INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **14** (2002) 9445–9460 PII: S0953-8984(02)36792-4

Hydrophobic interactions: an overview

Pieter Rein ten Wolde[1,2]

[1] Division of Physics and Astronomy, Vrije Universiteit, De Boelelaan 1081, 1081 HV, Amsterdam, The Netherlands

Received 9 May 2002, in final form 23 May 2002
Published 27 September 2002
Online at stacks.iop.org/JPhysCM/14/9445

Abstract

We present an overview of the recent progress that has been made in understanding the origin of hydrophobic interactions. We discuss the different character of the solvation behaviour of apolar solutes at small and large length scales. We emphasize that the crossover in the solvation behaviour arises from a collective effect, which means that implicit solvent models should be used with care. We then discuss a recently developed explicit solvent model, in which the solvent is not described at the atomic level, but rather at the level of a density field. The model is based upon a lattice-gas model, which describes density fluctuations in the solvent at large length scales, and a Gaussian model, which describes density fluctuations at smaller length scales. By integrating out the small-length-scale field, a Hamiltonian is obtained, which is a function of the binary, large-length-scale field only. This makes it possible to simulate much larger systems than was hitherto possible as demonstrated by the application of the model to the collapse of an ideal hydrophobic polymer. The results show that the collapse is dominated by the dynamics of the solvent, in particular the formation of a vapour bubble of critical size. Implications of these findings for the understanding of pressure denaturation of proteins are discussed.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Hydrophobic interactions are widely believed to play a dominant role in the formation of large biological structures [1, 2]. Yet, the mechanism of the hydrophobic effect is still under debate. The solvation of small apolar species is well understood [3–6]. However, the attraction between two such species in water is weak [3, 4, 6] and probably not responsible for the stability of biological structures. On the other hand, strong and long-ranged attractions have been measured between extended hydrophobic surfaces [7, 8]. But here, the origin of the effect is still being discussed. It has been suggested that the interaction arises from electrostatic

[2] Address for correspondence: FOM-Instiute for Atomic and Molecular Physics (AMOLF), Kruislaan 407, 1098 SJ, Amsterdam, The Netherlands.

fluctuations [9], changes in water structure [10], bridging (sub)microscopic bubbles [11, 12], and a 'drying' transition induced by the hydrophobic surfaces [13–15].

Recently, Lum *et al* [15] have developed a theory of hydrophobicity which suggests that the nature of the hydrophobic effect precisely arises from the interplay of density fluctuations at both small and large length scales. Small apolar species only affect density fluctuations in water at small length scales. Concomitantly, water can only induce a relatively weak and short-ranged attraction between small hydrophobic objects. In contrast, large hydrophobic species can affect density fluctuations at large length scales. At ambient conditions, water is close to phase coexistence. A sufficiently large hydrophobic object or, more importantly, an assembly of several small apolar species, can therefore induce a depletion of water relative to the bulk density [16, 17]. Recently, it has been demonstrated that this drying transition can induce a strong attraction between hydrophobic objects and provide a strong driving force for protein folding [18].

In this paper, we give an overview of the recent progress that has been made in understanding the origin of the hydrophobic effect. In section 2 we discuss the statistics of density fluctuations at small and large length scales. Understanding these fluctuations is important, because it provides insight not only into how the solvation free energy of solutes scales with their size, but also into which models are needed to describe their solvation behaviour. In the next section, we briefly discuss a recently developed model. In section 4, we see that this model gives a reasonable prediction of the solvation free energy of apolar solutes. In particular, it predicts that in the small-length-scale regime the solvation free energy scales with the size of the excluded volume of the solute, whereas in the large-length-scale regime it scales with the area of the excluded volume. The model also shows that the crossover in the solvation free energy arises from a collective effect in the solvent. Implicit solvent models cannot conveniently describe this collective effect. It thus appears that explicit solvent models are needed. Explicit atomistic solvent models, however, are computationally demanding. The model discussed in section 3 lays the foundation for a scheme in which the solvent is not described at the atomic level, but rather at the level of a coarse-grained density field. This scheme, which is discussed in section 5, allows us to simulate the solvent much more efficiently.

In section 6 we discuss the role of attractive interactions between the solutes and the solvent. Most of the theoretical work on the hydrophobic effect has focused on the solvation behaviour of ideal hydrophobic solutes [15, 19, 20]. These are objects that exclude solvent from a certain region in space, but have no attractive interactions with the solvent. In section 6, however, it is seen that the presence of weak dispersive interactions does not significantly affect the solvation behaviour of hydrophobic solutes.

