'\frac{1}{|\mathbf{r} - \mathbf{r}'|}\right). \tag{4.86}$$

The constrained fluctuation spectrum, solvation free energy and average polarization field then follow directly from the results of Table 4.2. Song and Chandler [100] discuss a numerical scheme to calculate these properties for cavities of fixed shape.

We thus expect that a model for the polarization can be constructed, similar to the one we have built in this Chapter for the solvent density. However, there are serious obstacles in doing this that we have not attempted to address. First, while the solvent density fluctuation spectrum was quite local, the polarization fluctuation spectrum is distinctly non-local. Thus, we can't approximate $\chi^{(\mathrm{m})}(\mathbf{r}, \mathbf{r}')$ as $\chi(\mathbf{r}, \mathbf{r}')$ for $\mathbf{r}$ and $\mathbf{r}'$ outside of v and zero otherwise, as we could with the solvent density. Equally problematic, the "one-basis approximation" that we used to simplify $H_S[n(\mathbf{r})]$ is thoroughly inadequate in the case of polarization. To see this, we implement this approximation as a generalization of Equation (4.31) and obtain,

$$\Delta G_{\mathrm{sol}} \approx \sum_{i \neq j} \frac{q_i q_j}{r_{ij}} + k_{\mathrm{B}}T \ln\sqrt{\det 2\pi\chi_{\mathbf{v}}} + \frac{1}{2}\Big[\int_{\mathbf{r}\in c} \mathbf{P}_0(\mathbf{r})\Big] \cdot \chi_c^{-1} \cdot \Big[\int_{\mathbf{r}\in c} \mathbf{P}_0(\mathbf{r})\Big], \tag{4.87}$$

where $\mathbf{P}_0(\mathbf{r}) = -\int_{\mathbf{r}'} \chi(\mathbf{r}, \mathbf{r}')\mathbf{E}_0(\mathbf{r}) = (\epsilon - 1)\mathbf{E}_0(\mathbf{r})$ is the polarization field in the absence of a constraint, the matrix χ_c is the 3×3 matrix given by

$$\chi_c = \int_{\mathbf{r}\in c}\int_{\mathbf{r}'\in c} \chi(\mathbf{r}, \mathbf{r}'), \tag{4.88}$$

with $\chi(\mathbf{r}, \mathbf{r}')$ the uniform fluctuation spectrum of Equation (4.86), and we've omitted the last term in Equation (4.31) that corresponds to the self-energy of the charges in the dielectric medium. In the limit of widely separated spherical cavities around each charge (Equation (4.74)), $\int_{\mathbf{r}\in c} \mathbf{P}_0(\mathbf{r})$ is zero, so in this limit, Equation (4.87) fails to capture *any* polarization effects. Thus, implementing the constrained Gaussian field theory for dielectric media in a way that allows for large-length-scale solvent density fluctuations remains an open problem.

Chapter 5

Hydrophobic effects at surfaces

In this chapter, we explore how hydrophobic forces are modified near large surfaces of different chemical character [86]. As discussed in Chapter 3 and modeled in Chapter 4, solutes in bulk are accommodated in water through a combination of two mechanisms: spontaneous density fluctuations and inclusion within vapor bubbles. The balance between these two mechanisms shifts from the former to the latter as the size of the solute increases. In making this statement, the term “solute” means everything that is dissolved in the solvent. This perspective is not always the most natural one. One might want to think of the “solute” as only a small part of the system, and regard the remaining part as some external constraint, like a large surface. From this perspective, the manifestations of the hydrophobic effect differ considerably with respect to when the solute is isolated in bulk water.

In Section 3.2.3, we briefly assumed this perspective when calculating water number distributions near a large hydrophobic plate. There, we found that the existing liquid-vapor interface around the plate facilitates solvating a cavity next to the solute, compared with solvating the same cavity in bulk. In this chapter, we delve much more deeply into these issues, and rationalize them to near-quantitative accuracy using a minimal model based on LCW ideas.

We begin by calculating the solvation free energy of cavities of varying size next to model surfaces of varying chemical character and hydrophobicities. We further calculate the temperature dependence of these free energies to discover whether entropic or enthalpic considerations prevail in each case. These two sets of observations capture the phenomenology of solvation at surfaces.

Next, we construct a minimal solvation model, adapted to solvation near surfaces, based on LCW ideas, and discuss the pertinent implementation details. Using this model, we calculate a set of solvation free energies and their temperature dependence analogous to the ones described above. In all cases, we find reasonable agreement in both qualitative and quantitative aspects. Having reduced the hydrophobic effect to its essence, we rationalize all the observed behavior in terms of the simple picture of a precarious balance between spontaneous density fluctuations and liquid-vapor interfacial fluctuations.

We conclude the chapter by considering the broader implications of our results. In particular, we argue that hydrophobic surfaces can be regarded as generic catalysts for disassembly and unfolding of hydrophobically collapsed polymers. This finding may be relevant in the functioning of chaperonin proteins.

5.1 Phenomenology

5.1.1 Setup

Our simulation setup is depicted in Figure 5.1. In a $7 \times 7 \times 9\,\mathrm{nm}^3$ simulation box at temperature $T = 300\,\mathrm{K}$, we prepare a slab of SPC/E water that is open at the top and is adjacent to a model self-assembled monolayer (SAM) at the bottom. The model SAM, similar to one used in previous studies [39], consists of a periodic, crystalline array of fixed gold atoms, to each of which are attached two alkyl chains (12 carbons in length) capped by a head group of our choosing. The particular head group controls the SAM's hydrophobicity. The head groups we use, in ascending hydrophilicity, are CH_3, OCH_3, CH_2CN, $CONH_2$ and OH. The contact angle calculated in simulations between a droplet of water and the SAM agrees well with experiment [39].

In this setup, we calculate the solvation free energy, ΔG_{sol}, of an $L \times L \times W\,\mathrm{nm}^3$ probe volume, V, immediately adjacent to the SAM (shown in green in Figure 5.1) with the relation $\Delta G_{\mathrm{sol}} = -k_{\mathrm{B}}T \ln P_V(0)$. The width W is fixed at $0.3\,\mathrm{nm}$ to specifically capture hydrophobic effects at the SAM surface, and the side length L ranges from 0.5 to 3.0 nm. For $L < 1\,\mathrm{nm}$, we measured $P_V(0)$ directly in an unperturbed system, whereas for larger L, we used the method of Section 3.2.1. For comparison, we also calculate ΔG_{sol} for the same cavity in bulk water.

5.1.2 Solvation free energies at surfaces

Figure 5.2 shows the calculated solvation free energies, scaled by the cavity's surface area, $A = 2L^2 + 4WL$. In bulk, we observe the expected length-scale dependence of the quantity $\Delta G_{\mathrm{sol}}/A$ (Figure 2.3), namely a gradual transition from ΔG_{sol} scaling as the volume of the cavity to it scaling as the cavity's surface area. The solvation free energies next to the -OH-capped SAM are virtually identical, as expected for a hydrophilic surface (Figure 3.6(b)) and are substantially reduced as the SAM is made more hydrophobic. For the most part, this reflects the reduced attraction between the SAM and the water when the head groups are hydrophobic, although we argue below that enhance fluctuations at the SAM-water surface also contribute to this reduced solvation free energy. For the most hydrophobic SAM, $\Delta G_{\mathrm{sol}}/A$ is essentially constant. Below, we argue that this behavior is a consequence of solvation near a hydrophobic SAM-water interface being dominated by interfacial fluctuations at both small and large length scales.

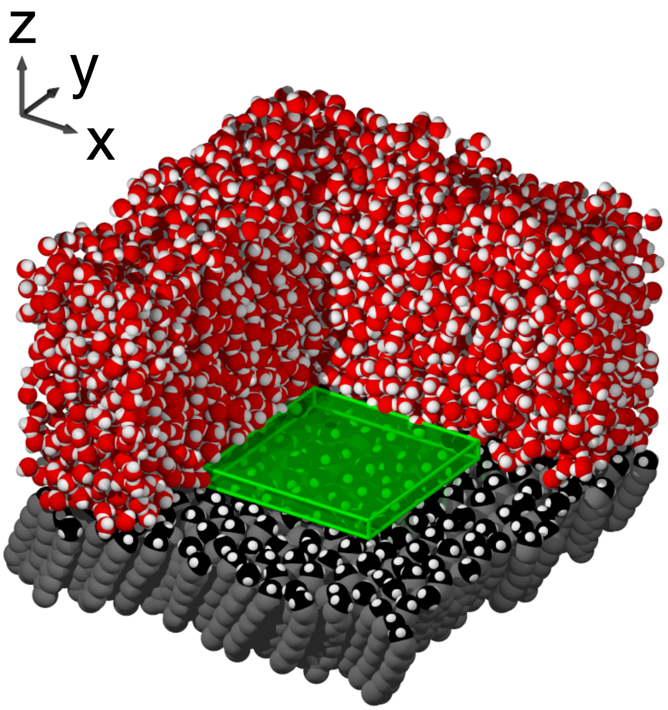


Figure 5.1: Simulation setup for calculating ΔG_{sol} for a thin probe cavity (green) in water (red and white spheres, partially removed for clarity) atop a model SAM (top alkyl chains in grey; gold atoms and bottom alkyl chain not shown). The SAM head group was varied from hydrophobic (-CH_3, shown here in black and white) to hydrophilic (-OH).

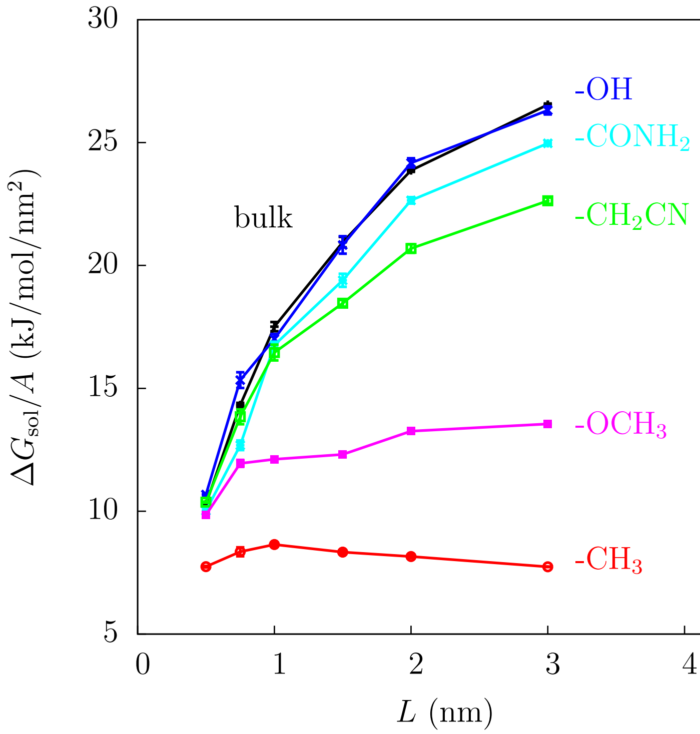


Figure 5.2: Solvation free energy per unit area, $\Delta G_{\mathrm{sol}}/A$, for an $L \times L \times 0.3\,\mathrm{nm}^3$ probe volume next to SAM surfaces with different head groups, at temperature $T = 300\,\mathrm{K}$. The quantity $\Delta G_{\mathrm{sol}}/A$ for the same cavity in bulk water is shown in black.

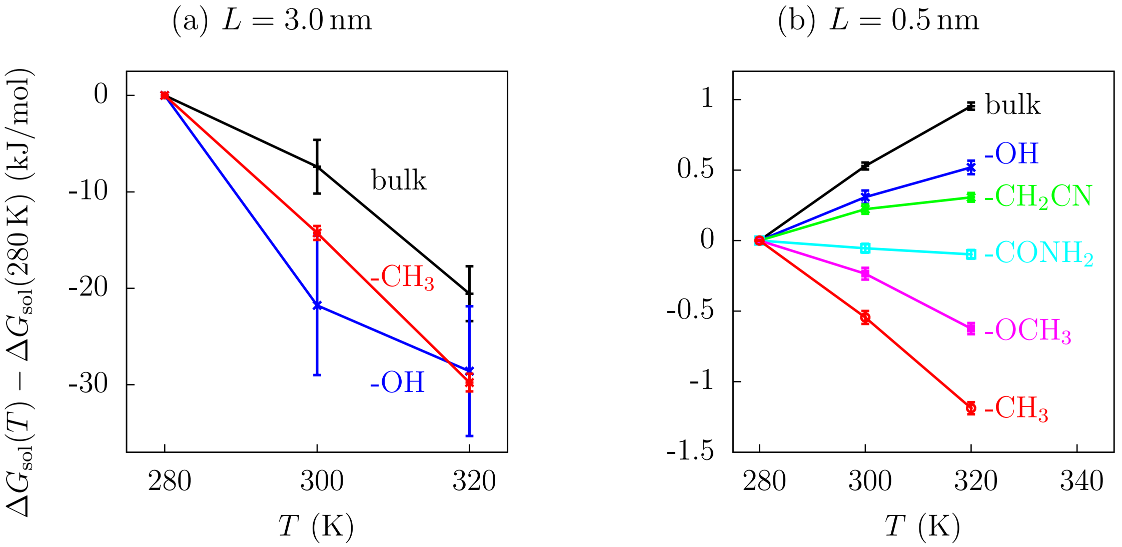


Figure 5.3: Solvation free energy per unit area, $\Delta G_{\mathrm{sol}}/A$, for an $L \times L \times 0.3\,\mathrm{nm}^3$ probe volume next to SAM surface and in bulk water at temperature T, relative to its value at $T = 280\,\mathrm{K}$.

5.1.3 Temperature dependence of solvation free energies

To assess the temperature dependence of $\Delta G_{\mathrm{sol}}/A$, one could in principle calculate the solvation entropy $\Delta S_{\mathrm{sol}} = -\partial \Delta G_{\mathrm{sol}}/\partial T$ directly from one simulation at $T = 300\,\mathrm{K}$, using the relation $\Delta S_{\mathrm{sol}} = (\Delta H_{\mathrm{sol}} - \Delta G_{\mathrm{sol}})/T$, where $\Delta H_{\mathrm{sol}} = \langle H \rangle_{N=0} - \langle H \rangle$. However, this direct estimate has large statistical errors, since a small quantity is estimated as a difference between two large quantities. Instead, we have calculated ΔG_{sol} at two other temperatures, 280 K and 320 K, and approximated the derivative numerically, as would be done in an experiment. When L is large, calculating ΔG_{sol} is expensive, so we have only done so for the most hydrophobic and most hydrophilic SAMs, as well as in bulk water.

Figure 5.3(a) summarizes the results at large length scales ($L = 3.0\,\mathrm{nm}$). In this regime, all solvation free energies decrease with increasing temperature. This result is consistent with interfacial physics dominating solvation at these length scales, since the liquid-vapor surface tension of water decreases with temperature (Equation (2.20)). The absolute decrease in ΔG_{sol} is equal, within errors, for the three systems depicted, even though ΔG_{sol} is about four times as large for bulk water and next to the -OH-capped SAM than next to the -CH$_3$-capped SAM. We rationalize this observation below.

Figure 5.3(b) summarizes the results at small length scales ($L = 0.5\,\mathrm{nm}$). In this regime, the solvation thermodynamics is more subtle. On the one hand, in bulk and next to the hydrophilic SAMs, ΔG_{sol} increases with temperature, which implies that solvating the cavity primarily reduces the entropy of the surrounding water (Equation (2.27)). On the other hand, next to the most hydrophobic SAM, ΔG_{sol} decreases with temperature, as would occur if

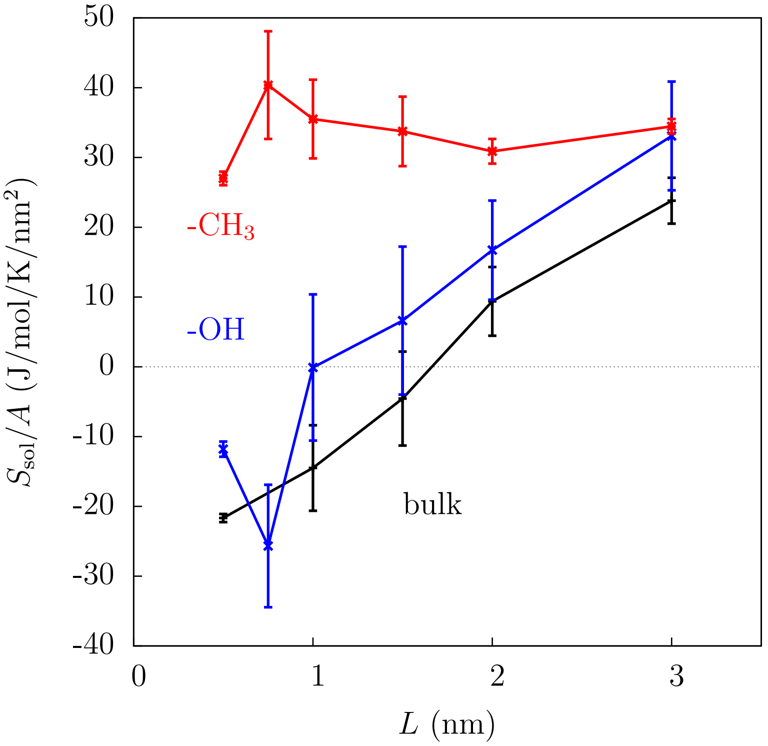


Figure 5.4: Solvation entropy per unit area, $\Delta S_{\mathrm{sol}}/A$, at $T = 300\,\mathrm{K}$, for an $L \times L \times 0.3\,\mathrm{nm}^3$ probe volume next to SAM surface. The solvation entropy per unit area of the same cavity in bulk water is shown for comparison.

interfacial physics dominated even at this length scale. For intermediate hydrophobicities, an intermediate behavior is observed.

These observations are summarized for all length scales in Figure 5.4. For the most hydrophobic SAM, the solvation entropy is nearly constant at all length scales, whereas for the most hydrophilic SAM, the solvation entropy crosses over from negative to positive as length scales grow, as happens in bulk.

5.2 Theory

5.2.1 Minimal model

To rationalize the observations of the previous section, we have constructed a minimal solvation model, adapted to the geometry in Figure 5.1, based on LCW ideas. Specifically, given the thin probe volumes that we consider, we expect that the slowly-varying field of

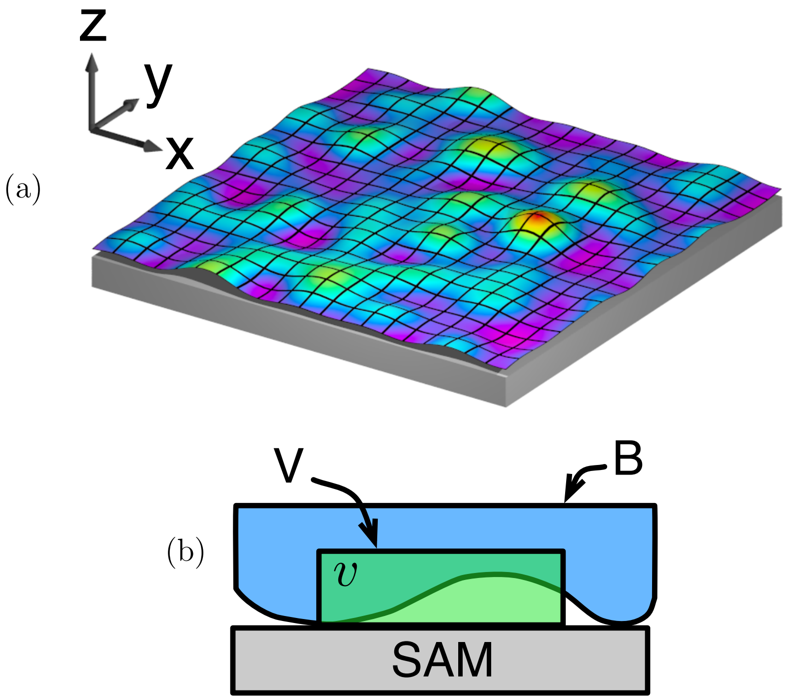


Figure 5.5: (a) A typical configuration of the model interface $z = h(x, y)$ (color-coded by z). The grey region represents the model surface. (b) Important volumes in estimating the free energy ΔG_{sol} of emptying the probe volume V (green) via Equation (5.2). The region $z > h(x, y)$ is the volume B (blue), and the intersection of V and B is v (dark green).

LCW theory, $\rho_\ell n(\mathbf{r})$, will be uniformly 1 in the bulk and uniformly 0 inside the SAM, with a single, continuous interface separating the two regions. Hence, instead of modeling the entire field $\rho_\ell n(\mathbf{r})$, we model only the interface, which we describe as a periodic elastic membrane, $z = h(x, y)$. The associated Hamiltonian, $H[h(x, y)]$, captures its surface tension and its interaction with the underlying surface,

$$H[h(x,y)] = \int_{x,y} \left[\frac{\gamma}{2} |\nabla h(x,y)|^2 + \int_{z=h(x,y)}^{\infty} \rho_\ell U(\mathbf{r}) \right]. \tag{5.1}$$

Here, γ is the experimental liquid-vapor surface tension of water, ρ_ℓ is the bulk number density of water, and $U(\mathbf{r})$ is the interaction potential between the surface below and a water molecule at position $\mathbf{r} = (x, y, z)$. The square-gradient term in Equation (5.1) accurately captures the energetics of interfacial capillary waves only for wavelengths that are large compared to atomic dimensions, so we restrict $h(x, y)$ to contain only modes with wavevectors below $2\pi/9\,\text{Å}$ (see Figure 3.20). Figure 5.5(a) depicts a typical configuration of $h(x, y)$.

Figure 5.5(b) shows a side view of an $L \times L \times W$ probe volume, V, located at the surface-water interface. At any instant in time, part of V can be empty due to an interfacial fluctuation. The number of waters, N, in the remaining volume, v, fluctuates according to a distribution $P_v(N)$. For V to be empty, N must be 0. Thus, we estimate the free energy for emptying V to be

$$\Delta G_{\mathrm{sol}}(V) = -k_{\mathrm{B}}T \ln \int \mathcal{D}h\, Z^{-1} e^{-\beta H[h(x,y)]} P_v(0), \tag{5.2}$$

where the constant $Z = \int \mathcal{D}h \, \exp\{-\beta H[h(x,y)]\}$ is the partition function of the membrane. The volume v depends on the interfacial configuration $h(x,y)$, i.e., $v = v[h(x,y)]$.

For small v, we know that $P_v(N)$ is well-approximated by the Gaussian model (Equation (2.25)). If water were far from liquid-vapor coexistence, then $P_v(N)$ would also be close to Gaussian for arbitrarily large v. The fact that water at ambient conditions is near liquid-vapor coexistence, and that there is a liquid-vapor-like interface near the SAM, is captured by the additional interfacial energy factor $Z^{-1}\exp\{-\beta H[h(x,y)]\}$ in Equation (5.2). The net result is that the thermal average in Equation (5.2) is dominated by interface configurations where v is small, so that even at ambient conditions, we can approximate $P_v(N)$ using the Gaussian model. Explicitly,

$$P_v(N) \approx (2\pi\sigma_v)^{-1/2} \exp\left[-(N - \langle N\rangle_v)^2/2\sigma_v\right], \tag{5.3}$$

where $\langle N\rangle_v$ is the average number of waters in v and $\sigma_v = \langle(\delta N)^2\rangle_v$ is its variance. We estimate these by assuming that the solvent density responds linearly to the attractive potential, $U(\mathbf{r})$, in the volume occupied by the water, B, depicted in Figure 5.5(b). Hence,

$$\begin{aligned} \langle N\rangle_v &\approx \rho_\ell v - \int_{\mathbf{r}\in v}\int_{\mathbf{r}'\in B} \chi(\mathbf{r},\mathbf{r}')\beta U(\mathbf{r}'), \\ \sigma_v &\approx \int_{\mathbf{r}\in v}\int_{\mathbf{r}'\in v} \chi(\mathbf{r},\mathbf{r}'), \end{aligned}$$

where

$$\chi(\mathbf{r},\mathbf{r}') \approx \rho_\ell \delta(\mathbf{r}-\mathbf{r}') + \rho_\ell^2[g(|\mathbf{r}-\mathbf{r}'|) - 1]. \tag{5.4}$$

Here, $g(r)$ is the oxygen-oxygen radial distribution function of water [81].

The surface-water interaction is modeled by a potential, $U(\mathbf{r})$, chosen to closely mimic the attractive potential that the -CH_3 SAM exerts on the water above it. It is composed of three terms,

$$U(\mathbf{r}) = U_{\mathrm{wall}}(\mathbf{r}) + \eta U_{\mathrm{head}}(\mathbf{r}) + U_{\mathrm{tail}}(\mathbf{r}).$$

The first term, $U_{\mathrm{wall}}(\mathbf{r})$, is a sharply repulsive potential in the region $z < R_0$ that captures the hard-core exclusion of a plane of head groups at $z = 0$ with hard-sphere radius R_0. The second term, $U_{\mathrm{head}}(\mathbf{r})$, captures the head group–water interaction, with the head groups modeled as a plane of Lennard-Jones (LJ) interaction sites of area density μ_{head} at $z = 0$

and the LJ parameters of an OPLS/UA CH_3 group interacting with an SPC/E water oxygen atom [58]. It is scaled by a parameter η that we discuss shortly. The final term, $U_{\text{tail}}(\mathbf{r})$, similarly captures the alkane tail–water interaction, with the tail groups modeled as a uniform half-space of LJ interaction sites of volume density ρ_{tail} at a distance ζ below the head groups and the LJ parameters of an OPLS/UA CH_2 group interacting with an SPC/E water oxygen atom. The parameters R_0, ζ, μ_{head} and ρ_{tail} are dictated by the geometry of the SAM. Their values and the explicit form of the potential are given below.

The parameter η is used to tune the strength of the surface-water interaction. A surface with $\eta = 1$ resembles a -CH_3-capped SAM, while higher values of η result in more hydrophilic surfaces. Such surfaces, however, lack the specific details of hydrogen bonding interactions, such as those between an -OH-capped SAM and water, so comparisons between high-η model surfaces and hydrophilic SAMs are qualitative in nature.

5.2.2 Implementation details

This section describes in detail how we perform computations with the minimal solvation model. The results we obtain in Sections 5.2.3 and 5.2.4 can be understood independently of the material in this section.

Interface description

We describe the liquid-vapor-like interface next to the model surface by a periodic height function $h(\mathbf{a})$, with $\mathbf{a} = (x, y)$ and $-D/2 \leq x, y < D/2$. This function is sampled discretely at a resolution Δ, at points satisfying

$$\mathbf{a} = (n_x\Delta, n_y\Delta), \qquad -\frac{D}{2\Delta} \leq n_x, n_y < \frac{D}{2\Delta}.$$

This results in N^2 discrete sampling points $\{\mathbf{a}\}$, with $N = D/\Delta$. In the following, sums over $\mathbf{a}$ denote sums over these N^2 sampling points. We have used $D = 60\,\text{Å}$ and $\Delta = 1\,\text{Å}$.

The discrete variables $\{h_{\mathbf{a}}\}$ represent the interface height at each sample point $\mathbf{a}$, so that

$$h_{\mathbf{a}} = h(\mathbf{a}), \quad \text{for } \mathbf{a} = (n_x\Delta, n_y\Delta).$$

This notation clearly distinguishes between the N^2 height variables $h_{\mathbf{a}}$ and the continuous height function $h(\mathbf{a})$ that they represent.

The discrete Fourier transform of $\{h_{\mathbf{a}}\}$ is denoted by $\{\tilde{h}_{\mathbf{k}}\}$, and is defined at wavevectors $\mathbf{k} = (2\pi/L)(m_x, m_y)$, with $-N/2 \leq m_x, m_y < N/2$. We use the symmetric normalization convention throughout for Fourier transforms.

Energetics

The essential property of the liquid-vapor-like interface is its surface tension. For a free interface, the resulting capillary-wave Hamiltonian [8] is

$$H_0[\{h_{\mathbf{a}}\}] \approx \frac{\gamma\Delta^2}{2}\sum_{\mathbf{a}} |\nabla h_{\mathbf{a}}|^2 \approx \frac{\gamma\Delta^2}{2}\sum_{\mathbf{k}} k^2 |\tilde{h}_{\mathbf{k}}|^2,$$

where $\nabla h_{\mathbf{a}}$ is a finite-difference approximation to $\nabla h(\mathbf{a})$.

With the Willard-Chandler definition of the instantaneous liquid-vapor interface [125], the power spectrum of capillary waves in SPC/E water agrees with the spectrum predicted by the above Hamiltonian for wavevectors smaller than about $2\pi/\ell$, but is substantially lower for higher wavevectors (see Figure 3.20). This result is consistent with the liquid-vapor-like interfaces being sensitive to molecular detail at high wavevectors [94]. At $T = 300\,K$, we have found that $\ell \approx 9\,\text{Å}$. We thus constrain all Fourier components $\tilde{h}_{\mathbf{k}}$ to be zero for high $\mathbf{k}$, i.e.

$$\tilde{h}_{\mathbf{k}} = 0, \qquad |\mathbf{k}| > 2\pi/\ell. \tag{5.5}$$

In our model, the liquid-vapor-like interface interacts with a model surface via a potential that depends on $\{h_{\mathbf{a}}\}$. As discussed below, it is also convenient to introduce additional umbrella potentials to aid in sampling. The Hamiltonian of the interface subject to this additional potential energy $H'[\{h_{\mathbf{a}}\}]$ is

$$H[\{h_{\mathbf{a}}\}] = \frac{\gamma\Delta^2}{2}\sum_{\mathbf{k}} k^2 |\tilde{h}_{\mathbf{k}}|^2 + H'[\{h_{\mathbf{a}}\}]. \tag{5.6}$$

When expressed as a function of the Fourier components $\{\tilde{h}_{\mathbf{k}}\}$, we denote the Hamiltonian by $\tilde{H}[\{\tilde{h}_{\mathbf{k}}\}]$ and the external potential by $\tilde{H}'[\{\tilde{h}_{\mathbf{k}}\}]$, so that

$$\tilde{H}[\{\tilde{h}_{\mathbf{k}}\}] = \frac{\gamma\Delta^2}{2}\sum_{\mathbf{k}} k^2 |\tilde{h}_{\mathbf{k}}|^2 + \tilde{H}'[\{\tilde{h}_{\mathbf{k}}\}].$$

Dynamics

We calculate thermal averages of interface configurations by introducing a fictitious Langevin dynamics and replacing thermal averages by trajectory averages. We first assign a mass per unit area μ to the interface. The Lagrangian in real space is

$$L[\{h_{\mathbf{a}}, \dot{h}_{\mathbf{a}}\}] = \frac{\mu\Delta^2}{2}\sum_{\mathbf{a}} \dot{h}_{\mathbf{a}}^2 - H[\{h_{\mathbf{a}}\}].$$

The corresponding Lagrangian in Fourier space is

$$\tilde{L}[\{\tilde{h}_{\mathbf{k}}, \dot{\tilde{h}}_{\mathbf{k}}\}] = \frac{\mu\Delta^2}{2}\sum_{\mathbf{k}} |\dot{\tilde{h}}_{\mathbf{k}}|^2 - \tilde{H}[\{\tilde{h}_{\mathbf{k}}\}].$$

Since all $h_{\mathbf{a}}$ are real, the amplitudes of modes $\mathbf{k}$ and $-\mathbf{k}$ are related, $\tilde{h}_{\mathbf{k}} = \tilde{h}^*_{-\mathbf{k}}$. Taking this constraint and Equation (5.5) into account, the Euler-Lagrange equations yield equations of motion in Fourier space. To thermostat each mode, we add Langevin damping and noise terms. The final equation of motion has the form

$$\mu\Delta^2\ddot{\tilde{h}}_{\mathbf{k}} = -\gamma\Delta^2|\mathbf{k}|^2\tilde{h}_{\mathbf{k}} - \frac{\partial\tilde{H}'[\{\tilde{h}_{\mathbf{k}}\}]}{\partial\tilde{h}_{\mathbf{k}}} - \eta\dot{\tilde{h}}_{\mathbf{k}} + \tilde{\xi}_{\mathbf{k}}(t), \qquad (|\mathbf{k}| < 2\pi/\ell), \tag{5.7}$$

The Langevin damping constant η is chosen to decorrelate momenta over a time scale τ, so $\eta = \mu\Delta^2/\tau$. The zero-mean Gaussian noise terms $\{\tilde{\xi}_{\mathbf{k}}(t)\}$ have variance such that

$$\langle\tilde{\xi}^*_{\mathbf{k}}(t)\tilde{\xi}_{\mathbf{k}'}(t')\rangle = 2\eta k_{\mathrm{B}}T\delta(t-t')\delta_{\mathbf{k},\mathbf{k}'}.$$

As with $\tilde{h}_{\mathbf{k}}$, the $\tilde{\xi}_{\mathbf{k}}$ variables satisfy the related constraint $\tilde{\xi}_{\mathbf{k}} = \tilde{\xi}^*_{-\mathbf{k}}$. Hence, for $\mathbf{k} = \mathbf{0}$, the noise is purely real and its variance is twice that of the real and imaginary components of all other modes.[1]

We propagate these equations of motion using the Velocity Verlet algorithm [34]. At each force evaluation, we use a Fast Fourier Transform (FFT [35]) to calculate $\{h_{\mathbf{a}}\}$ from $\{\tilde{h}_{\mathbf{k}}\}$. We then calculate $\partial H'[\{h_{\mathbf{a}}\}]/\partial h_{\mathbf{a}}$ in real space and perform an inverse FFT to obtain the force $\partial\tilde{H}'[\{\tilde{h}_{\mathbf{k}}\}]/\partial\tilde{h}_{\mathbf{k}}$ on mode $\tilde{h}_{\mathbf{k}}$ due to $H'[\{h_{\mathbf{a}}\}]$. We then add the forces due to surface tension, Langevin damping and thermal noise, as in Equation (5.7).

For the Velocity Verlet algorithm to be stable, we choose a timestep equal to 1/20th of the typical time scale of the highest-frequency mode of the free interface, $\Delta t = \frac{1}{20}\sqrt{\mu\ell^2/\gamma}$. To equilibrate the system quickly but still permit natural oscillations, we choose the Langevin damping time scale so that $\tau = 100\Delta t$. Finally, we choose a value of μ close to the mass of a single water layer, $\mu = 100\,\mathrm{amu/nm^2}$.

This interface dynamics is entirely fictitious. However, it correctly samples configurations of the interface, Boltzmann-weighted by the Hamiltonian $H[\{h_{\mathbf{a}}\}]$. This is true irrespective of the exact values of μ, Δt and τ, so our choices have no effect on the results. We have simply chosen reasonable values that do not lead to large discretization errors when solving the system's equations of motion.

Surface-interface interactions

The liquid-vapor-like interface interacts with the model surface via a potential $H'_{\mathrm{surf}}[\{h_{\mathbf{a}}\}]$. In the atomistic simulations, the SAM sets up an interaction potential $U(\mathbf{r})$ felt by the atoms in the water molecules. To model this interaction potential, we smear out the atomistic detail of the SAM and replace it with three elements:

---

[1] The constraint on the magnitude of $\mathbf{k}$ ensures that no Nyquist modes, i.e., modes with k_x or k_y equal to $\pm\pi/D$, are ever excited. If they were included, these modes would also be purely real, and the variance of the real component of their noise terms would likewise be twice that of the real component of the interior modes.

- A uniform area density μ_{head} of Lennard-Jones sites (with length and energy scales σ_{head} and ϵ_{head}) in the $z = 0$ plane to represent the SAM head groups.
- A uniform volume density ρ_{tail} of Lennard-Jones sites (with length and energy scales σ_{tail} and ϵ_{tail}) in the half-space $z < -\zeta$ to represent the SAM tail groups.
- Coarse-graining the head-group atoms into a uniform area density results in a softer repulsive potential allowing the interface to penetrate far deeper into the model surface than would be possible in the actual SAM. To rectify this, we apply a strongly repulsive linear potential in the half-space $z < R_0$, where R_0 is the radius of the head group's hard core. The repulsive potential is chosen to be $1\,k_{\text{B}}T$ when $1\,\text{nm}^2$ of interface penetrates the region $z < R_0$ by a "skin depth" δ.

The head groups are thus modeled by the following potential acting on a water molecule at position $\mathbf{r}$:

$$U_{\text{head}}(x, y, z \geq R_0) = \mu_{\text{head}} \int_{-\infty}^{\infty} \mathrm{d}x' \int_{-\infty}^{\infty} \mathrm{d}y'\, u_{\text{LJ}}\left(|\mathbf{r} - \mathbf{r}'|; \epsilon_{\text{head}}, \sigma_{\text{head}}\right)\Big|_{z'=0},$$

where $u_{\text{LJ}}(r; \epsilon, \sigma) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$ is the Lennard-Jones pair potential. Similarly, the effect of the tail groups is captured by

$$U_{\text{tail}}(x, y, z \geq R_0) = \rho_{\text{tail}} \int_{-\infty}^{\infty} \mathrm{d}x' \int_{-\infty}^{\infty} \mathrm{d}y' \int_{-\infty}^{-\zeta} \mathrm{d}z'\, u_{\text{LJ}}\left(|\mathbf{r} - \mathbf{r}'|; \epsilon_{\text{tail}}, \sigma_{\text{tail}}\right).$$

Finally, the repulsive wall is modeled by the potential

$$U_{\text{wall}}(x, y, z < R_0) = 2\rho_\ell^{-1} \cdot (1\,k_{\text{B}}T/1\,\text{nm}^2)(R_0 - z)/\delta,$$

where $\rho_\ell = 0.03333\,\text{Å}^{-3}$ is the number density of liquid water.

These smeared interaction potentials depend only on z, not on x or y. As described in the main text, we also scale the head-group interaction by a parameter η. Putting everything together, we obtain an explicit expression for the surface-interface interaction potential,

$$H'_{\text{surf}}[\{h_{\mathbf{a}}\}] = \rho_\ell \Delta^2 \sum_{\mathbf{a}} h'_{\text{surf}}(h_{\mathbf{a}}),$$

where

$$\begin{aligned} h'_{\text{surf}}(h_{\mathbf{a}}) &= \int_{h_{\mathbf{a}}}^{\infty} \mathrm{d}z\, \eta U_{\text{head}}(z) + U_{\text{tail}}(z) + U_{\text{wall}}(z), \\ &= \begin{cases} \eta\pi\epsilon_{\text{head}}\mu_{\text{head}}\sigma_{\text{head}}^3 \left[\frac{4}{45}(\sigma_{\text{head}}/z)^9 - \frac{2}{3}(\sigma_{\text{head}}/z)^3\right] & \\ \quad +\pi\epsilon_{\text{tail}}\rho_{\text{tail}}\sigma_{\text{tail}}^4 \left[\frac{1}{90}(\sigma_{\text{head}}/[z+\zeta])^8 - \frac{1}{3}(\sigma_{\text{head}}/[z+\zeta])^2\right], & z \geq R_0, \\ h'_{\text{surf}}(R_0) + \rho_\ell^{-1} \cdot (1\,k_{\text{B}}T/1\,\text{nm}^2)([R_0 - z]/\delta)^2, & z < R_0. \end{cases} \end{aligned}$$

To model the -CH$_3$-capped SAM, we chose the following values for the parameters

- The head groups are modeled as OPLS united-atom CH_3 groups interacting with SPC/E water, so $\sigma_{\text{head}} = 3.5355\,\text{Å}$ and $\epsilon_{\text{tail}} = 0.68976\,\text{kJ/mol}$.
- The tail groups are modeled as OPLS united-atom CH_2 groups (sp^3-hybridized) interacting with SPC/E water, so $\sigma_{\text{tail}} = 3.5355\,\text{Å}$ and $\epsilon_{\text{tail}} = 0.5664\,\text{kJ/mol}$.
- The tail region is inset from the plane of the head groups by a distance equal to a CH_2-CH_3 bond length ($1.50\,\text{Å}$), minus the van der Waals radius of a CH_2 group ($1.9525\,\text{Å}$), so $\zeta = -0.4525\,\text{Å}$.
- The head group density is known from the atomistic SAM geometry to be $\mu_{\text{head}} = 0.0462\,\text{Å}^{-2}$. The mass density of the SAM tails was estimated to be $935\,\text{kg/m}^3$ [39], resulting in a CH_2 group number density of $\rho_{\text{tail}} = 0.0402\,\text{Å}^{-3}$.
- The equivalent hard sphere radius R_0 of a -CH_3 group at room temperature was estimated to be $3.37\,\text{Å}$, as described in Section 4.3.1. It has a small temperature dependence, which we neglect.
- The wall skin depth δ was set to $0.1\,\text{Å}$, which is small enough so that the repulsive potential is essentially a hard wall at $z = R_0$, but large enough that we can propagate the interfacial dynamics with a reasonable timestep.

Umbrella sampling

Calculating $\Delta G_{\text{sol}}(V)$ from Equation (5.2) as a thermal average $\langle P_v(0)\rangle$ over Boltzmann-weighted configurations of $h(\mathbf{a})$ is impractical for large V. The configurations that dominate this average simply have a vanishingly small Boltzmann weight. To solve this problem, and in analogy to what we do in atomistic simulations, we perform umbrella sampling on the size of the sub-volume v of the probe cavity V that is above the interface.

We begin by defining the volume V corresponding to a probe cavity of dimensions $L \times L \times W$ as the set of points satisfying $|x|, |y| \leq L/2$ and $R_0 \leq z \leq R_0 + W$. We then define $v[\{h_{\mathbf{a}}\}]$ as the size of the sub-volume of V that is above the interface. Using umbrella sampling and MBAR (Section 3.1.1), we calculate the probability distribution for v, $P(v)$, down to $v = 0$. To do this, we use quadratic umbrellas defined by a center $\bar{v}$ and width $(\delta v)^2$, which result in the addition to the Hamiltonian of

$$H'_{\text{umb}}[\{h_{\mathbf{a}}\}] = k_{\text{B}}T\,\frac{(v[\{h_{\mathbf{a}}\}] - \bar{v})^2}{2(\delta v)^2}.$$

During each umbrella run, we also record the configurations $\{h_{\mathbf{a}}\}$ which yield each observed value of v. We then approximate the right-hand side of Equation (5.2) by summing over these configurations with appropriate weights, and obtain

$$\Delta G_{\text{sol}}(V) \approx -k_{\text{B}}T \ln \frac{\sum_{\{h_{\mathbf{a}}\}} P_v(0) P(v[\{h_{\mathbf{a}}\}])}{\sum_{\{h_{\mathbf{a}}\}} P(v[\{h_{\mathbf{a}}\}])},$$

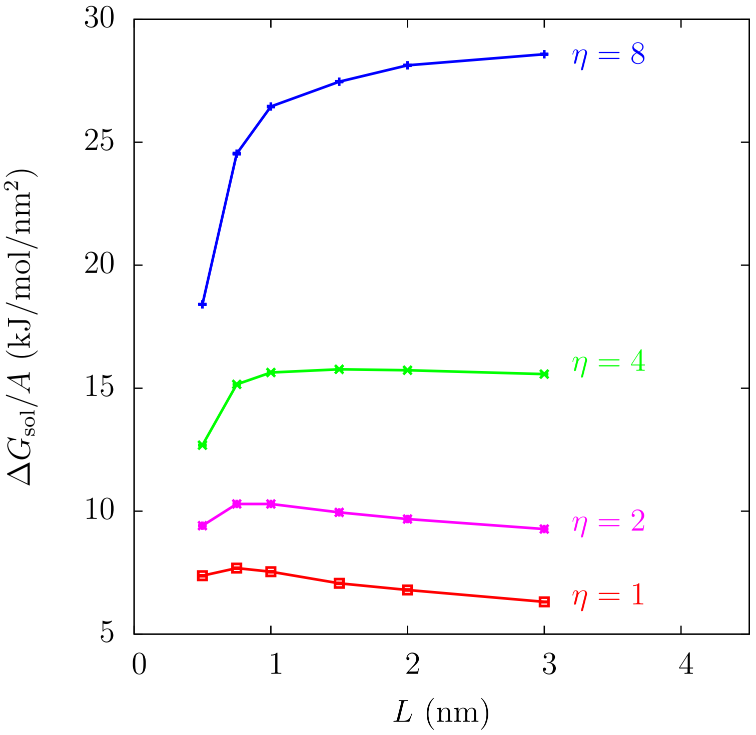


Figure 5.6: Solvation free energy per unit areas calculated using Equation (5.2) for the same cavities as in Figure 5.2.

where, as in the main text, the term $P_v(0)$ depends on the interface configuration $\{h_{\mathbf{a}}\}$, and the sum $\{h_{\mathbf{a}}\}$ is over all interface configurations in all the different umbrellas. To evaluate $P_v(0)$, we use the discrete versions of the integrals defining σ_v and $\langle N \rangle_v$ analogous to Equations (4.54) and (4.55).

5.2.3 Solvation free energies at surfaces

Figure 5.6, analogous to Figure 5.2, shows the solvation free energies calculated using Equation (5.2). The essential features of solvation next to the SAM surfaces are captured well by this model. This is particularly true for the hydrophobic surfaces (with η around 1), where the potential $U(\mathbf{r})$ closely mimics the effect of the real SAM on the adjacent water, and the agreement between theory and simulation is nearly quantitative. For the more hydrophilic SAMs, as stated above, the comparison is qualitative, because the simple form for $U(\mathbf{r})$ does not represent dipolar interactions well.

5.2.4 Temperature dependence of solvation free energies

Figure 5.7(a), analogous to Figure 5.3(a), shows that the observed trends in temperature dependence of ΔG_{sol} are captured nearly quantitatively by the simple model of Equation (5.2). In the model, the solvation free energies of the large probe volumes have large but mostly athermal contributions from the attractions between water and the model surface, since the average solvent density in the probe volume is nearly equal to its value in bulk. The main temperature-dependent contribution to ΔG_{sol} is the cost to deform the liquid-vapor-like interface near the surface to accommodate the large probe volume. Since the necessary deformation is similar, regardless of the hydrophobicity of the surface, the variation of ΔG_{sol} with temperature is similar as well.

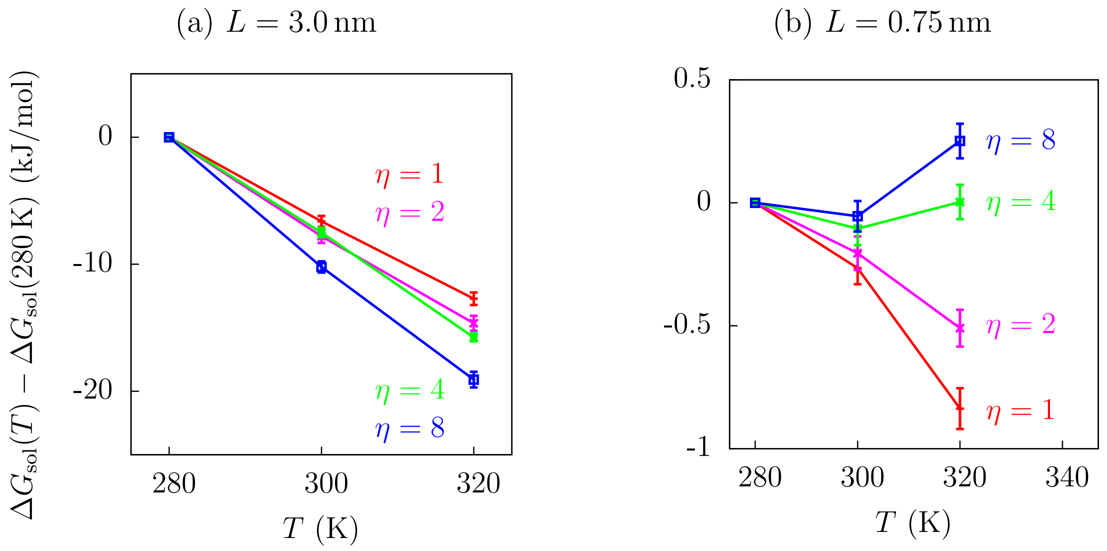
Figure 5.7(b), analogous to Figure 5.3(b), shows a similar comparison of temperature dependence of ΔG_{sol} for small probe volumes. The phenomenology of a crossover in sign is recovered by the simple model of Equation (5.2), though the correspondence is clearest at a slightly larger cavity size ($L = 0.75\,\text{nm}$). Near hydrophilic model surfaces, the interface is pulled close to the surface by a strong attraction, so it is costly to deform it. As a result, the cavity is emptied through bulk-like spontaneous density fluctuations, which is associated with a negative solvation entropy. In contrast, near a hydrophobic surface, the interface is easy to deform, which provides an additional mechanism for creating cavities. In fact, this mechanism dominates near sufficiently hydrophobic surfaces, and since the surface tension of water decreases with increasing temperature, so does ΔG_{sol}. Hence, even small cavities have a positive solvation entropy near hydrophobic surfaces. The continuous spectrum of negative to positive solvation entropies observed in Figures 5.3(b) and 5.7(b) is thus revealed to be a direct consequence of the balance between bulk-like water density fluctuations and liquid-vapor-like interfacial fluctuations, represented explicitly by the two factors in the integrand of Equation (5.2).

Figure 5.8 shows that our simple implementation of LCW ideas recovers many of the observed trends in solvation entropy. It is everywhere positive for the smallest attraction strength η, and a thermodynamic crossover length of just under 1 nm emerges for the more hydrophilic model surfaces, similar to that in bulk water. Nevertheless, the agreement between Figures 5.8 and 5.4 is somewhat qualitative, mostly as a result of the crude form of $U(\mathbf{r})$ used to model hydrophilic surfaces.

5.3 Discussion: hydrophobic surfaces are catalysts for unfolding

The results of our systematic study of solvation near surfaces expand upon the results of Section 3.2.3 in three important ways, which we now discuss.

First, our results reveal and explain the temperature dependence of the binding free energy of a hydrophobe to an extended surface. The binding free energies of the probe


(a) $L = 3.0\,\mathrm{nm}$

(b) $L = 0.75\,\mathrm{nm}$

Figure 5.7: Temperature dependence of solvation free energies in the model of Equation (5.2). Analogous simulation results are shown in Figure 5.3.

cavities to each SAM, as measured by the difference in ΔG_{sol} in bulk and next to the SAM (Figure 5.2) are all positive, so there is a net hydrophobic attraction between the cavities and the SAMs. As expected from Section 3.2.3, the binding strength is correlated with the SAM's hydrophobicity. However, the temperature dependences shown in Figure 5.3 reveal that when the temperature rises, small objects will bind more strongly to the SAMs, but the binding free energy of large objects remains mostly unchanged. At heart, this difference arises because solvation of even small objects is dominated by interface formation. Because we have rationalized this result, we can make non-trivial predictions about the effect of other perturbations on binding free energies. For example, adding table salt to the water raises the surface tension of the solution (because the small, charged ions are repelled from the interface) but makes little difference to the probabilities of cavity formation (at 1 M salt concentration, there are, on average, only 0.1 ions in a $5 \times 5 \times 3\,\text{Å}$ cavity). Hence, we expect that adding salt to the solution would reduce the binding free energy of small objects to a hydrophobic surface. Similar changes can be confidently predicted for the conclusions that follow.

Second, our results show that the force for hydrophobic assembly is weakened near a surface. When n monomers are dispersed, their total solvation free energy scales linearly with n. When aggregated, in bulk or near a surface, their total solvation free energy scales sublinearly with n. Hence, for large enough n, the free energy of assembly is proportional to the solvation free energy of the isolated monomers. Since ΔG_{sol} is smaller for monomers near hydrophobic surfaces than in bulk, so is the free energy of assembly there. The difference

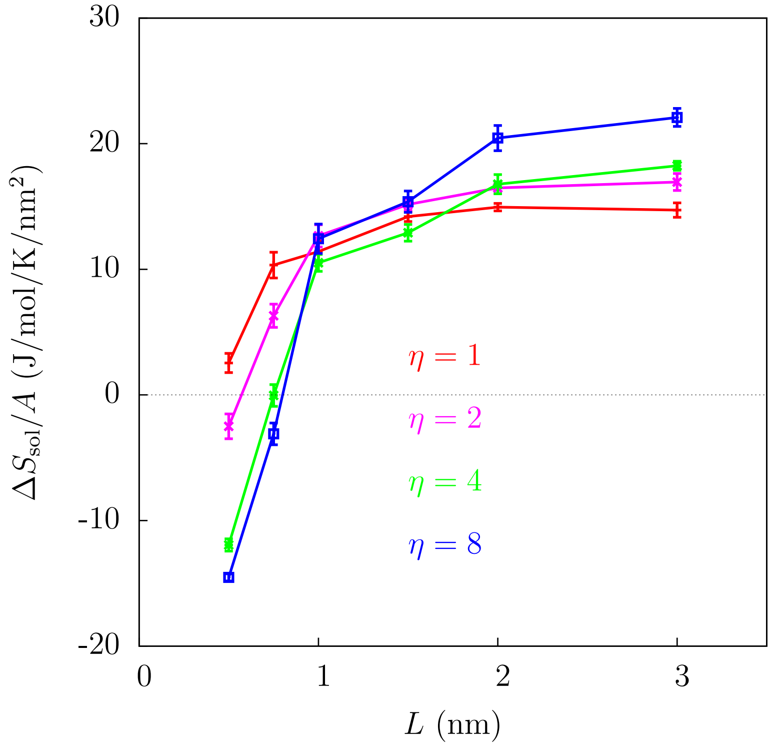


Figure 5.8: Solvation entropy per unit area at $T = 300\,\mathrm{K}$, deduced from the model of Equation (5.2). Analogous simulation results are shown in Figure 5.4.

arises from interfacial physics dominating small-object solvation at surfaces, as rationalized by the LCW-like model. We thus expect our observation about the free energy of assembly to hold generically at any surface and with a variety of monomers, not simply the surfaces and cavities we have explicitly investigated numerically.

Third, the free energy barrier separating the dispersed and assembled states differs significantly in bulk and at surfaces. In bulk, the principal free energy barrier to assembly is the nucleation of a liquid-vapor interface in the solvent (Figure 2.4). At a surface, there is a pre-existing liquid-vapor interface solvating the dispersed monomers, which can be deformed continuously to solvate the assembled state. We thus expect no solvent-dominated free energy barrier to assembly.

Since the relative stability of the assembled and dispersed states is shifted towards the dispersed state at a surface, and the free energy barrier between these two states is removed, we suggest that hydrophobic surfaces can generically acts as catalysts for disassembly of structures that are assembled in bulk. In the context of protein folding, our findings suggest that hydrophobic surfaces may generically catalyze the unfolding of proteins.

Our suggested catalytic activity of hydrophobic surfaces may play a role in chaperonin function [31]. The interior walls of chaperonins in the open conformation are hydrophobic and can bind misfolded proteins, whereupon their unfolding is catalyzed [27,57]. Subsequent ATP-driven conformational changes render the chaperonin walls hydrophilic [27,31]. As a result, the unfolded protein is released from the wall, as the free energy for a hydrophobe to bind to a hydrophilic surface is much lower than that to bind to a hydrophobic one (Figure 5.2).

Chapter 6

A microscopic view of water evaporation

In this chapter, we examine the process of water evaporation in molecular detail. Our goal is to determine whether evaporation is an activated process, and if so, characterize the relevant barriers. In a sample of water at equilibrium with its vapor, the rate of evaporation is equal to the rate of condensation. If every vapor molecule that hit the liquid surface condensed, then the rate of condensation would simply be the rate at which vapor particles hit the liquid surface. This collision rate, which can be estimated by treating the vapor as an ideal gas, places an upper bound on the evaporation rate. To the extent that the two differ, there is a barrier to evaporation (or condensation) beyond the mere cohesive strength of the liquid.

We are inspired by a set of water evaporation rate measurements carried out at Berkeley, themselves motivated by the importance of water evaporation and condensation in atmospheric science and climate prediction. Surprisingly, measurements of the evaporation rate of pure water have been controversial, with estimates varying by as much as 3 orders of magnitude over the past century [25], and converging to within 1 order of magnitude in the last decade [24]. The two main difficulties faced by experimenters are the effect of impurities in the water samples, and avoiding or accounting for recondensation of the evaporated molecules. The Saykally and Cohen groups have attempted these measurements using Raman thermometry of a continuous jet of liquid microdroplets. They have examined H_2O and D_2O, with and without dissolved salts, under carefully controlled conditions [22–24, 99], ingeniously selected to avoid both the impurity and recondensation problems, so we suspect that their results are among the most trustworthy in the literature. Evaporation rates are usually reported as an evaporation coefficient, defined as the ratio of the evaporation rate to the ideal gas collision rate. The Saykally and Cohen groups report an evaporation coefficient of $62 \pm 9\,\%$.

To uncover the molecular origins of this result through simulation, we use transition path sampling [7] to harvest a statistically representative collection of short trajectories of rare evaporation events, with Boltzmann-distributed initial conditions and energy-conserving dynamics. An alternate strategy pursued by others is to examine the complementary process, condensation, which is not rare at all [110]. However, this introduces the uncontrolled approximation that the velocities and angular momenta of the evaporated water molecule are Boltzmann-distributed, with a temperature equal to that of the liquid. Our approach sidesteps this issue by specifying the initial ensemble of the liquid, with the velocity and angular momentum distributions of the evaporated water molecules being an output of the calculation. The result, not knowable *a priori*, is that the velocities and angular momenta are indeed Boltzmann-distributed. As such, the fact that others have observed near-unity condensation coefficients in simulations suggests that the evaporation coefficient for water is also close to unity in simulations, at odds with the experiments of the Saykally and Cohen groups.

To explore the details further, we examine properties of the harvested evaporation trajectories to determine what barriers, if any, are being overcome. In particular, we estimate the positions of the transition states along each trajectory, and project these transition states onto various possible reaction coordinates. We find that the most relevant coordinates controlling evaporation are the distance of the water molecule from the water surface and its component of velocity along the average surface normal. Motivated by this find, we measure the effective free energy of a water molecule as a function of its distance from the water surface, i.e., the effective cohesive potential of the liquid. Along this coordinate, we find no significant barrier. Regarding the free energy approximately as a potential energy surface, we further show that transition states cluster around states of threshold energy. This suggests that evaporation can usefully be thought of as a thermal escape from a deep, barrierless potential well.

A barrierless evaporation process can nonetheless be characterized by the collective fluctuations observed in the rare configurations visited during evaporation (not rare during condensation). In particular, we find that the mean curvature of the liquid-vapor interface correlates with the distance of the evaporated water molecule from the surface. This finding might be of relevance if the curvature of the interface could be controlled externally.

Having found that our simulations disagree qualitatively from the experimental results, we critically examine the experiments. We find, in particular, that the Raman thermometry used relies on a linear extrapolation of calibration data into supercooled temperatures, and that the extrapolation need not be accurate. We suggest that the deviation from linearity may have been observed before, and that only a small deviation from linearity would be needed to reconcile the experimental measurements with an evaporation rate equal to the ideal-gas collision rate. Our suggestions motivate both repeating the experiments at higher temperature and extending the calibration procedure down to supercooled temperatures, so that the range of temperatures observed in the experimental data matches that of the calibration data. Both suggestions are currently being implemented by our experimental colleagues.

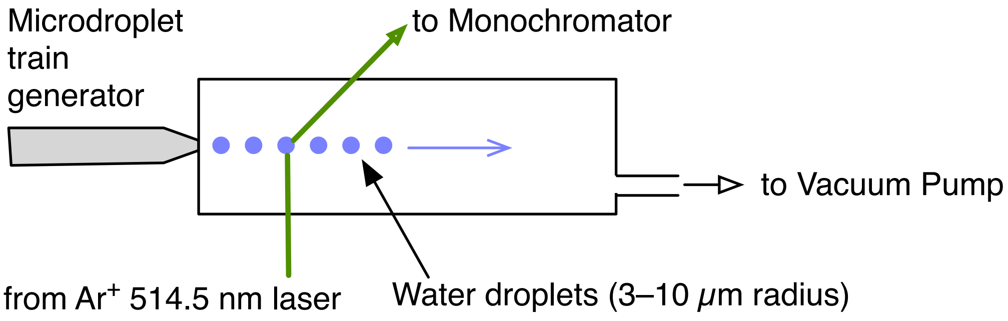


Figure 6.1: Simplified schematic of experimental setup for measuring evaporation rate of water. Adapted from Ref. [99].

6.1 Preliminaries

6.1.1 Experimental motivation

The full details of the Saykally and Cohen groups' experiment are given in Ref. [99]. Here, we summarize the aspects that are relevant to our study. A simplified schematic of the experimental apparatus is shown in Figure 6.1. The experimenters inject a continuous train of liquid microdroplets of tunable radius, r_0, ranging from 3 to 10 μm $\pm 0.1\,\mu$m, into a chamber that is continuously evacuated to a pressure of less than 5×10^{-4} Torr, or about 7×10^{-7} atm. At the end of the chamber, a liquid nitrogen pool collects and freezes the droplets, so as to maintain the chamber vacuum. The droplets are injected with uniform velocity, and the center-to-center distance between consecutive droplets is $6r_0$. As a droplet progresses through the vacuum chamber, it undergoes evaporative cooling, as described below. At any point along the droplet's flight, its temperature is measured indirectly by measuring its total Raman scattering intensity, $I(\omega)$, as a function of angular frequency ω. This spectrum is a useful thermometer because it varies sensitively with temperature. The end result of the experiment is the average temperature of a microdroplet as a function of time elapsed since injection. From this cooling curve, the rate of evaporation can be extracted, as described below. A typical cooling curve is shown in Figure 6.2.

The flux of evaporating water molecules at the droplet surface, J_e, can be expressed as a fraction γ, called the "evaporation coefficient", of the ideal-gas collision rate, $J_e^{\max}$,

$$J_e = \gamma J_e^{\max}, \tag{6.1}$$

where $J_e^{\max}$, given by

$$J_e^{\max} = \frac{P_{\text{vap}}}{\sqrt{2\pi m k_{\mathrm{B}} T}}. \tag{6.2}$$

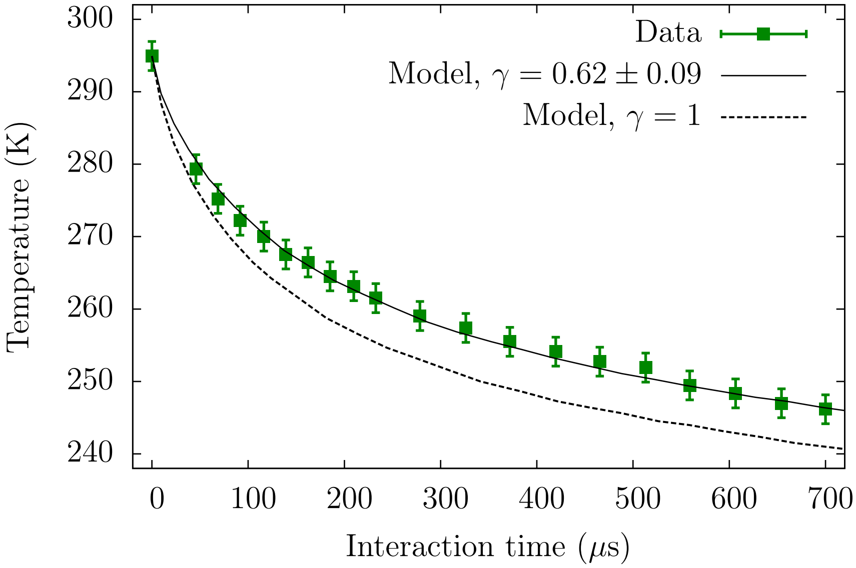


Figure 6.2: Example cooling curve of H_2O measured with the setup of Figure 6.1. Adapted from Ref. [99].

Here, P_{vap} is the vapor pressure of water at a given temperature, and m is the mass of a water molecule. This result follows from calculating the flux of mass-m particles of an ideal gas at pressure P_{vap} and temperature T through a flat surface. The substantial temperature and isotope dependence of $J_e^{\max}$ reflect thermodynamic properties of the liquid and the vapor, whereas the evaporation coefficient γ reveals information about barriers to evaporation.

An evaporation coefficient can be deduced from a cooling curve by modeling the evaporation process. Each evaporating water molecule removes from the droplet an amount of energy equal to, on average, ΔH_{vap}, the enthalpy of vaporization of water[1]. On average, this evaporation event lowers the temperature of the droplet by an amount $\Delta T = \Delta H_{\text{vap}}/C_P M$, where C_P is the constant-pressure specific heat capacity of water and M is the mass of the droplet. Hence, the rate of change of droplet temperature is given by

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\gamma J_e^{\max} \frac{\Delta H_{\text{vap}}}{C_P M}. \tag{6.3}$$

---

[1] Strictly speaking, the per-molecule enthalpy on vaporization, $\Delta H_{\text{vap}} \approx 16\,k_{\text{B}}T$ at $T = 300\,\text{K}$, includes the work needed to increase the volume of the system by an amount $1/\rho_g$ against the external pressure p, where ρ_g is the number density of the gas at pressure p. This work, equal to $p/\rho_g = k_{\text{B}}T$ at any pressure for an ideal gas, is absent in the experiment, so should not be included in the model of Equation (6.3). Implementing this correction results in an estimate of the rate of evaporation that is about 7 % larger than reported in Ref. [99].

The only unknown in this differential equation is the evaporation coefficient γ. If γ is assumed constant, the experimental cooling curves may be fit numerically to the cooling curves predicted by Equation (6.3). The main result of Ref. [99] is that for pure H_2O,

$$\gamma = 0.62 \pm 0.09. \tag{6.4}$$

This result is actually derived from a more elaborate model that accounts for a small (about 4° C) radial temperature gradient arising in the droplet, coupled to evaporation through thermal diffusion. However, Smith et al. [99] remark that this elaboration on the simple cooling curve model of Equation (6.3), and others such as introducing an Arrhenius dependence of γ on temperature, do not change the result appreciably.

Before proceeding to describe our theoretical investigations, we dwell on one aspect of the experiment in further detail. In particular, Raman thermometry relies on the Raman scattering intensity, $I(\omega)$, exhibiting so-called van't Hoff behavior [98]. That is, for an arbitrary frequency ω^*, the quantity $\ln\big(\int_{-\infty}^{\omega^*} \mathrm{d}\omega\, I(\omega) / \int_{\omega^*}^{\infty} \mathrm{d}\omega\, I(\omega)\big)$ is approximately linear in inverse temperature, $1/T$, around a small temperature range. An example set of calibration data from Ref. [99] is shown in Figure 6.3. As shown by Geissler [36], van't Hoff behavior is expected in any spectra that suffers from inhomogeneous broadening, and the linear relation to inverse temperature holds only approximately over a small enough temperature range, such as 0° C to 100° C for water [98]. It is important to note that while Smith et al. calibrated their Raman thermometer over the temperature range 0° C to 50° C, most of the observed droplet temperatures are below 0° C (Figure 6.2).

6.1.2 Transition path sampling

We now turn to our simulation study of evaporation. Since evaporation is a rare event, happening about once per 1 ns per 1 nm^2 at ambient conditions, we need special techniques to collect a statistically relevant number of examples of evaporation. The alternate strategy of running a single simulation for an extended period of time is computationally expensive, and samples only an anecdotal number of events at undesirable conditions that favor evaporation (e.g., in Ref. [75], 70 evaporation events were collected from a simulation of water at 350 K). The strategy we pursue here is to use transition path aampling (TPS), a method we now summarize.

To sample static configurations of a system that are Boltzmann-distributed, one can use the Metropolis algorithm, which implements a biased random walk through configuration space with a move set and acceptance criteria chosen to sample the Boltzmann distribution. Similarly, transition path sampling implements a biased random walk in trajectory space, with a move set and acceptance criteria chosen to sample a specific path ensemble [7]. Specifically, let $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_T\}$ denote the configurations (positions and momenta) along a dynamical trajectory of length T. To each trajectory, we assign a path weight $\mathcal{P}[\mathbf{X}]$ given by the initial ensemble of the system and its dynamics. For example, a system in the

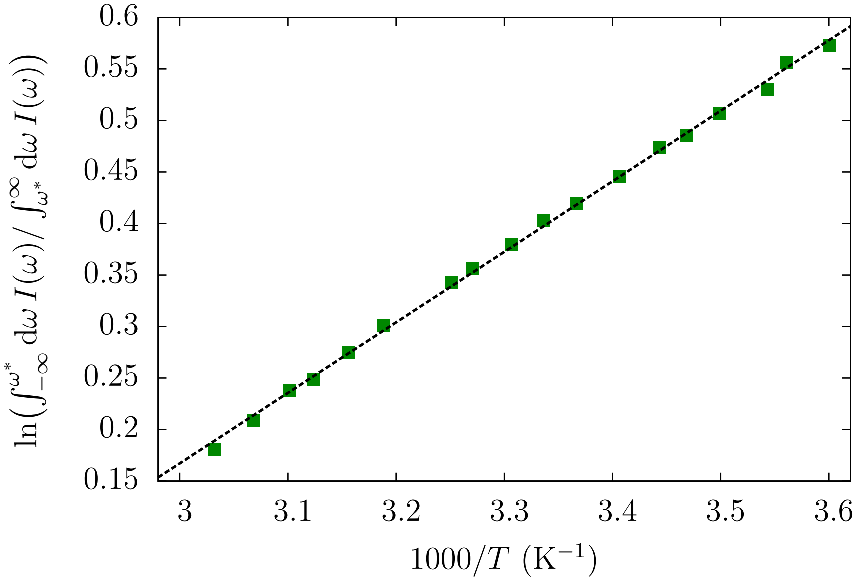


Figure 6.3: Calibration curve for Raman thermometry. The Raman scattering intensity, $I(\omega)$, at angular frequency ω is measured experimentally. The temperature of the water is then extracted by using the approximate linear relation between $\ln\left(\int_{-\infty}^{\omega^*} \mathrm{d}\omega\, I(\omega) / \int_{\omega^*}^{\infty} \mathrm{d}\omega\, I(\omega)\right)$ and inverse temperature $1/T$. The calibration temperatures here range from 4.4° C ($1000/T = 3.60\,\mathrm{K}^{-1}$) to 56.5° C ($1000/T = 3.03\mathrm{K}^{-1}$). Adapted from Ref. [99].

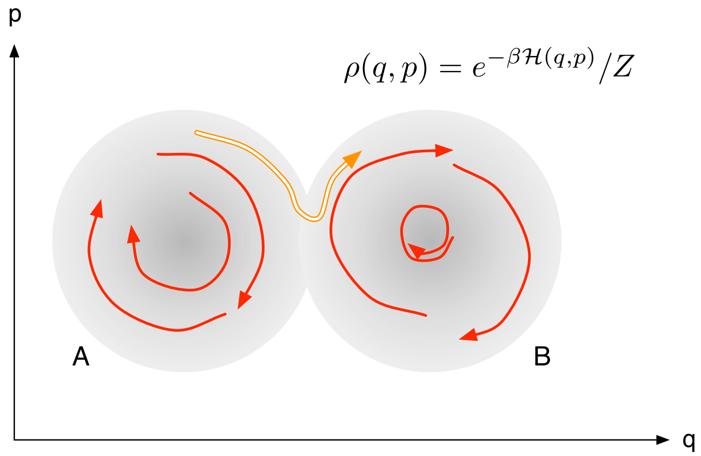


Figure 6.4: Schematic representation of the path ensemble $\mathcal{P}_0[\mathbf{X}]$ in a one-dimensional system. The intensity of the grey background at any point in phase space is proportional to its Boltzmann weight. Typical paths of length T are shown as traces over this phase space. When there is a natural partition of phase space into two basins, A and B, most short trajectories remain in one of these basins. Only the trajectories connecting A and B (orange) are members of the restricted path ensemble, $\mathcal{P}_{AB}[\mathbf{X}]$ (Equation (6.10)).

canonical ensemble propagated with Newtonian dynamics has a path weight functional

$$\mathcal{P}_0[\mathbf{X}] = \frac{e^{-\beta\mathcal{H}(\mathbf{x}_1)}}{Z} \prod_{i=1}^{T-1} \delta(\mathbf{x}_{i+1} - \mathcal{G}_1 \mathbf{x}_i), \tag{6.5}$$

where $\mathcal{G}_1$ is the propagator that evolves the configuration of the system for one time unit (e.g., one step in the Velocity Verlet algorithm). This path weight is chosen to integrate to unity over all possible trajectories. Its value is the probability that an arbitrary subtrajectory of length T extracted from a long trajectory of the system is identical to $\mathbf{X}$. As with the Boltzmann weight, the choice of path weight $\mathcal{P}[\mathbf{X}]$ defines a path ensemble. A schematic path ensemble is shown in Figure 6.4.

Given an "old" path $\mathbf{X}^{(\mathrm{o})}$, we can construct a related "new" path $\mathbf{X}^{(\mathrm{n})}$ in this path ensemble in many different ways, called "moves" in analogy to typical Monte Carlo moves. Two particularly simple and efficient moves are *shifting* and *shooting*. In a shifting move, the old path is altered by advancing or regressing it in time, in much the same way that a snake propagates on the ground: an initial or final end of the path is discarded, and the system is evolved dynamically at the other end to construct a new path of length T. In a shooting move, a time point along the path is picked, the system's configuration is slightly perturbed, and the system's dynamics are evolved forwards and backwards in time to construct a new path of length T.

As in the Metropolis algorithm, we can use moves to construct a biased random walk in trajectory space. Let $P(\mathrm{o} \to \mathrm{n})$ be the probability that a single step in this random walk starting with a path $\mathbf{X}^{(\mathrm{o})}$ yields the path $\mathbf{X}^{(\mathrm{n})}$, and define $P(\mathrm{n} \to \mathrm{o})$ analogously. The trajectories observed in this random walk occur with a probability given by the path ensemble $\mathcal{P}[\mathbf{X}]$ if the following detailed balance condition [34] is satisfied:

$$\mathcal{P}[\mathbf{X}^{(\mathrm{o})}]P(\mathrm{o} \to \mathrm{n}) = \mathcal{P}[\mathbf{X}^{(\mathrm{n})}]P(\mathrm{n} \to \mathrm{o}). \tag{6.6}$$

The probability $P(\mathrm{o} \to \mathrm{n})$ can be expressed as a product, $P_{\mathrm{gen}}(\mathrm{o} \to \mathrm{n})P_{\mathrm{acc}}(\mathrm{o} \to \mathrm{n})$, of a generation probability, $P_{\mathrm{gen}}(\mathrm{o} \to \mathrm{n})$, which one is free to choose, and an acceptance probability, $P_{\mathrm{acc}}(\mathrm{o} \to \mathrm{n})$, chosen to satisfy the detailed balance condition. If the generation probabilities are chosen symmetrically, i.e., $P_{\mathrm{gen}}(\mathrm{o} \to \mathrm{n}) = P_{\mathrm{gen}}(\mathrm{n} \to \mathrm{o})$, as is usually the case, then the usual Metropolis acceptance probability guarantees detailed balance:

$$P_{\mathrm{acc}}(\mathrm{o} \to \mathrm{n}) = \min\left[1, \frac{\mathcal{P}[\mathbf{X}^{(\mathrm{n})}]}{\mathcal{P}[\mathbf{X}^{(\mathrm{o})}]}\right]. \tag{6.7}$$

In the path ensemble $\mathcal{P}_0[\mathbf{X}]$, a shifting move has a trivial acceptance probability,

$$P_{\mathrm{acc}}^{\mathrm{shift}}(\mathrm{o} \to \mathrm{n}) = 1, \qquad \text{(path ensemble: } \mathcal{P}_0[\mathbf{X}]\text{)}. \tag{6.8}$$

The acceptance probability of a shooting move only depends on the configuration at the shooting point, denoted here by s:

$$P_{\mathrm{acc}}^{\mathrm{shoot}}(\mathrm{o} \to \mathrm{n}) = \min\Big[1, \exp\big\{-\beta(\mathcal{H}\big(\mathbf{x}_s^{(\mathrm{n})}\big) - \mathcal{H}\big(\mathbf{x}_s^{(\mathrm{o})}\big))\big\}\Big], \qquad \text{(path ensemble: } \mathcal{P}_0[\mathbf{X}]\text{)}. \tag{6.9}$$

Though one can sample paths from the unrestricted path ensemble $\mathcal{P}_0[\mathbf{X}]$ with the above algorithm, doing so is not very useful, since a long trajectory samples the same trajectories with similar efficiency. However, we can define a path ensemble that is restricted to trajectories, or transition paths, that connect two distinct regions of phase space, commonly denoted A and B. For a system with equal ensemble and dynamics to that used to define $\mathcal{P}_0[\mathbf{X}]$, the restricted path ensemble weight, $\mathcal{P}_{AB}[\mathbf{X}]$, is given by

$$\mathcal{P}_{AB}[\mathbf{X}] = h_A(\mathbf{x}_1)\frac{e^{-\beta\mathcal{H}(\mathbf{x}_1)}}{Z_{AB}}\prod_{i=1}^{T-1}\delta(\mathbf{x}_{i+1} - \mathcal{G}_1\mathbf{x}_i)h_B(\mathbf{x}_T), \tag{6.10}$$

where $h_A(\mathbf{x})$ and $h_B(\mathbf{x})$ are indicator functions that are 1 if $\mathbf{x}$ is in A or B, respectively, and 0 otherwise. The normalization constant Z_{AB} ensures that the path weight integrates to unity over all possible paths. This change in ensemble results in analogous changes to the shifting and shooting acceptance probabilities. The resulting random walk efficiently samples trajectories of length T that start at A and end in B. If A and B are separated by a large free energy maximum, so that short paths that cross from A to B are rare, then sampling these trajectories with TPS is orders of magnitude more efficient than running a long simulation and waiting for the crossing to happen spontaneously.

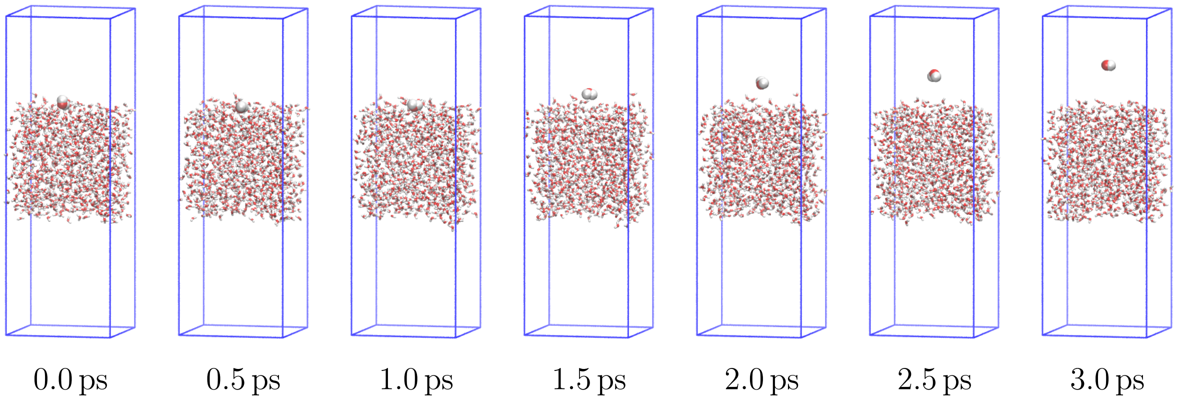


Figure 6.5: Snapshots of a 3 ps-long initial evaporation trajectory, used to start the TPS random walk through trajectory space. See text for details.

6.1.3 Simulation setup

To examine water evaporation, we set up a $30 \times 30 \times 30$ Å^3 simulation box in LAMMPS [89] filled with SPC/E water (about 900 molecules) at a number density 0.0333 waters/Å^3 (mass density 0.997 g/ml). Lennard-Jones interactions are truncated and shifted at a distance of 10 Å. Electrostatic interactions are calculated using the particle-particle particle-mesh (PPPM) method [45]. The geometry of each water molecule is kept fixed using the SETTLE algorithm [80]. The system is evolved using Langevin dynamics, with a timestep of 2 fs and a velocity damping time constant of 3 ps.

After an initial equilibration of 50 ps, the size of the simulation box in the z direction is tripled, resulting in a $30 \times 30 \times 90$ Å^3 system with a slab of water at its center. This system is then equilibrated again for another 50 ps, using a Langevin thermostat with velocity damping time constant of 2 ps. Finally, to generate an initial evaporation trajectory in this system, we remove the thermostat, create a condensation trajectory, and then reverse this trajectory in time. In detail, we add a new water molecule to the system at a point that is 15 Å above the top of the water slab, with a thermal momentum in the negative z-direction, no momentum in the x or y directions, and zero angular momentum. We then evolve this trajectory for 3 ps in the absence of a thermostat. We check that the water molecule has indeed condensed (see below): if it has not, we extend the trajectory by up to 6 ps more, until condensation occurs, and use the last 3 ps of the extended trajectory. In principle, it is possible for this procedure to fail if the inserted water "bounces" off of the water slab and we never considered it to have condensed, but in practice, we have not seen this happen. Finally, we reverse the condensation trajectory in time to obtain an initial evaporation trajectory, which we use to start the TPS random walk in trajectory space. An example initial trajectory is depicted in Figure 6.5.

To perform a TPS study, the initial and final states of the rare event being sampled must be precisely defined. Here, basin A, encompassing the fully "condensed" configurations of the system, consists of all configurations where every water that is not hydrogen-bonded to any other water[2] is at most 4 Å away from the nearest water[3]. Basin B, where exactly one water molecule is indisputably detached from the bulk liquid, is defined as all configurations of the system where there is exactly one water molecule with no hydrogen-bonding partner that is more than 8 Å away from its nearest neighbor. While the 4 Å distance used to define basin A is merely convenient, the 8 Å distance used to define basin B is critical. We have noted that an isolated water molecule that is less than about 8 Å away from its nearest neighbor will spontaneously recondense into the slab of liquid water. This observation is rationalized in Section 6.2.2.

The TPS random walk is performed as follows. At every step, we choose to perform a shifting move 90 % of the time, and a shooting move 10 % of the time, reflecting the low cost of shifting versus shooting. In a shifting move, we shift the trajectory forwards or backwards by a time Δt uniformly distributed between -1 and 1 ps. Shooting moves are performed as in the appendices of Refs. [38] and [37]. Briefly, a shooting point is picked uniformly along the length of the trajectory. Then, to each component of velocity of each atom of mass m is added a Gaussian perturbation of zero mean and variance $c_1^2 \times k_{\mathrm{B}}T/m$, where c_1 is a small parameter used to tune the size of the shooting perturbation, here chosen to be 0.5. The components of the resulting velocities along fixed bond directions are then removed. The center-of-mass velocity is also set to $\mathbf{0}$. The velocities are then rescaled to yield the same kinetic energy as before the shooting move. This portion of the move is symmetric, and amounts to a random rotation of the $3N$-dimensional mass-weighted velocity vector of the system that respects the bond constraints of each water molecule. Finally, we attempt to change the kinetic energy from K to K' by adding to it a Gaussian perturbation of zero mean and variance $c_2^2 \times k_{\mathrm{B}}T \times d/2$, where d is the number of degrees of freedom in the system (here equal to $6N_w - 3$, with N_w equal to the number of waters) and c_2 is a small parameter used to tune the size of the kinetic energy change in a shooting move, here chosen to be 0.5 as well. The change in kinetic energy is accepted with probability

$$P[K \to K'] = \min\left[1, e^{-\beta(K'-K)}\left(\frac{K'}{K}\right)^{d/2-1}\right]. \tag{6.11}$$

When the new configuration is generated with this procedure, the generation probabilities $P_{\mathrm{gen}}(\mathrm{o} \to \mathrm{n})$ and $P_{\mathrm{gen}}(\mathrm{n} \to \mathrm{o})$ are not symmetric, due to phase space factors. It can be shown [37] that detailed balance (Equation (6.7)) is satisfied if the acceptance probability

---

[2] Two waters are considered hydrogen bonded if the distance between their oxygen atoms is below 3.5 Å and the angle between the OH bond of the donor and the line connecting the two oxygen atoms is below 30°, as was done in Ref. [69]. For our purposes, any other reasonable definition of a hydrogen bond should yield nearly identical results.

[3] The "position of a water" means the position of its oxygen atom, unless otherwise stated.

of such asymmetric shooting moves is 1, provided that the new trajectory starts in basin A and ends in basin B.

For each set of initial conditions, we performed 10,000 TPS steps, recording a trajectory every 100 TPS steps. Each recorded trajectory is reasonably independent of the previous one. To further improve the sampling, we repeated the entire procedure outlined in this section 40 times. The final outcome of this exercise is a set of thousands of mostly uncorrelated trajectories of evaporation events, with initial conditions drawn from a canonical ensemble at temperature 300 K and evolved in time with energy-conserving Newtonian dynamics.

6.2 Results

We now analyze the properties of the collected evaporation trajectories. We estimate the value of each observable independently in each of the 40 full TPS runs. The reported estimate is the mean of these values, and the error bar is the standard error of the mean.

We first consider the time that it takes for evaporation events to occur, to verify that a TPS trajectory length of 3 ps is sufficiently large. For each trajectory, let t_A be the largest time for which $\mathbf{x}(t_A)$ is in basin A, and let t_B be the largest time for which $\mathbf{x}(t_B)$ is not in basin B. These times roughly characterize the points along the trajectory at which the evaporation rare event begins and ends. Figure 6.6 shows the distribution of the time difference $t_B - t_A$. Most evaporation events take under 1 ps, and very few take just under 3 ps. Hence, the 3 ps trajectory length we chose to use for our TPS sampling is long enough. Correcting for the bias towards short evaporation events owing to their larger number of possible starting times does not change this conclusion. This is demonstrated in Figure 6.7, which shows the distributions of times t_B. If the TPS trajectory length is sufficiently long, then this distribution should rise from zero at small t_B and plateau to a constant for t_B much larger than the typical time for an A-to-B transition to occur. This is indeed observed. Were the TPS trajectory length too short, there would be no plateau region.

For all subsequent results, we consider only the trajectories for which $t_B > 2\,\text{ps}$, i.e., the plateau region of Figure 6.7. This choice avoids any biases towards unusually short evaporation trajectories that are overrepresented in the ensemble of trajectories with $t_B \ll 2\,\text{ps}$.

6.2.1 Post-evaporation momenta are Boltzmann-distributed

Here, we examine the center-of-mass velocities and angular momenta at the end of each trajectory. Figure 6.8 shows the distributions of the component of the center of mass velocity along a direction perpendicular to $\hat{\mathbf{z}}$. These velocities are consistent with a Boltzmann distribution at temperature $T = 300\,\text{K}$.

Figure 6.9 shows that, similarly, the distribution of the observed components of angular momenta along the principal axes of inertia of the evaporating water molecule are compatible

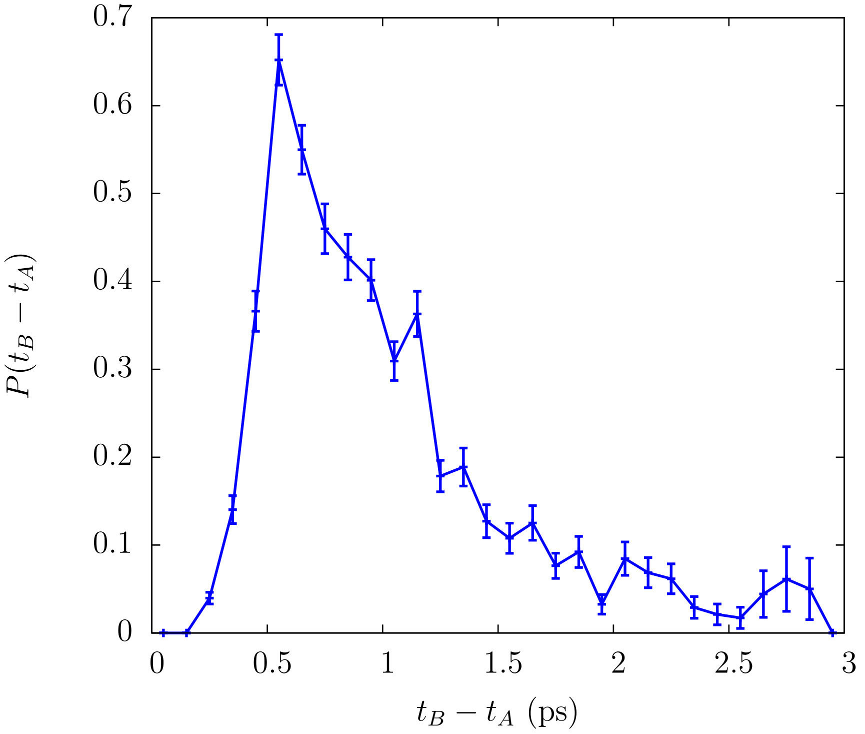


Figure 6.6: Distribution of evaporation event durations.

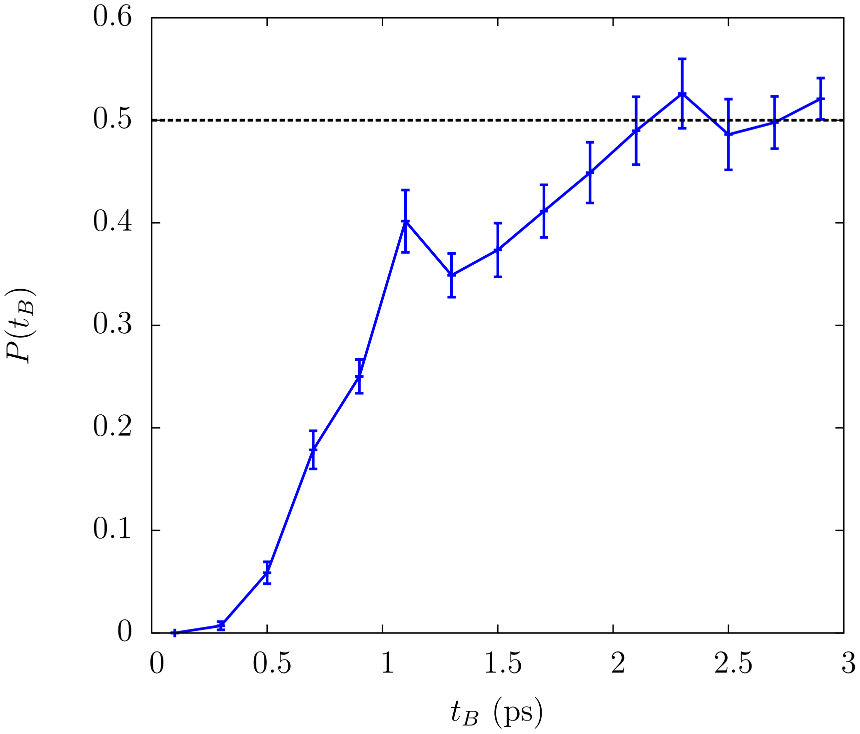


Figure 6.7: Distribution of times at which the evaporation event completes.

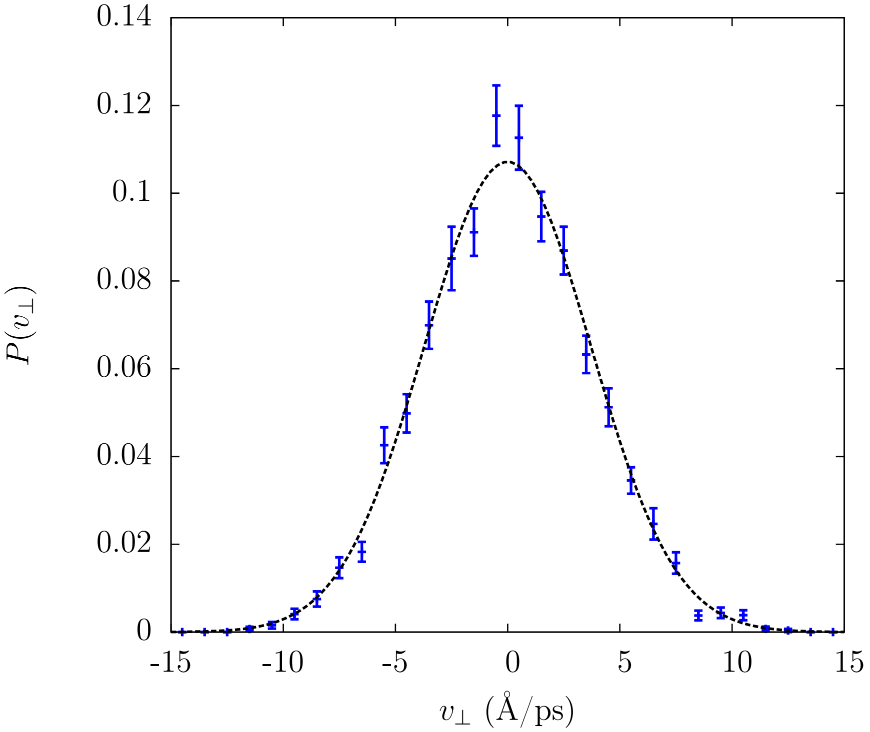


Figure 6.8: Distribution of the component of center-of-mass velocity perpendicular to $\hat{\mathbf{z}}$ (error bars). For comparison, the corresponding Boltzmann distribution at temperature $T = 300\,\mathrm{K}$ is shown (dashed line).

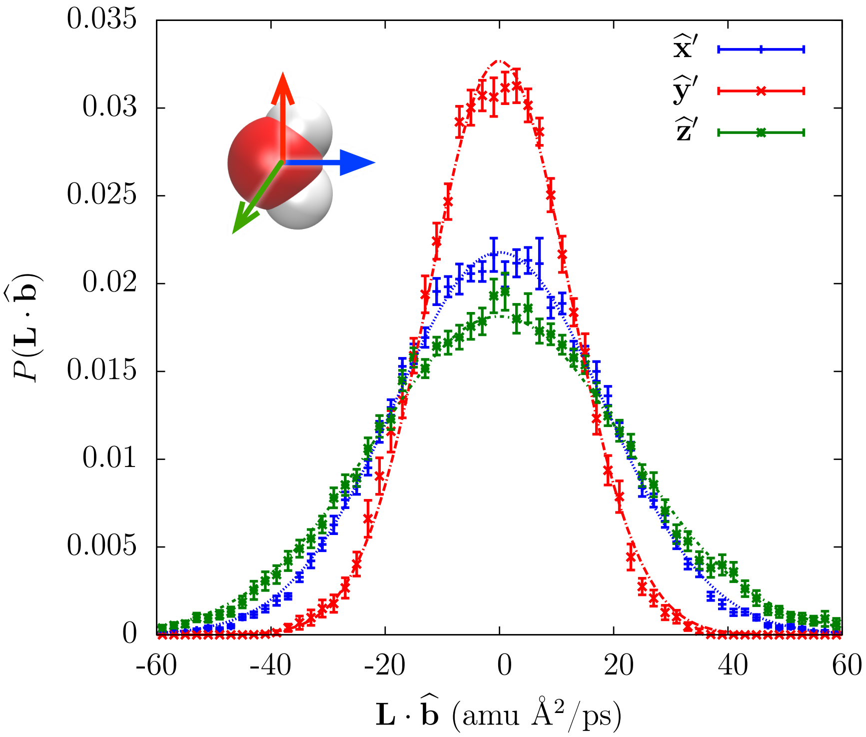


Figure 6.9: Distribution of the components of angular momentum of the evaporated water along the principal axes of inertia (error bars). For comparison, the corresponding Boltzmann distributions at temperature $T = 300\,\mathrm{K}$ are shown (dashed line). Inset: principal axes of inertia of a water molecule.

with a Boltzmann distribution at $T = 300\,\mathrm{K}$.

The component of the velocity along the z direction has a more interesting distribution, shown in Figure 6.10. By construction, water molecules evaporate in the $+\hat{\mathbf{z}}$ direction, so no water molecule should have a negative v_z if the definition of basin B were sufficiently strict[4], Naively, one might thus expect the distribution of v_z to be Boltzmann for $v_z > 0$, unlike what is actually observed in Figure 6.10.

To get a sense for the expected distribution, we examine the simpler situation, depicted in Figure 6.11, of thermal ideal gas particles evaporating from a deep, barrierless potential well of depth ΔU. Particles inside the well have a thermal distribution of velocities, $P(v_i)$,

---

[4] In practice, the definition of basin B used here does not perfectly discriminate between the evaporated states and states where recondensation will occur. Since the trajectories examined here are finite, a trajectory where the system that transiently enters B before recondensing may appear as an evaporation event, but with $v_z < 0$ at the end of the trajectory. Only about 1% of our trajectories exhibit this problem, which can in principle be mitigated by using longer trajectories and a stricter definition of basin B.

given by

$$P(v_i) \propto \exp\left(-\frac{1}{2}\beta m v_i^2\right). \tag{6.12}$$

A particle with initial velocity velocity v_i can only escape the well if v_i is above a threshold velocity, v_t, given by $\frac{1}{2}mv_t^2 = \Delta U$. Were there a barrier, this threshold velocity would be higher, but the remainder of this discussion would carry through unchanged. The final velocity of this particle, v_f, is determined by conservation of energy,

$$\frac{1}{2}mv_i^2 = \frac{1}{2}mv_f^2 + \Delta U. \tag{6.13}$$

This equation relates the distributions of initial and final velocities, $P(v_i)$ and $P(v_f)$ respectively. In particular,

$$P(v_f)\mathrm{d}v_f \propto P(v_i)\mathrm{d}v_i, \tag{6.14}$$

so

$$P(v_f) \propto P(v_i)\frac{\mathrm{d}v_i}{\mathrm{d}v_f} \propto \exp\left(-\frac{1}{2}\beta m v_i^2\right)\frac{v_f}{\sqrt{v_f^2 + \frac{2\Delta U}{m}}}. \tag{6.15}$$

For ΔU sufficiently large, the denominator in the last fraction is essentially constant over the range where the exponential factor is appreciable. Using this approximation, Equation (6.13), and the requirement that v_i exceed v_t, we obtain the approximation

$$P(v_f) \approx \begin{cases} \frac{m}{k_\mathrm{B}T} v_f \exp\left(-\frac{1}{2}\beta m v_f^2\right), & v_f > 0, \\ 0, & v_f \leq 0. \end{cases} \tag{6.16}$$

While Equation (6.16) applies only to an ideal gas of thermal particles escaping from a deep, barrierless potential well, it describes the observed distribution of v_z for evaporating molecules surprisingly well (dashed line in Figure 6.10). The small but systematic deviations are presumably due to the evaporating water molecule dissipating a small amount of energy to the remainder of the liquid as it escapes from the effective well set up by the bulk (see Section 6.2.2). It is worth noting that similar velocity distributions have been reported in simulations of evaporation of argon, where evaporation can be observed straightforwardly without special sampling techniques like TPS [110].

This distribution of evaporating water velocities has an apparent problem: it results in $\left\langle\frac{1}{2}mv_z^2\right\rangle = 1\,k_\mathrm{B}T$. Thus, whereas evaporating water molecules have a mean kinetic energy of $7/2\,k_\mathrm{B}T$ (not shown), the mean kinetic energy of water molecules in the vapor that recondense is only $6/2\,k_\mathrm{B}T$! Naively, it would seem that this situation is incompatible with the liquid and the vapor being at equilibrium. Of course, this contradiction is illusory. While it is true that the mean kinetic energy of an arbitrary water in the vapor is $6/2\,k_\mathrm{B}T$, the mean per-molecule *flow* of kinetic energy in the $-\hat{\mathbf{z}}$ direction through the $z = 0$ plane is actually $7/2\,k_\mathrm{B}T$. This is because for a particle to flow through this plane in a time Δt, it must be in the region

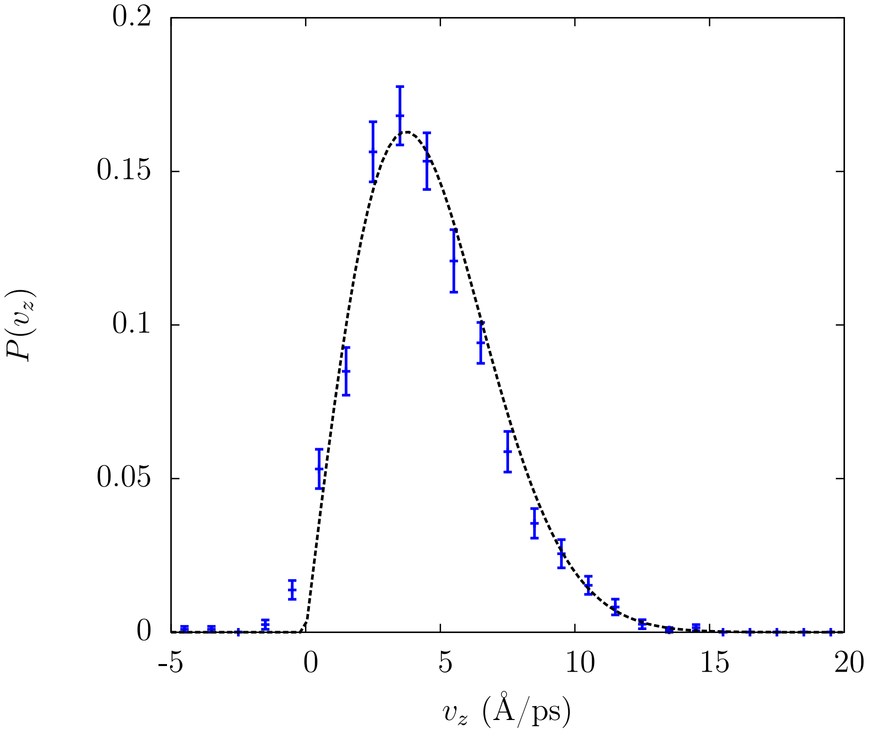


Figure 6.10: Observed distribution of z-component of velocity of evaporating water molecules (error bars). This distribution compares favorably to that of thermal ideal gas particles evaporating from a deep, barrierless potential well (Equation (6.16), dashed line).

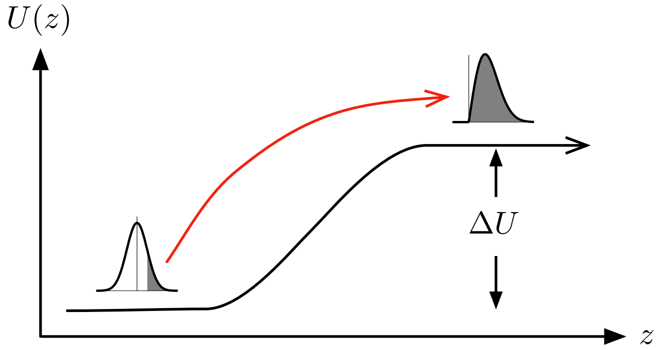


Figure 6.11: Ideal gas particles at the bottom of a deep, barrierless potential well have a Boltzmann distribution of velocities. Only a fraction of particles have enough energy to escape the well. After evaporating, but before thermalizing outside the well, the distribution of velocities of these particles is given by Equation (6.16).

defined by $z + v_z\Delta t < 0$. Hence, water molecules with higher v_z are overrepresented in the ensemble of water molecules flowing through the $z = 0$ plane per unit time. In fact, the distribution of observed velocities is given, up to a sign difference, by an expression identical to Equation (6.16). As a result, the rate at which energy and mass flow from the vapor to the liquid can exactly compensate the flows due to evaporation.

Non-equilibrium evaporation experiments have reported steady state temperature distributions where the temperature of the vapor immediately above the liquid is higher than that of the liquid [29]. Our discussion on velocity distributions may have some bearing on this observation. Moreover, it has been observed in molecular dynamics simulations of water condensation [110] that the condensation coefficient decreases with increasing temperature. The ideal gas evaporation model of Equation (6.15) is only compatible with a unit condensation coefficient if the well depth is much larger than $k_\mathrm{B}T$. As temperature is raised, the relative depth of the potential well set up by the liquid is reduced. In such a case, it may be possible to use Equation (6.15), together with the known distribution of velocities for condensing particles, to derive an upper bound on the condensation and evaporation coefficients of the liquid at high temperature. We have not pursued this possibility here.

6.2.2 Potential of mean force for removing a water molecule from bulk is barrierless

The height of a water molecule i above (or below) the liquid-vapor interface, a_i^*, is a natural order parameter for describing the evaporation of this molecule. Following Ref. [125], we define this height independently of the orientation of the slab as follows. We first construct the instantaneous liquid-vapor interface formed by the other water molecules, as described in Section 3.4. Next, we find the point $\mathbf{s}_i$ on the interface closest to water molecule i at $\mathbf{r}_i$, and denote by $\hat{\mathbf{n}}_i$ the vapor-pointing normal vector to the interface at $\mathbf{s}_i$. Then, a_i^* is defined as the distance from $\mathbf{r}_i$ to $\mathbf{s}_i$ projected along the $\hat{\mathbf{n}}_i$ direction,

$$a_i^* = \hat{\mathbf{n}}_i \cdot (\mathbf{r}_i - \mathbf{s}_i). \tag{6.17}$$

Figure 6.12 shows the free energy, $F(a^*)$, of an arbitrary water molecule in our system as a function of a^*, calculated using the umbrella potential described in Section 3.4 and MBAR (Section 3.1.1). The essential feature of this free energy is that it is barrierless. Apart from density layering in the bulk, manifest as oscillations in $F(a^*)$ for $a^* \lesssim 0\,\text{Å}$, the bulk liquid simply sets up a deep potential well for any individual water molecule, and a molecule in the vapor can simply roll downhill into this well. While the absence of a barrier along the a^* coordinate does not preclude the existence of barriers along other coordinates, we demonstrate below that the transition states of the evaporation trajectories are consistent with a^* describing the majority of the evaporation reaction coordinate.

The depth of the well in $F(a^*)$, denoted by ΔF^*, quantifies the cohesiveness of the liquid with respect to the vapor. Indeed, if we regard a single water molecule as an independent

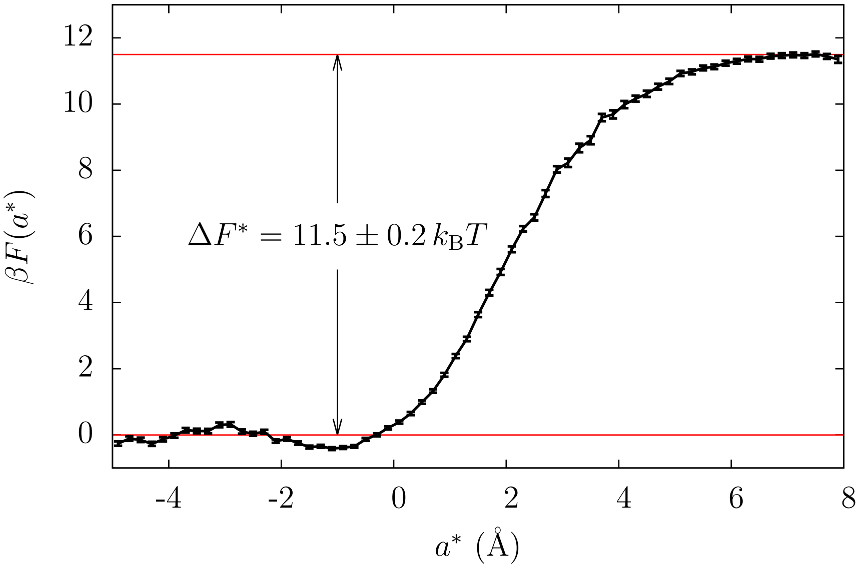


Figure 6.12: Free energy for a single water molecule at a height a^* from the liquid-vapor interface defined by the remaining water molecules. The red lines are the free energies of the stable liquid and vapor phases, and are guides to the eye.

particle moving in the potential well $F(a^*)$, then the relative density of this particle in the liquid, ρ_ℓ, with respect to that in the vapor, ρ_g, is given by

$$\rho_g = \rho_\ell e^{-\beta \Delta F^*}. \tag{6.18}$$

We estimate from Figure 6.12 a value of ΔF^* of $11.5 \pm 0.2\, k_{\mathrm{B}}T$. This compares favorably with the value of $11.8\, k_{\mathrm{B}}T$ obtained by setting $\rho_g = P_{\mathrm{vap}}/k_{\mathrm{B}}T$ and using the computed value of P_{vap} for SPC/E water at a temperature of 300 K and pressure of 1 atm [28]. For real water, the analogous calculation yields $\Delta F^* = 10.5\, k_{\mathrm{B}}T$.

The range of $F(a^*)$ also characterizes the effective range of attraction between a molecule in the vapor and the bulk slab. It is this range that motivates the definition of basin B described in Section 6.1.3.

6.2.3 Transition states are consistent with diffusion out of a deep well

We now focus on characterizing the transition states between condensed and evaporated states. Ordinarily [7], a transition states are identified using committor functions. The committor, $p_B(\mathbf{x})$, of a spatial configuration $\mathbf{x}$ is defined as the fraction of short trajectories that start at $\mathbf{x}$ with random thermal velocities, and finish in basin B. At most point in a transition path, this function is either 0 or 1, with a quick crossover around the configurations that dominate the dynamical bottleneck between A and B. Thus, a pragmatic definition of a transition state along a trajectory is the point where $p_B(\mathbf{x}) = 0.5$.

Implicit in the above definition is the assumption that momenta are not important in characterizing the transition state. In a dense system, this assumption is generally true, since the velocity of any particle decorrelates rapidly, usually within 1 ps [10]. When examining evaporation, the assumption breaks down, since the velocity of an evaporated water molecule decorrelates over much longer timescales. The clearest manifestation of the problem is that $p_B < 1$ for a configuration containing a single, clearly evaporated water molecule, since the water can will recondense if its velocity points towards the liquid slab.

As a compromise, we have chosen to redefine the committor function to include the z-component of the velocity of the evaporated water molecule. Strictly speaking, it's impossible to tell which water molecule is "the evaporated molecule" in an *arbitrary* configuration, but this is not a problem for identifying transition states along a transition path. Figure 6.13 illustrates the typical behavior of $p_B(t) = p_B(\mathbf{x}(t), v_z^{\mathrm{evap}}(t))$ defined in this way, estimated by spawning 10 short trajectories at every time point.

We have identified the transition states in a moderately-sized subset of the evaporation trajectories by finding the time along each trajectory where $p_B(t) \approx 0.5$. Since calculating $p_B(t)$ in full for each trajectory is expensive, we have instead implemented a bisection algorithm in t to locate the transition state with reasonable accuracy. At each value of t, we estimate t with 10 short trajectories. If $p_B(t) < 0.3$ or $p_B(t) > 0.7$, we eliminate the

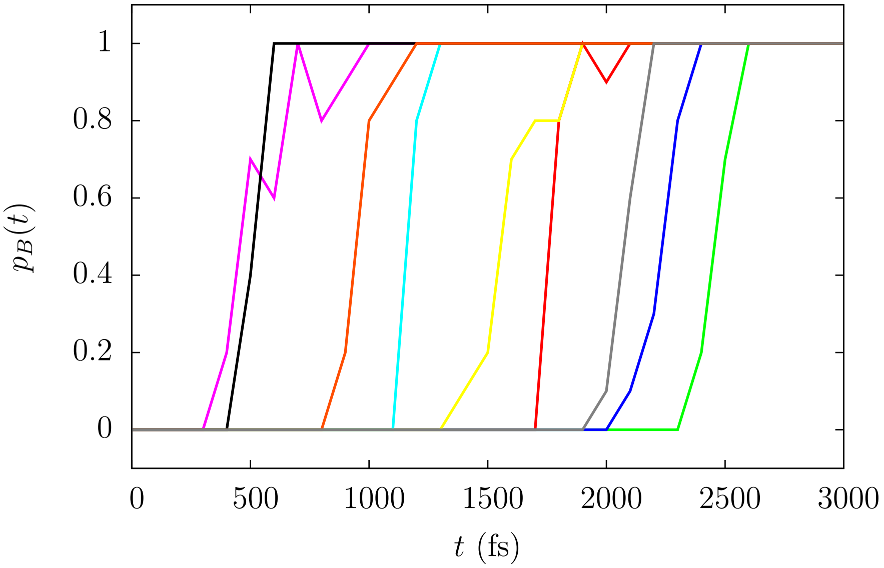


Figure 6.13: Estimated committor, $p_B(t)$, sampled at 100 fs intervals along a small random selection of evaporation trajectories. To obtain this estimate, all velocities but the z-component of the evaporated water's velocity are randomized independently 10 times, after which a short trajectory is evolved forwards in time for up to 5 ps until the systems enters either basin A or basin B.

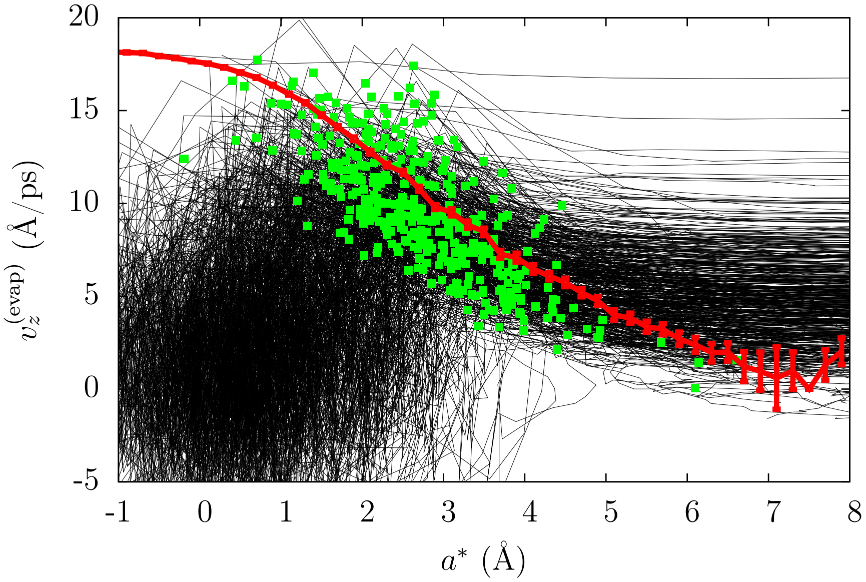


Figure 6.14: Evaporation trajectory traces projected onto variables $v_z^{(\mathrm{evap})}$ and a^* (black lines). The transition state of each trajectory, identified as described in the text, is highlighted by a green dot. Red line: the expected transition state ensemble for a coarse model of ballistic escape from a potential shaped as in Figure 6.12, given by Equation 6.19.

configurations at all times before or after t, respectively, as candidates for the transition state. Otherwise, if $0.3 \leq p_B(t) \leq 0.7$, we abort the bisection search, and calculate estimates of $p_B(t)$ at times $\pm 40\,\mathrm{fs}$ and $\pm 80\,\mathrm{fs}$ from where the bisection algorithm was stopped. Finally, we fit all the $p_B(t)$ estimates to the functional form $p_B(t) = (1 + \tanh[(t - t_c)/\tau])/2$, with t_c and τ being the fit parameters. The resulting estimate for the transition state is the configuration at time closest to t_c.

Figure 6.14 depicts traces of many evaporation trajectories projected onto the two coordinates $v_z^{(\mathrm{evap})}$ and a^*, with the transition state of each trajectory highlighted in green. Unlike similar traces onto many other pairs of coordinates (not shown), there is a definite correlation between the distance of the evaporated water from the liquid-vapor interface and its speed in the z direction. We can partially rationalize this dependence by conceiving of the free energy along a^* (Figure 6.12) as an actual potential energy well, and approximating the velocity along the a^* direction with $v_z^{(\mathrm{evap})}$. If evaporation were a ballistic escape from this well, then the transition states would satisfy the condition

$$\frac{1}{2} m (v_z^{(\mathrm{evap})})^2 = F(a^*). \tag{6.19}$$

The points satisfying this relation are shown as a thick red line in Figure 6.14. The transition states clearly cluster around this line.

Together with the distribution of the z-component of the velocities of evaporated water molecules (Section 6.2.1), the fact that the transition states are well-captured by the coarse model of an escape from a deep, barrierless potential well suggests that there is practically no additional barrier to evaporation for SPC/E water at 300 K. This is consistent with the near-unit condensation coefficient measured in simulations in Ref. [110]. This finding is in apparent contradiction with the experimental results, which suggested a barrier of around $-k_\mathrm{B}T\ln(\gamma) \approx 0.5\, k_\mathrm{B}T$. We discuss this point further in Section 6.3.1.

6.2.4 Evaporation correlates with negative mean curvature

Although it appears that evaporation is barrierless in simulations, we can nonetheless characterize some of the microscopic details of this process. Our motivation was initially to characterize potential barriers to evaporation, which seem to be absent under the conditions studied here. However, we anticipate that external influences might be used to alter the microscopic details we describe, and so may perhaps be used to exert control over evaporation. Additionally, our characterization establishes a baseline for understanding evaporation under different conditions where barriers *are* observed in simulations, such as at higher temperatures [110] or in the presence of surfactants [104]. Here, we focus on how curvature of the liquid-vapor interfacial couples to evaporation.

We approximate the liquid-vapor interface using the procedure of Willard and Chandler [125], and focus on the point $\mathbf{s}$ on the surface that is closest to the evaporating water molecule. The mean curvature, K_M, of the surface at $\mathbf{s}$ serves as a concise characterization of collective fluctuations of water molecules at the liquid-vapor surface. The mean curvature is defined as [61]

$$K_M = \frac{k_1 + k_2}{2}, \tag{6.20}$$

where k_1 and k_2 are the principal curvatures at $\mathbf{s}$. The magnitude of a principal curvature is the reciprocals of the principal radius of curvature, and its sign specifies whether the surface curves towards (positive) or away (negative) from the normal direction along the corresponding principal direction. Our interest in mean curvature stems from its relation to infinitesimal surface area changes. An deformation along the normal direction by an infinitesimal distance ϵ changes the area element $\mathrm{d}A$ as follows[5]

$$\mathrm{d}A \mapsto (1 - 2\epsilon K_M)\mathrm{d}A. \tag{6.21}$$

---

[5] One way to understand this result is as follows. Consider an infinitesimal patch $\mathrm{d}A$ of a surface that curves away from the normal in every direction. Its radii of curvature are R_1 and R_2, so the principal curvatures are $-1/R_1$ and $-1/R_2$. Before the deformation, the boundaries of the area element are arcs of length $R_1\mathrm{d}\theta_1$ and $R_2\mathrm{d}\theta_2$, so that $\mathrm{d}A = R_1R_2\,\mathrm{d}\theta_1\mathrm{d}\theta_2$. After the deformation, the radii of curvature are $R_1+\epsilon$ and $R_2+\epsilon$, so that the area element is $\mathrm{d}A(1+\epsilon/R_1+\epsilon/R_2) = \mathrm{d}A(1-2\epsilon K_M)$.

This relation justifies regarding K_M as a local characterization of the force of surface tension on the liquid-vapor surface.

A technical complication arises when calculating the curvature of the liquid-vapor interface defined by Willard and Chandler. Owing to the truncation in the Gaussian smearing kernel used to define the smooth density field $\tilde{\rho}(\mathbf{r})$ (Equation (3.63)), the surface has discontinuous gradients. Since curvature captures second-derivative information about a surface, the curvature at these points is infinite. To avoid these complications, here we define $\tilde{\rho}(\mathbf{r})$ using a smoothing function $\phi(\mathbf{r}) = \varphi(r)$ that is slightly different than the one used in Ref. [125] and in the preceding chapters. Our chosen $\varphi(r)$ stitches together two continuous cubic functions at a point $r = 0.7\xi$ so that $\varphi(r)$ is continuous and has continuous first and second derivatives. It also satisfies $\varphi(3\xi) = 0$, $\varphi'(0) = \varphi'(3\xi) = 0$ and $\varphi''(0) = \varphi''(3\xi) = 0$, and the resulting smearing function, $\phi(\mathbf{r})$, is normalized to 1. These properties uniquely specify $\varphi(r)$. The location of the stitching point is chosen empirically so that $\phi(r)$ closely resembles a Gaussian with standard deviation ξ.

To establish a baseline, we first calculate the distribution of K_M as a function of the height a^* of a probe water molecule from the liquid-vapor interface. In doing this, the probe water molecule can be either included in or exclude from the calculation of the liquid-vapor surface. It is not clear that either choice is superior, so we have repeated the calculation with both choices. Figure 6.15 shows the results as a joint free energy for K_M and a^*. At very low and very high a^*, only a trivial bias in K_M is seen as a function of a^*, resulting from the nearest point on the surface being preferentially one where the surface is bending most towards the probe water molecule. However, an evident additional bias towards negative mean curvature can be seen for a^* just above the surface, indicating that a water molecule suspended there significantly deforms the surface below it. Figure 6.16 shows an example of this kind of deformation in one of the harvested evaporation trajectories.

Figure 6.17 overlays the transition states of the evaporation trajectories on the free energies of Figure 6.15. To a certain extent, the transition states exhibit some of the bias towards negative curvature that can be seen in the equilibrium free energies. The bias is slight when the probe molecule is not included in the definition of the liquid-vapor surface, but is clearer when the probe molecule is included. The definition of a liquid-vapor interface during the evaporation process is somewhat ambiguous, and we regard full inclusion and full exclusion as the two limiting extremes for a suitable definition. Since the bias towards negative curvature is present in both cases, our finding should be robust with respect to reasonable changes in the definition of the interface.

6.3 Discussion

We have examined the process of evaporation of SPC/E water in detail, and all the evidence suggests that there is no barrier to evaporation in this model. In other words, to evaporate, a water molecule near the surface need simply spontaneously acquire enough ki-

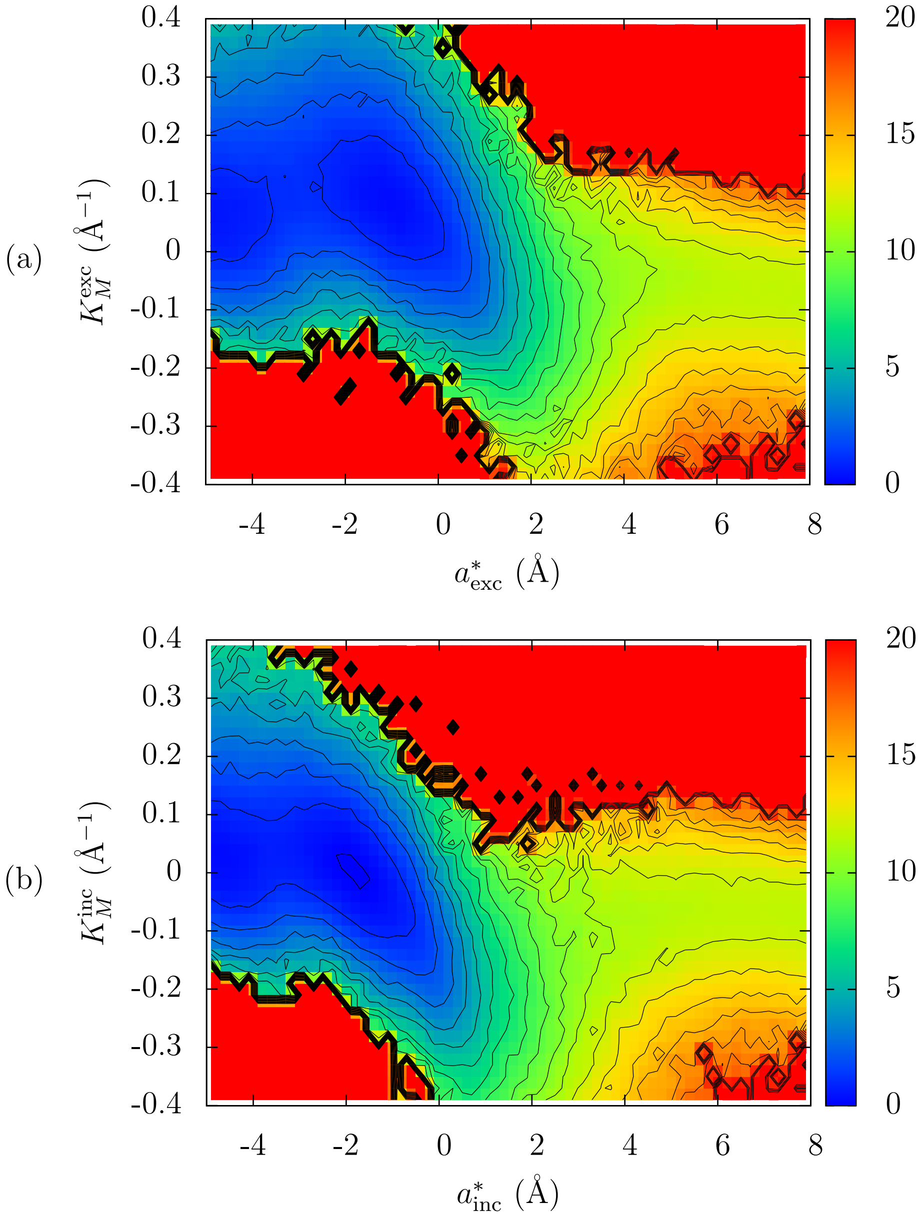
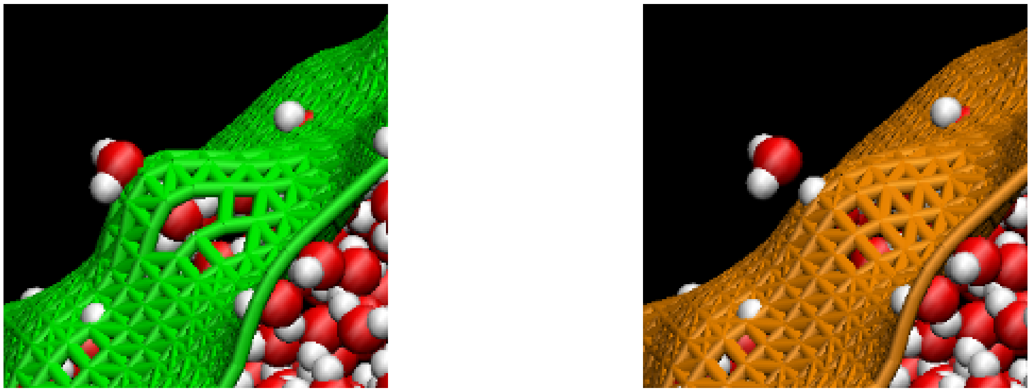


Figure 6.15: Free energy for height a^* of a probe water molecule and the mean curvature, K_M, at the nearest point on the liquid-vapor surface. Contours are spaced at $1\,k_{\mathrm{B}}T$. The probe water molecule can either be (a) excluded from, or (b) included in the definition of the liquid-vapor surface.



Evaporated water included

Evaporated water excluded

Figure 6.16: Example of the interfacial deformation that accompanies an evaporation event. The system is depicted at the time when the last hydrogen bond between the bulk and the evaporated water breaks.

netic energy in the direction of the liquid-vapor interface normal. This view is consistent with the distribution of v_z for the final velocities (Figure 6.10), the fact that the potential of mean force along a coordinate a^* perpendicular to the liquid-vapor surface is barrierless (Figure 6.12) and the fact that the transition states cluster around values of v_z and a^* that have a threshold amount of energy to escape from the potential well set up by the remainder of the bulk (Figure 6.14). It is difficult, if not impossible, for evaporation to be a mildly activated process and still be consistent with these three pieces of evidence.

One possible explanation for the discrepancy is that the experimentally observed barrier is a fundamentally quantum effect. By construction, these effects are beyond the scope of the classical molecular dynamics simulations used here. Important quantum effects are plausible because librational motions of water, with wavenumbers around $500\,\mathrm{cm}^{-1}$, have strong quantization effects (a photon with energy $k_\mathrm{B}T$ at $T = 300\,\mathrm{K}$ has a wavenumber of $1300\,\mathrm{cm}^{-1}$). If this were the source of the discrepancy, a simulation study that explicitly accounts for quantum effect (e.g., using ring-polymer molecular dynamics [17]) could elucidate the details of the barrier. However, any account of such quantum effects playing a dominant role would have to be compatible with the observation that the evaporation coefficient of D_2O is equal, within errors, to that of H_2O [22].

An alternate explanation that we pursue here is that of a systematic error in the calibration of the experiments.

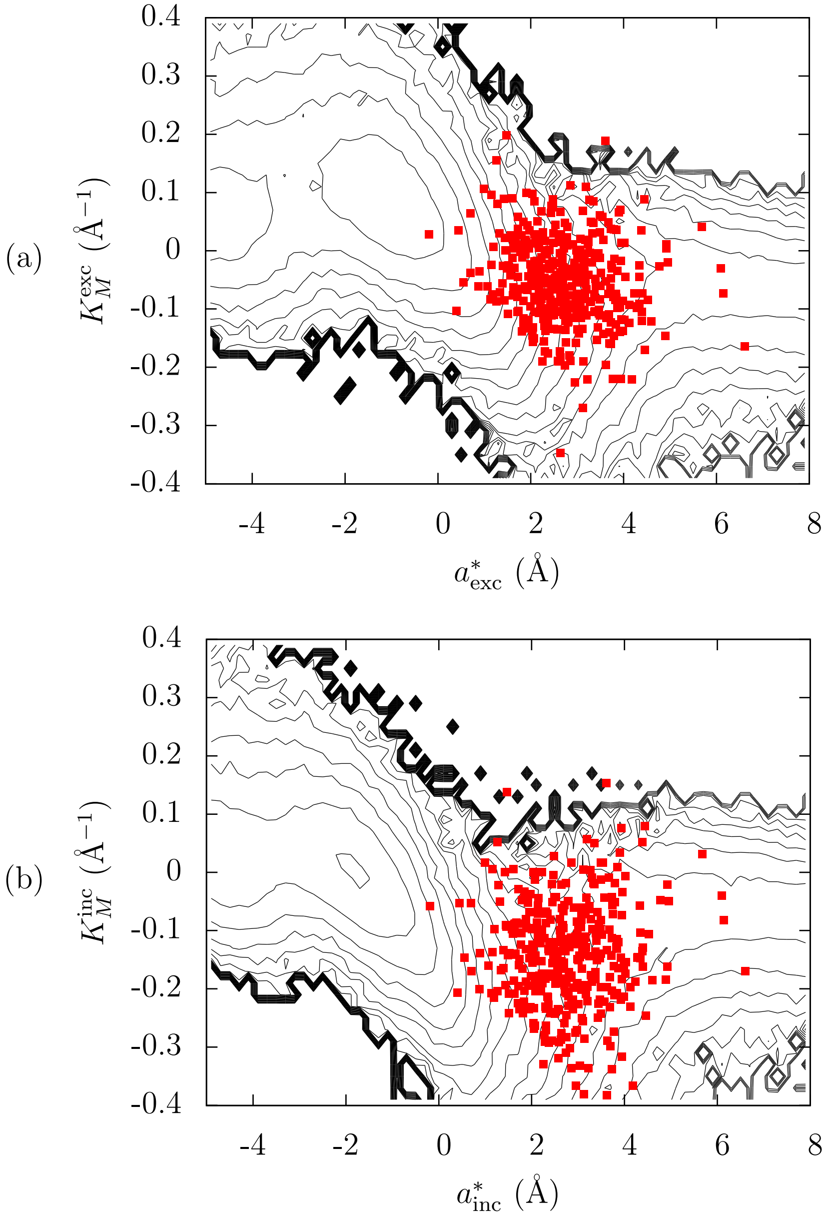


Figure 6.17: Representative transition states of evaporation trajectories (red) projected onto the K_M and a^* coordinates. The free energies of these coordinates are shown for comparison. Labels as in Figure 6.15.

6.3.1 Experimental temperature calibration may explain discrepancies between simulations and experiments

As described in Section 6.1.1, in the experiment that motivated this study, the evaporation coefficient is deduced from cooling curves of microdroplets. The droplet temperature is estimated by measuring its Raman scattering spectrum and assuming a linear van't Hoff relation between ratios of integrated scattering intensities and inverse temperature (Figure 6.3). However, this linear behavior is an approximation that only holds over small temperature ranges [36]. Further, whereas the temperature is calibrated using the Raman spectra of water samples at temperatures above 0° C, most of the observed temperatures are below 0° C.

We first determine whether significant deviations from linear van't Hoff behavior can be observed in water. Very few experiments have measured the Raman scattering spectrum of supercooled water, with the notable exception of the work of D'Arrigo et al [19]. They report a Raman spectrum of water at -24° C, as well as five other temperatures ranging from 20° C to 95° C. Using their data, we have produced a van't Hoff calibration curve analogous to that of Figure 6.3. Our result, shown in Figure 6.18, clearly shows that while the five points data point taken above 0° can be fit to a straight line, the data point at deeply supercooled temperatures deviates from this line significantly. The conditions under which the spectra of D'Arrigo et al and those of Smith et al [99] are taken are not directly comparable, and troublingly, to the author's knowledge, the D'Arrigo et al data has not been reproduced, but we interpret the deviation from linearity observed in Figure 6.18 as evidence of the plausibility of calibration errors affecting the results in Ref. [99].

To bolster this argument, we have reinterpreted the cooling curves reported in Ref. [99] as reporters of calibration non-linearities. That is, if we assume that the temperature T of the droplet is in fact the expected temperature when $\gamma = 1$, then the "measured" temperature can be viewed as a calibration measurement of the ratio of integrated Raman scattering intensities at temperature T. The resulting calibration curve is shown in Figure 6.19. This figure illustrates that the amount of deviation from linearity required for the data to be compatible with $\gamma = 1$ is small, and within the range suggested by the above results.

One way to compensate for non-linearity empirically is to use a quadratic fit, instead of a linear fit, to describe the calibration data: the calibration data of Smith et al already has a slight but measurable curvature. The result of this fit is also shown in Figure 6.19. The deviation from linearity observed for $T < 0^\circ$ C is indeed roughly consistent with the proposed nonlinearity (Figure 6.19)[6].

Ultimately, the best way to address the issues with the current calibration procedure is to either inject warmer microdroplets into the experimental observation chamber to begin with, or to measure calibration data for water at supercooled conditions, or both. Our experimental collaborators are presently pursuing both of these avenues.

---

[6] Recently, Kaitlin Duffey in Richard Saykally's group has implemented this quadratic fit on a more accurate calibration dataset, and reports that the deduced evaporation coefficient is closer to 1, but that a result of exactly 1 would still be incompatible with the measured cooling curves.

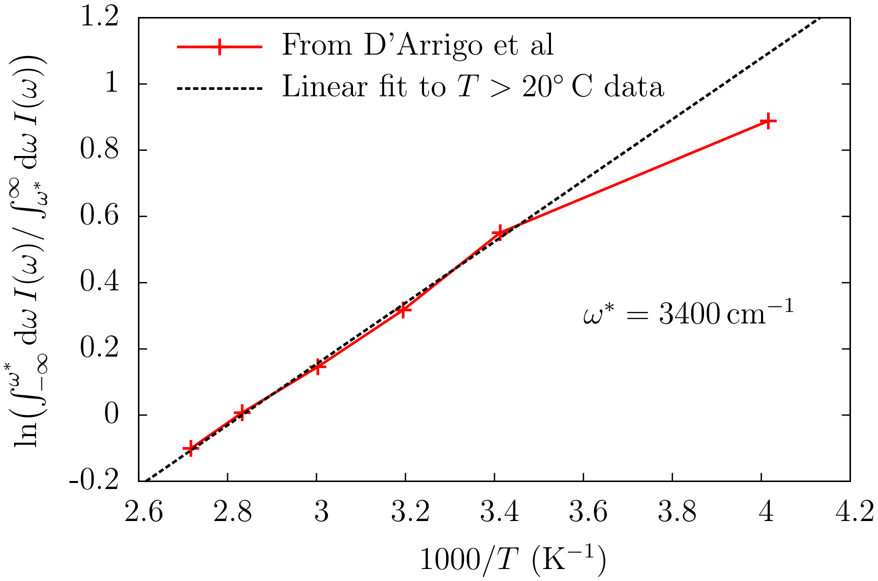


Figure 6.18: Van't Hoff calibration curve, analogous to Figure 6.3, derived from the data of D'Arrigo et al [19].

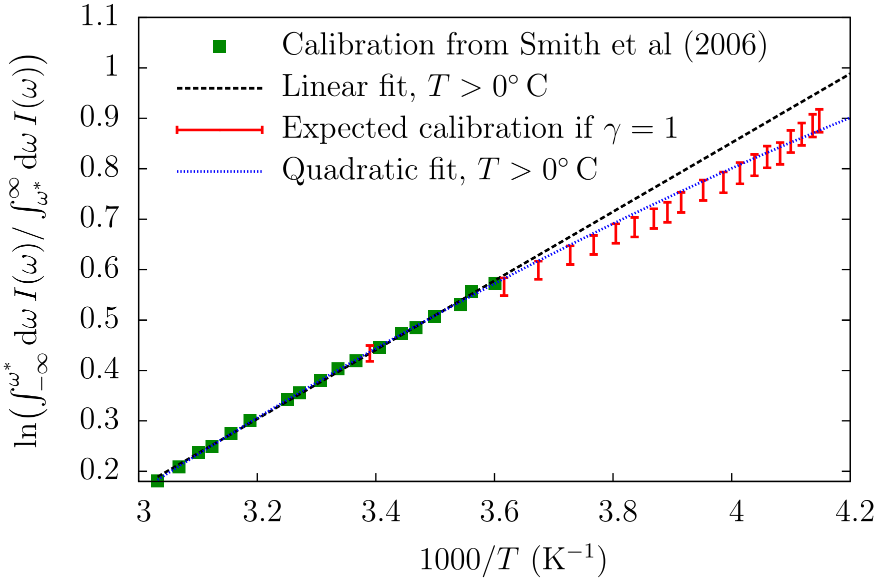


Figure 6.19: Expected van't Hoff calibration if $\gamma = 1$, derived from reported temperature measurement in Ref. [99].

Bibliography

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell.* Garland Science, 5th edition, 2007.

[2] C. Bartels. Analyzing biased Monte Carlo and molecular dynamics simulations. *Chem. Phys. Lett.*, 331(5–6):446–454, 2000.

[3] D. Bashford and D. Case. Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.*, 51:129–152, 2000.

[4] C. H. Bennett. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.*, 22(2):245–268, 1976.

[5] H. Berendsen, J. Grigera, and T. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91(24):6269–6271, 1987.

[6] P. G. Bolhuis and D. Chandler. Transition path sampling of cavitation between molecular scale solvophobic surfaces. *J. Chem. Phys.*, 113(18):8154–8160, 2000.

[7] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler. Transition path sampling: throwing ropes over rough mountain passes, in the dark. *Annu. Rev. Phys. Chem.*, 53:291–318, 2002.

[8] F. Buff, R. Lovett, and F. H. Stillinger. Interfacial density profile for fluids in the critical region. *Phys. Rev. Lett.*, 15(15):621–623, 1965.

[9] R. Car. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55(22):2471–2474, 1985.

[10] D. Chandler. *Introduction to Modern Statistical Mechanics.* Oxford University Press, 1987.

[11] D. Chandler. Gaussian field model of fluids with an application to polymeric fluids. *Phys. Rev. E*, 48(4):2898–2905, 1993.

[12] D. Chandler. Interfaces and the driving force of hydrophobic assembly. *Nature*, 437(7059):640–7, 2005.

[13] D. Chandler and P. Varilly. Lectures on molecular- and nano-scale fluctuations in water. In *Course CLXXVI—"Complex materials in physics and biology"*, Varenna, Italy, 2011. International School of Physics "Enrico Fermi". (to appear, pre-print at `http://arxiv.org/abs/1101.2235`).

[14] J. Chen and C. L. Brooks III. Implicit modeling of nonpolar solvation for simulating protein folding and conformational transitions. *Phys. Chem. Chem. Phys.*, 10(4):471–481, 2008.

[15] H. Conroy. Molecular Schrödinger equation. VIII. A new method for the evaluation of multidimensional integrals. *J. Chem. Phys.*, 47(12):5307–5318, 1967.

[16] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117(19):5179–5197, 1995.

[17] I. R. Craig and D. E. Manolopoulos. Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics. *J. Chem. Phys.*, 121(8):3368–3373, 2004.

[18] G. E. Crooks and D. Chandler. Gaussian statistics of the hard-sphere fluid. *Phys. Rev. E*, 56(4):4217–4221, 1997.

[19] G. D'Arrigo, G. Maisano, F. Mallamace, P. Migliardo, and F. Wanderlingh. Raman scattering and structure of normal and supercooled water. *J. Chem. Phys.*, 75(9):4264, 1981.

[20] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars. *Computational Geometry: Algorithms and Applications.* Springer, 3rd edition, 2008.

[21] C. E. Dempsey. The actions of melittin on membranes. *Biochim. Biophys. Acta*, 1031(2):143–161, 1990.

[22] W. S. Drisdell, C. D. Cappa, J. D. Smith, R. J. Saykally, and R. C. Cohen. Determination of the evaporation coefficient of D_2O. *Atmos. Chem. Phys.*, 8:6699–6706, 2008.

[23] W. S. Drisdell, R. J. Saykally, and R. C. Cohen. On the evaporation of ammonium sulfate solution. *P. Natl. Acad. Sci. U.S.A.*, 106(45):18897–18901, 2009.

[24] W. S. Drisdell, R. J. Saykally, and R. C. Cohen. Effect of surface active ions on the rate of water evaporation. *J. Phys. Chem. C*, 114(27):11880–11885, 2010.

[25] I. Eames, N. Marr, and H. Sabir. The evaporation coefficient of water: A review. *Int. J. Heat Mass Tran.*, 40(12):2963–2973, 1997.

[26] D. Earl and M. Deem. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.*, 7(23):3910–3916, 2005.

[27] J. L. England, D. Lucent, and V. Pande. Rattling the cage: computational models of chaperonin-mediated protein folding. *Curr. Opin. Struc. Biol.*, 18(2):163–169, 2008.

[28] J. Errington and A. Panagiotopoulos. A fixed point charge model for water optimized to the vapor-liquid coexistence properties. *J. Phys. Chem. B*, 102(38):7470–7475, 1998.

[29] G. Fang and C. Ward. Examination of the statistical rate theory expression for liquid evaporation rates. *Phys. Rev. E*, 59(1):441–453, 1999.

[30] B. U. Felderhof. Fluctuations of polarization and magnetization in dielectric and magnetic media. *J. Chem. Phys.*, 67(2):493–500, 1977.

[31] W. A. Fenton and A. L. Horwich. Chaperonin-mediated protein folding: fate of substrate polypeptide. *Q. Rev. Biophys.*, 36(2):229–256, 2003.

[32] A. M. Ferrenberg and R. H. Swendsen. Optimized Monte Carlo data analysis. *Phys. Rev. Lett.*, 63(12):1195–1198, 1989.

[33] D. Frenkel. Speed-up of Monte Carlo simulations by sampling of rejected states. *P. Natl. Acad. Sci. U.S.A.*, 101(51):17571–17575, 2004.

[34] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, 2nd edition, 2001.

[35] M. Frigo and S. G. Johnson. The design and implementation of FFTW3. *P. IEEE*, 93(2):216–231, 2005.

[36] P. L. Geissler. Temperature dependence of inhomogeneous broadening: On the meaning of isosbestic points. *J. Am. Chem. Soc.*, 127(42):14930–14935, 2005.

[37] P. L. Geissler and D. Chandler. Importance sampling and theory of nonequilibrium solvation dynamics in water. *J. Chem. Phys.*, 113(21):9759–9765, 2000.

[38] P. L. Geissler, C. Dellago, and D. Chandler. Chemical dynamics of the protonated water trimer analyzed by transition path sampling. *Phys. Chem. Chem. Phys.*, 1(6):1317–1322, 1999.

[39] R. Godawat, S. N. Jamadagni, and S. Garde. Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *P. Natl. Acad. Sci. U.S.A.*, 106(36):15119–15124, 2009.

[40] D. J. Griffiths. *Introduction to Electrodynamics.* Benjamin Cummings, 3rd edition, 1999.

[41] J.-P. Hansen and I. R. McDonald. *Theory of Simple Liquids.* Academic Press, 3rd edition, 2006.

[42] C. B. Haselgrove. A method for numerical integration. *Math. Comput.*, 15(76):323–337, 1961.

[43] G. Hawkins, C. Cramer, and D. G. Truhlar. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem.*, 100(51):19824–19839, 1996.

[44] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4(3):435–447, 2008.

[45] R. W. Hockney and J. W. Eastwood. *Computer Simulation Using Particles.* Taylor & Francis, 1989.

[46] D. M. Huang and D. Chandler. Cavity formation and the drying transition in the Lennard-Jones fluid. *Phys. Rev. E*, 61(2):1501–1506, 2000.

[47] D. M. Huang and D. Chandler. Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding. *P. Natl. Acad. Sci. U.S.A.*, 97(15):8324–8327, 2000.

[48] D. M. Huang and D. Chandler. The hydrophobic effect and the influence of solute-solvent attractions. *J. Phys. Chem. B*, 106(8):2047–2053, 2002.

[49] D. M. Huang, P. L. Geissler, and D. Chandler. Scaling of hydrophobic solvation free energies. *J. Phys. Chem. B*, 105(28):6704–6709, 2001.

[50] H. W. Huang. Molecular mechanism of antimicrobial peptides: The origin of cooperativity. *Biochim. Biophys. Acta*, 1758(9):1292–1302, 2006.

[51] X. Huang, C. J. Margulis, and B. J. Berne. Dewetting-induced collapse of hydrophobic particles. *P. Natl. Acad. Sci. U.S.A.*, 100(21):11953–11958, 2003.

[52] X. Huang, R. Zhou, and B. J. Berne. Drying and hydrophobic collapse of paraffin plates. *J. Phys. Chem. B*, 109(8):3546–3552, 2005.

[53] G. Hummer, S. Garde, A. Garcia, A. Pohorille, and L. Pratt. An information theory model of hydrophobic interactions. *P. Natl. Acad. Sci. U.S.A.*, 93(17):8951–8955, 1996.

[54] G. Hummer, N. Gronbech-Jensen, and M. Neumann. Pressure calculation in polar and charged systems using Ewald summation: Results for the extended simple point charge model of water. *J. Chem. Phys.*, 109(7):2791–2797, 1998.

[55] W. Humphrey, A. Dalke, and K. Schulten. VMD: Visual molecular dynamics. *J. Mol. Graphics*, 14(1):33–38, 1996.

[56] P. Hünenberger. Calculation of the group-based pressure in molecular simulations. I. A general formulation including Ewald and particle-particle-particle-mesh electrostatics. *J. Chem. Phys.*, 116(16):6880–6897, 2002.

[57] A. I. Jewett and J.-E. Shea. Reconciling theories of chaperonin accelerated folding with experimental evidence. *Cell. Mol. Life. Sci.*, 67(2):255–276, 2010.

[58] W. L. Jorgensen, J. D. Madura, and C. J. Swenson. Optimized intermolecular potential functions for liquid hydrocarbons. *J. Am. Chem. Soc.*, 106(22):6638–6646, 1984.

[59] P. Jungwirth and D. Tobias. Ions at the air/water interface. *J. Phys. Chem. B*, 106(25):6361–6373, 2002.

[60] A. Kong, P. McCullagh, X. Meng, D. Nicolae, and Z. Tan. A theory of statistical models for Monte Carlo integration. *J. Roy. Stat. Soc. B*, 65:585–604, 2003.

[61] E. Kreyszig. *Differential Geometry.* Dover, 1991.

[62] J. Kyte and R. F. Doolittle. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, 157(1):105–132, 1982.

[63] E. Lemmon, M. McLinden, and D. Friend. Thermophysical properties of fluid systems. In P. Linstrom and W. Mallard, editors, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69.* National Institute of Standards and Technology, Gaithersburg MD, 20899, 2011. `http://webbook.nist.gov` (retrieved May 16, 2011).

[64] P. Liu, X. Huang, R. Zhou, and B. J. Berne. Observation of a dewetting transition in the collapse of the melittin tetramer. *Nature*, 437(7055):159–162, 2005.

[65] W. Lorensen and H. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Comp. Graph.*, 1987.

[66] M. Louwerse and E. Baerends. Calculation of pressure in case of periodic boundary conditions. *Chem. Phys. Lett.*, 421(1–3):138–141, 2006.

[67] K. Lum. *Hydrophobicity at Small and Large Length Scales*. PhD thesis, University of California, Berkeley, 1998.

[68] K. Lum, D. Chandler, and J. D. Weeks. Hydrophobicity at small and large length scales. *J. Phys. Chem. B*, 103(22):4570–4577, 1999.

[69] A. Luzar and D. Chandler. Hydrogen-bond kinetics in liquid water. *Nature*, 379(6560):55–57, 1996.

[70] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.

[71] L. Maibaum, A. R. Dinner, and D. Chandler. Micelle formation and the hydrophobic effect. *J. Phys. Chem. B*, 108(21):6778–6781, 2004.

[72] M. Marchi, D. Borgis, N. Levy, and P. Ballone. A dielectric continuum molecular dynamics method. *J. Chem. Phys.*, 114(10):4377–4385, 2001.

[73] R. Marcus. On the theory of oxidation-reduction reactions involving electron transfer. I. *J. Chem. Phys.*, 24(5):966–978, 1956.

[74] G. Martyna, D. Tobias, and M. Klein. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.*, 101:4177, 1994.

[75] P. E. Mason. Molecular dynamics study on the microscopic details of the evaporation of water. *J. Phys. Chem. A*, 115(23):6054–6058, 2011.

[76] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

[77] M. Mickler, M. Hessling, C. Ratzke, J. Buchner, and T. Hugel. The large conformational changes of Hsp90 are only weakly coupled to ATP hydrolysis. *Nat. Struct. Mol. Biol.*, 16(3):281–286, 2009.

[78] T. F. Miller, E. Vanden-Eijnden, and D. Chandler. Solvent coarse-graining and the string method applied to the hydrophobic collapse of a hydrated chain. *P. Natl. Acad. Sci. U.S.A.*, 104(37):14559–64, 2007.

[79] J. Mittal and G. Hummer. Static and dynamic correlations in water at hydrophobic interfaces. *P. Natl. Acad. Sci. U.S.A.*, 105(51):20130–20135, 2008.

[80] S. Miyamoto and P. A. Kollman. SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.*, 13(8):952–962, 1992.

[81] A. H. Narten and H. A. Levy. Liquid water: Molecular correlation functions from X-ray diffraction. *J. Chem. Phys.*, 55(5):2263–2269, 1971.

[82] J. Noah-Vanhoucke and P. L. Geissler. On the fluctuations that drive small ions toward, and away from, interfaces between polar liquids and their vapors. *P. Natl. Acad. Sci. U.S.A.*, 106(36):15125–15130, 2009.

[83] A. Onufriev, D. Bashford, and D. A. Case. Exploring protein native states and large-scale conformational changes with a modified Generalized Born model. *Proteins*, 55(2):383–394, 2004.

[84] A. J. Patel, P. Varilly, and D. Chandler. Fluctuations of water near extended hydrophobic and hydrophilic surfaces. *J. Phys. Chem. B*, 114(4):1632–1637, 2010.

[85] A. J. Patel, P. Varilly, D. Chandler, and S. Garde. Quantifying density fluctuations in volumes of all shapes and sizes using indirect umbrella sampling. (accepted for publication in *J. Stat. Phys.*, arXiv:1105.0895).

[86] A. J. Patel, P. Varilly, S. N. Jamadagni, H. Acharya, S. Garde, and D. Chandler. Hydrophobic effects in interfacial environments. (submitted, arXiv:1105.0907).

[87] P. Petersen and R. J. Saykally. On the nature of ions at the liquid water surface. *Annu. Rev. Phys. Chem.*, 57:333–364, 2006.

[88] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten. Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, 26(16):1781–1802, 2005.

[89] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.*, 117(1):1–19, 1995. (`http://lammps.sandia.gov`).

[90] A. Pohorille, C. Jarzynski, and C. Chipot. Good practices in free-energy calculations. *J. Phys. Chem. B*, 114(32):10235–10253, 2010.

[91] L. Pratt. Molecular theory of hydrophobic effects: "She is too mean to have her name repeated". *Annu. Rev. Phys. Chem.*, 53:409–436, 2002.

[92] B. Roux and T. Simonson. Implicit solvent models. *Biophys. Chem.*, 78(1–2):1–20, 1999.

[93] J. S. Rowlinson and B. Widom. *Molecular Theory of Capillarity.* Dover, 1982.

[94] F. Sedlmeier, D. Horinek, and R. R. Netz. Nanoroughness, intrinsic density profile, and rigidity of the air-water interface. *Phys. Rev. Lett.*, 103(13):136102, 2009.

[95] P. Setny and M. Zacharias. Hydration in discrete water. a mean field, cellular automata based approach to calculating hydration free energies. *J. Phys. Chem. B*, 114(26):8667–8675, 2010.

[96] M. R. Shirts, E. Bair, G. Hooker, and V. Pande. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Lett.*, 91(14):140601, 2003.

[97] M. R. Shirts and J. D. Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.*, 129(12):124105, 2008.

[98] J. D. Smith, C. Cappa, K. Wilson, R. C. Cohen, P. L. Geissler, and R. J. Saykally. Unified description of temperature-dependent hydrogen-bond rearrangements in liquid water. *P. Natl. Acad. Sci. U.S.A.*, 102(40):14171–14174, 2005.

[99] J. D. Smith, C. D. Cappa, W. S. Drisdell, R. C. Cohen, and R. J. Saykally. Raman thermometry measurements of free evaporation from liquid water droplets. *J. Am. Chem. Soc.*, 128(39):12892–12898, 2006.

[100] X. Song and D. Chandler. Dielectric solvation dynamics of molecules of arbitrary shape and charge distribution. *J. Chem. Phys.*, 108(6):2594–2600, 1998.

[101] X. Song, D. Chandler, and R. Marcus. Gaussian field model of dielectric solvation dynamics. *J. Phys. Chem.*, 100(29):11954–11959, 1996.

[102] W. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, 112(16):6127–6129, 1990.

[103] C. G. Sztrum-Vartash and E. Rabani. Lattice gas model for the drying-mediated self-assembly of nanorods. *J. Phys. Chem. C*, 114(25):11040–11049, 2010.

[104] S. Takahama and L. M. Russell. A molecular dynamics study of water mass accommodation on condensed phase water coated by fatty acid monolayers. *J. Geophys. Res.*, 116(D2):D02203, 2011.

[105] C. Tanford. *The hydrophobic effect: formation of micelles and biological membranes.* Wiley, 1973.

[106] P. R. ten Wolde and D. Chandler. Drying-induced hydrophobic polymer collapse. *P. Natl. Acad. Sci. U.S.A.*, 99(10):6539–6543, 2002.

[107] P. R. ten Wolde, S. Sun, and D. Chandler. Model of a fluid at small and large length scales and the hydrophobic effect. *Phys. Rev. E*, 65(1):011201, 2001.

[108] T. C. Terwilliger and D. Eisenberg. The structure of melittin: I. Structure determination and partial refinement. *J. Biol. Chem.*, 257(11):6010–6015, 1982.

[109] T. C. Terwilliger and D. Eisenberg. The structure of melittin: II. Interpretation of the structure. *J. Biol. Chem.*, 257(11):6016–6022, 1982.

[110] T. Tsuruta and G. Nagayama. Molecular dynamics studies on the condensation coefficient of water. *J. Phys. Chem. B*, 108(5):1736–1743, 2004.

[111] P. Varilly, A. J. Patel, and D. Chandler. An improved coarse-grained model of solvation and the hydrophobic effect. *J. Chem. Phys.*, 134(7):074109, 2011.

[112] C. Vega and E. de Miguel. Surface tension of the most popular models of water by using the test-area simulation method. *J. Chem. Phys.*, 126(15):154707, 2007.

[113] L. Verlet and J.-J. Weis. Equilibrium theory of simple liquids. *Phys. Rev. A*, 5(2):939–952, 1972.

[114] K. Vollmayr-Lee, K. Katsov, and J. D. Weeks. Using mean field theory to determine the structure of uniform fluids. *J. Chem. Phys.*, 114(1):416–425, 2001.

[115] A. Wallqvist and B. J. Berne. Computer simulation of hydrophobic hydration forces on stacked plates at short range. *J. Phys. Chem.*, 99(9):2893–2899, 1995.

[116] F. Wang and D. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.*, 86(10):2050–2053, 2001.

[117] J. D. Weeks. Structure and thermodynamics of the liquid-vapor interface. *J. Chem. Phys.*, 67(7):3106–3121, 1977.

[118] J. D. Weeks. Connecting local structure to interface formation: A molecular scale van der Waals theory of nonuniform liquids. *Annu. Rev. Phys. Chem.*, 53:533–562, 2002.

[119] J. D. Weeks, D. Chandler, and H. C. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.*, 54(12):5237, 1971.

[120] J. D. Weeks, K. Katsov, and K. Vollmayr. Roles of repulsive and attractive forces in determining the structure of nonuniform liquids: Generalized mean field theory. *Phys. Rev. Lett.*, 81(20):4400–4403, 1998.

[121] J. D. Weeks, R. L. B. Selinger, and J. Q. Broughton. Self-consistent treatment of repulsive and attractive forces in nonuniform liquids. *Phys. Rev. Lett.*, 75(14):2694–2697, 1995.

[122] J. D. Weeks, K. Vollmayr, and K. Katsov. Intermolecular forces and the structure of uniform and nonuniform fluids. *Physica A*, 244(1–4):461–475, 1997.

[123] A. P. Willard and D. Chandler. The role of solvent fluctuations in hydrophobic assembly. *J. Phys. Chem. B*, 112(19):6187–6192, 2008.

[124] A. P. Willard and D. Chandler. Coarse-grained modeling of the interface between water and heterogeneous surfaces. *Faraday Discuss.*, 141:209–220, 2009.

[125] A. P. Willard and D. Chandler. Instantaneous liquid interfaces. *J. Phys. Chem. B*, 114(5):1954–1958, 2010.

[126] R. G. Winkler. Virial pressure of periodic systems with long range forces. *J. Chem. Phys.*, 117(5):2449–2450, 2002.

Appendix A

Unbalancing forces in LCW theory

Weeks and coworkers' work on modeling non-uniform liquids [114, 118, 120–122] provides the basis for the term in Equation (2.37) that couples small-length-scale and large-length-scale density fluctuations. This term can be justified rigorously (within certain approximations) in the case of simple (monatomic and isotropic) liquids, as we do in this appendix, and it can be regarded as a useful approximation in the case of more complex liquids like water.

Particles in a simple liquid interact via a pairwise potential $u(r)$ that may be decomposed into a repulsive (positive) portion, $u_0(r)$, and an attractive (negative) portion, $u_a(r)$, where $u_a(r)$ is finite and slowly-varying. A common separation when $u(r)$ is a Lennard-Jones potential is the Weeks-Chandler-Andersen separation [119], described in Section 4.1. In order to model the thermodynamics of non-uniform liquids, Weeks and coworkers considered two related liquid systems: the original liquid of particles that interact via the potential $u(r)$, and a related reference liquid R, whose particles interact with each other via the repulsive potential $u_0(r)$ alone, as well as with an external field $\psi(\mathbf{r})$. In both cases, the chemical potentials are chosen to yield a uniform density ρ_ℓ in the absence of external constraints or other fields. Using the first equation in the Yvon-Born-Green hierarchy with an approximate closure relation, they showed that there exists a unique choice of $\psi(\mathbf{r})$ that results in identical mean densities in both systems, i.e.,

$$\langle \rho(\mathbf{r}) \rangle = \langle \rho(\mathbf{r}) \rangle_R \,. \tag{A.1}$$

If the liquid is close to its triple point, so that its structure is essentially determined by packing constraints [119], then they found that

$$\psi(\mathbf{r}_1) \approx \int_{\mathbf{r}_2} u_a(r_{12})[\langle \rho(\mathbf{r}_2) \rangle - \rho_\ell] = \rho_\ell \int_{\mathbf{r}_2} u_a(r_{12})[\langle n(\mathbf{r}_2) \rangle - 1] + \int_{\mathbf{r}_2} u_a(r_{12}) \langle \delta\rho(\mathbf{r}_2) \rangle \,. \tag{A.2}$$

In order to discuss the consequences of introducing this potential, it is useful to establish some notation. Following Ref. [67], we define the positive moments a and m of the potential $u_a(r)$ by the relations

$$a = -\frac{1}{2}\int_{\mathbf{r}} u_a(r), \tag{A.3}$$

$$m = -\frac{1}{6}\int_{\mathbf{r}} r^2 u(r). \tag{A.4}$$

Given a field $f(\mathbf{r})$, we define the smeared field $\overline{f}(\mathbf{r})$ by the relation

$$\overline{f}(\mathbf{r}_1) = \frac{\int_{\mathbf{r}_2} u_a(r_{12}) f(\mathbf{r}_2)}{\int_{\mathbf{r}_2} u_a(r_{12})}. \tag{A.5}$$

With this notation, we can write

$$\psi(\mathbf{r}) = -2a\big[\langle\overline{\rho}(\mathbf{r})\rangle - \rho_\ell\big] = -2a\rho_\ell\big[\langle\overline{n}(\mathbf{r})\rangle - 1\big] - 2a\rho_\ell\left\langle\overline{\delta\rho}(\mathbf{r})\right\rangle. \tag{A.6}$$

Returning to our discussion, since the reference fluid only has repulsive forces, we can model its thermodynamics with a simple Hamiltonian H_R. In particular, consider splitting the density of the reference system into slowly and quickly varying parts as in Section 2.3. Because of the lack of attractive forces in the reference fluid, the free energy density of the slowly varying part at any given point depends approximately only on the local density and not on its gradients,[1] and its small-length-scale density fluctuations are approximately Gaussian [18], so

$$\begin{aligned} H_R[n(\mathbf{r}), \delta\rho(\mathbf{r})] = &\int_{\mathbf{r}} w_R(n(\mathbf{r})) \\ &+ \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \delta\rho(\mathbf{r})\chi_R^{-1}\big(\mathbf{r},\mathbf{r}';n(\mathbf{r})\big)\delta\rho(\mathbf{r}') + \int_{\mathbf{r}} \psi(\mathbf{r})[\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r})]. \end{aligned} \tag{A.7}$$

Here, $w_R(n)$ is the free-energy density of the uniform reference system at density $\rho_\ell n$. The fluctuation spectrum of the (non-uniform) reference system is $\chi_R\big(\mathbf{r},\mathbf{r}';n(\mathbf{r})\big)$, and as in the main text, we henceforth suppress its dependence on $n(\mathbf{r})$ for clarity. By definition, $\chi_R(\mathbf{r},\mathbf{r}') = \chi_R(\mathbf{r}',\mathbf{r})$.

Equation (A.1) holds only when $\psi(\mathbf{r})$ is chosen self-consistently, i.e. the fields $\langle n(\mathbf{r})\rangle$ and $\langle\delta\rho(\mathbf{r})\rangle$ obtained with the Hamiltonian $H_R[n(\mathbf{r}),\delta\rho(\mathbf{r})]$ are equal to the fields $\langle n(\mathbf{r})\rangle$ and $\langle\delta\rho(\mathbf{r})\rangle$ in the definition of $\psi(\mathbf{r})$. We impose this condition approximately by finding the fields $n(\mathbf{r})$ and $\delta\rho(\mathbf{r})$ that minimize $H_R[n(\mathbf{r}),\delta\rho(\mathbf{r})]$ and identifying them with $\langle n(\mathbf{r})\rangle$ and $\langle\delta\rho(\mathbf{r})\rangle$. To the extent that $H_R[n(\mathbf{r}),\delta\rho(\mathbf{r})]$ is quadratic about this minimum, this

---

[1] Strictly speaking, the free energy as a functional of density of the reference system is unknown. The approximation made here is equivalent to the hydrostatic approximation made by Weeks [118].

approximation is exact. The solution satisfies the relations

$$\begin{aligned} w_R'(n(\mathbf{r}_1)) + \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} \delta\rho(\mathbf{r})\frac{\delta\chi_R^{-1}(\mathbf{r},\mathbf{r}')}{\delta n(\mathbf{r}_1)}\delta\rho(\mathbf{r}') \\ +\rho_\ell^2\int_{\mathbf{r}_2} u_a(r_{12})\big[n(\mathbf{r}_2)-1\big] + \rho_\ell\int_{\mathbf{r}_2} u_a(r_{12})\delta\rho(\mathbf{r}_2) = 0, \end{aligned} \tag{A.8}$$

$$k_{\mathrm{B}}T\int_{\mathbf{r}_2}\chi_R^{-1}(\mathbf{r}_1,\mathbf{r}_2)\delta\rho(\mathbf{r}_2) + \rho_\ell\int_{\mathbf{r}_2} u_a(r_{12})[n(\mathbf{r}_2)-1] + \int_{\mathbf{r}_2} u_a(r_{12})\delta\rho(\mathbf{r}_2) = 0. \tag{A.9}$$

To account for fluctuations around the average behavior, we propose a Hamiltonian for the full fluid, $H[n(\mathbf{r}),\delta\rho(\mathbf{r})]$, whose minimum satisfies these two Equations, regardless of boundary conditions. In other words, the stable states of this Hamiltonian are identical to those of the reference fluid subject to the self-consistent field $\psi(\mathbf{r})$. Such a condition is satisfied by the Hamiltonian

$$\begin{aligned} H[n(\mathbf{r}),\delta\rho(\mathbf{r})] = \int_{\mathbf{r}} w_R(n(\mathbf{r})) + \frac{\rho_\ell^2}{2}\int_{\mathbf{r}_1}\int_{\mathbf{r}_2}\big[n(\mathbf{r}_1)-1\big]u_a(r_{12})\big[n(\mathbf{r}_2)-1\big] \\ + \frac{k_{\mathrm{B}}T}{2}\int_{\mathbf{r}_1}\int_{\mathbf{r}_2}\delta\rho(\mathbf{r}_1)\big[\chi_R^{-1}(\mathbf{r}_1,\mathbf{r}_2;n(\mathbf{r})) + \beta u_a(r_{12})\big]\delta\rho(\mathbf{r}_2) + \int_{\mathbf{r}}\phi(\mathbf{r})\delta\rho(\mathbf{r}), \end{aligned} \tag{A.10}$$

where

$$\phi(\mathbf{r}_1) \approx \rho_\ell\int_{\mathbf{r}_2} u_a(r_{12})[n(\mathbf{r}_2)-1] = -2a\rho_\ell\big[\overline{n}(\mathbf{r}_1)-1\big]. \tag{A.11}$$

This Hamiltonian is defined uniquely up to a physically irrelevant constant, which is chosen here so that when the liquid is uniform, i.e., when $n(\mathbf{r}) = 1$ and $\delta\rho(\mathbf{r}) = 0$, then $H_R[n(\mathbf{r}),\delta\rho(\mathbf{r})] = H[n(\mathbf{r}),\delta\rho(\mathbf{r})]$.

Equation (A.10) can be recast into a physically appealing form by a few manipulations. First, since $n(\mathbf{r})$ varies slowly in space and $u_1(r)$ is short-ranged, we can Taylor expand $n(\mathbf{r}_2)$ about $\mathbf{r}_2 = \mathbf{r}_1$ in the second term of the right-hand side of Equation (A.10) and truncate the expansion to second order [67], so

$$\begin{aligned} &\frac{\rho_\ell^2}{2}\int_{\mathbf{r}_1}\int_{\mathbf{r}_2}\big[n(\mathbf{r}_1)-1\big]u(r_{12})\big[n(\mathbf{r}_2)-1\big] \\ &\quad\approx \frac{\rho_\ell^2}{2}\int_{\mathbf{r}_1}\big[n(\mathbf{r}_1)-1\big]^2\left\{\int_{\mathbf{r}} u(r)\right\} + \frac{\rho_\ell^2}{2}\int_{\mathbf{r}_1}\int_{\mathbf{r}_2}\big[n(\mathbf{r}_1)-1\big]u(r_{12})(\mathbf{r}_{12}\cdot\nabla)\big[n(\mathbf{r}_1)-1\big] \\ &\qquad + \frac{\rho_\ell^2}{4}\int_{\mathbf{r}_1}\int_{\mathbf{r}_2}\big[n(\mathbf{r}_1)-1\big]u(r_{12})(\mathbf{r}_{12}\cdot\nabla)^2\big[n(\mathbf{r}_1)-1\big] \end{aligned} \tag{A.12}$$

In the right-hand side of this expression, the first term depends only locally on $n(\mathbf{r})$, so it can be viewed as a change to the free energy density from $w_R(n)$ to $w(n)$, given by[2]

$$w(n) = w_R(n) - a\rho_\ell^2(n-1)^2. \tag{A.13}$$

The second term on the right-hand side of Equation (A.12) is zero due to symmetry. Finally, the third term can be written as

$$\begin{aligned}\frac{\rho_\ell^2}{4}\int_{\mathbf{r}_1}\int_{\mathbf{r}_2}[n(\mathbf{r}_1)-1]u_1(r_{12})(\mathbf{r}_{21}\cdot\nabla)^2 n(\mathbf{r}_1) &= \frac{\rho_\ell^2}{4}\int_{\mathbf{r}_1}\int_{\mathbf{r}_{21}}[n(\mathbf{r}_1)-1]u_1(r_{12})(\mathbf{r}_{21}\cdot\nabla)^2 n(\mathbf{r}_1)\\ &= \frac{\rho_\ell^2}{12}\int_{\mathbf{r}_1}[n(\mathbf{r}_1)-1]\nabla^2 n(\mathbf{r}_1)\int_{\mathbf{r}_{21}} r_{21}^2 u_1(r_{12}),\\ &= \int_{\mathbf{r}}\frac{m}{2}|\nabla\rho_\ell n(\mathbf{r})|^2, \end{aligned} \tag{A.14}$$

where in the second line, we've used the result that for any tensor $\mathbf{A}$,

$$\int_{\mathbf{r}} u_a(r)\,\mathbf{r}^T\mathbf{A}\mathbf{r} = \frac{1}{3}\operatorname{Tr}\mathbf{A}\int_{\mathbf{r}} r^2 u_a(r). \tag{A.15}$$

In other words, the third term in Equation (A.12) gives rise to a square-gradient term. Finally, we define the symmetric operator $\chi\big(\mathbf{r}_1,\mathbf{r}_2;n(\mathbf{r})\big)$ so that

$$\chi^{-1}\big(\mathbf{r}_1,\mathbf{r}_2;n(\mathbf{r})\big) = \chi_R^{-1}\big(\mathbf{r}_1,\mathbf{r}_2;n(\mathbf{r})\big) + \beta u_1(r_{12}). \tag{A.16}$$

The form of Equation (A.10) indicates that $\chi(\mathbf{r}_1,\mathbf{r}_2)$ is the fluctuation spectrum of the full liquid. Equation (A.16) corresponds to the random phase approximation of standard liquid-state theory (Equations (3.5.9) and (3.5.17) in Ref. [41]).

Assembling the above results, we find that the Hamiltonian of Equation (A.10) may be written in the form

$$\begin{aligned}H[n(\mathbf{r}),\delta\rho(\mathbf{r})] \approx \int_{\mathbf{r}} w(n(\mathbf{r})) + \frac{m}{2}|\nabla\rho_\ell n(\mathbf{r})|^2&\\ + \int_{\mathbf{r}}\int_{\mathbf{r}'}\delta\rho(\mathbf{r})\chi^{-1}(\mathbf{r},\mathbf{r}';n(\mathbf{r}))\delta\rho(\mathbf{r}') + \int_{\mathbf{r}}\phi(\mathbf{r})\delta\rho(\mathbf{r}),&\end{aligned} \tag{A.17}$$

We can identify $w(n(\mathbf{r}))$ as the free energy density of the full fluid, $\chi(\mathbf{r},\mathbf{r}')$ as the fluctuation spectrum of the full fluid and the $\phi(\mathbf{r})\delta\rho(\mathbf{r})$ term as an effective coupling between the small- and large-length-scale density fields. This approximate form of the Hamiltonian, derived

---

[2] The difference between $w(n)$ and $w_R(n)$ has a physical interpretation. The part that is proportional to n^2 corresponds to an additional energy density $-a\rho_\ell^2 n^2$ resulting from solvent-solvent attractions that stabilize the fluid phase, analogous to the energy density in traditional van der Waals theory [93]. The part that is linear in n corresponds to a chemical potential difference $\Delta\mu = -2a\rho_\ell$ between the reference and full fluid. The remaining part reflects a choice of integration constant and is not meaningful.

above for simple dense liquids with isotropic interactions, serves as the starting point for the phenomenological modeling of water in LCW theory.[3]

We conclude this Appendix by making the explicit connection between the original LCW theory, Equations (5) and (6) of Ref. [68], and the formulation used in this thesis. To this end, we use an expansion analogous to that of Equation (A.12) to rewrite Equations (A.8) and (A.9). We obtain

$$w'(n(\mathbf{r}_1)) + \frac{k_\mathrm{B}T}{2}\int_\mathbf{r}\int_{\mathbf{r}'} \delta\rho(\mathbf{r})\frac{\delta\chi_R^{-1}(\mathbf{r},\mathbf{r}')}{\delta n(\mathbf{r}_1)}\delta\rho(\mathbf{r}') - m\rho_\ell^2\nabla^2 n(\mathbf{r}_1) = 2a\rho_\ell\overline{\delta\rho(\mathbf{r}_1)}, \tag{A.18}$$

$$k_\mathrm{B}T\int_{\mathbf{r}_2}\chi^{-1}(\mathbf{r}_1,\mathbf{r}_2)\delta\rho(\mathbf{r}_2) = -\phi(\mathbf{r}_1). \tag{A.19}$$

In the notation of Ref. [68], the slowly varying density field is defined as $n_\mathrm{s}(\mathbf{r}) = \rho_\ell n(\mathbf{r})$, and the free energy density $w_\mathrm{LCW}(n_\mathrm{s})$ is defined as a function of n_s, not n. Hence, $w'_\mathrm{LCW}(n_\mathrm{s}) = w'(n)/\rho_\ell$. Dividing Equation (A.18) throughout by ρ_ℓ, we obtain

$$w'_\mathrm{LCW}(n_\mathrm{s}(\mathbf{r}_1)) + \frac{k_\mathrm{B}T}{2}\int_\mathbf{r}\int_{\mathbf{r}'} \delta\rho(\mathbf{r})\frac{\delta\chi_R^{-1}(\mathbf{r},\mathbf{r}')}{\delta n_\mathrm{s}(\mathbf{r}_1)}\delta\rho(\mathbf{r}') = m\nabla^2 n_\mathrm{s}(\mathbf{r}_1) + 2a\overline{\delta\rho(\mathbf{r}_1)}. \tag{A.20}$$

This Equation corresponds to Equation (5) in Ref. [68] when the dependence of $\chi_R(\mathbf{r},\mathbf{r}')$ on $n_\mathrm{s}(\mathbf{r})$ is neglected.

As discussed in Chapter 4, solvent exclusion from a volume v results in the addition to the right-hand side of Equation (A.19) of a term $-k_\mathrm{B}Tc(\mathbf{r}_1)$ that is zero for $\mathbf{r}_1$ outside v, and chosen so that $\rho_\ell n(\mathbf{r}) + \delta\rho(\mathbf{r}) = 0$ for $\mathbf{r}$ inside v. Rearranging this new equation yields to give

$$\delta\rho(\mathbf{r}_1) = -\int_{\mathbf{r}_2}\chi(\mathbf{r}_1,\mathbf{r}_2)\big[\beta\phi(\mathbf{r}_1) + c(\mathbf{r}_2)\big], \qquad \text{(excluded solvent constraint)}. \tag{A.21}$$

Equation (A.21) thus corresponds to Equation (6) in Ref. [68] if the contribution of the field $\phi(\mathbf{r})$ is neglected with respect to the much larger contribution due to $c(\mathbf{r})$.

Our discussion here shows that the original LCW theory [68] corresponds to a mean-field construction of fields $n(\mathbf{r})$ and $\delta\rho(\mathbf{r})$ that approximately minimize Equation (A.17), as hinted to in Ref. [107].

---

[3] In the main text, a factor of ρ_ℓ^2 has been subsumed into the constant m.