Finally, in section 7 we apply the scheme discussed in section 5 to study the collapse of an ideal hydrophobic polymer. The simulations reveal that the dynamics of the collapse transition is dominated by the dynamics of the solvent. In particular, the rate-limiting step is the formation of a vapour bubble of critical size. In addition, we show that during the collapse the chain and the solvent remain out of equilibrium. Both observations imply that implicit solvent models should be used with great care. In the past, a statistically meaningful study of protein folding using explicit solvent models seemed impractical. The current analysis, however, suggests that such studies become feasible by using a statistical field model like the one presented in sections 3 and 5.

2. Density fluctuations at small and large length scales

A theory of hydrophobicity should be able to correctly predict the solvation free energy of apolar solutes. The solvation free energy of a hard sphere, an example of an ideal hydrophobic object, is given by

$$\Delta\mu(v_{\mathrm{ex}}) = -k_B T \ln[P(N = 0;\, v_{\mathrm{ex}})]. \tag{1}$$

Here $k_B T$ is Boltzmann's constant times temperature and $P(N;\, v_{\mathrm{ex}})$ is the probability of finding N solvent particles inside a region, of volume v_{ex}, from which the solvent would be excluded by the solute; note that $P(N = 0;\, v_{\mathrm{ex}})$ corresponds to the insertion probability of a hard sphere. Computer simulations have revealed that density fluctuations that are entropic in origin obey Gaussian statistics. For the hard-sphere fluid, the distribution $P(N;\, v_{\mathrm{ex}})$ (as a function of N) is found to be almost exactly Gaussian for spherical volumes v_{ex} that have a diameter d up to four times the diameter of the solvent molecules, σ (i.e. hard-sphere diameter) [21]. Conceivably, the distribution functions remain Gaussian for volumes that are larger than those that could be studied in the simulations, although even larger voids could induce the formation of a solid–fluid-like interface if the fluid is close to freezing. On the other hand, for volumes in SPC/E water at ambient conditions [6] and a Lennard-Jones liquid close to phase coexistence with the vapour [22], the distribution $P(N;\, v_{\mathrm{ex}})$ remains Gaussian up to approximately $d = (1.0\text{–}2.0)\sigma$. For such small volumes, the occupation statistics is dominated by entropic effects. The solvation free energy of small solutes is determined by the entropic cost of permitting only those density fluctuations that do not violate the effect of the strong forces that are present in the system. In a simple fluid, these strong forces arise from the repulsive interactions between the particles; in water, these are the hydrogen bond forces. As the occupation statistics is Gaussian for small volumes, the insertion probability $P(N = 0;\, v_{\mathrm{ex}})$, and hence the solvation free energy of small solutes, can be obtained from a Gaussian theory. Examples of such Gaussian theories are the Pratt–Chandler theory of hydrophobicity [3, 5] and the theory of Hummer, Pratt and co-workers [6, 23].

For larger volumes the situation changes significantly. In a simple fluid, the attractive forces between the solvent molecules no longer cancel each other when the volume v_{ex} is excavated. The unbalanced attractive forces give rise to an unbalancing potential, which tends to push the solvent molecules away from the void. For water, a large excavated volume will drastically disrupt the water structure. Close proximity of water to the void is energetically unfavourable, because the hydrogen bond network can no longer be maintained close to the surface of the void, as pointed out by Stillinger [16]. In both cases, solvent density is depleted near the surface of the void. This drying transition is a collective effect and can be interpreted as a microscopic manifestation of a phase transition. A Gaussian model cannot support such a transition as it only describes density fluctuations at microscopic length scales. For larger solutes, however, coarse-grained models, such as a Landau–Ginzburg model or a lattice-gas model, become useful. Such coarse-grained models describe density fluctuations at length scales larger than the bulk correlation length and can support phase transitions and sustain gas–liquid interfaces. The essence of our model, discussed in section 3, is to combine a Gaussian model with a lattice-gas model.

3. Theory of solvation

Water at ambient conditions is a cold liquid. A cold liquid is a fluid that is well below the critical temperature. For such a cold liquid, the density of the vapour is typically orders of magnitude lower than that of the liquid. It is then natural to divide space into a (cubic) grid of cells and, taking the grid spacing, l, to be of the order of the bulk correlation length, ξ, describe the states of the cells by a binary field, n_i, where $n_i = 1$ if cell i is filled with 'liquid' and $n_i = 0$ if cell i contains 'vapour' [20]. This binary field describes density fluctuations at length scales larger than the grid spacing and it can support phase transitions. Density fluctuations at length scales smaller than the grid spacing are described by a second field, $\delta\rho$. This field does not support phase transitions, but it can describe the microscopic granularity of matter. As discussed above, we can assume that this field obeys Gaussian statistics. Hence, the full density $\rho(r_i)$ at point r_i is decomposed as

$$\rho(r_i) = \rho_l n_i + \delta\rho(r_i), \tag{2}$$

where ρ_l is the bulk liquid density.

The partition function for our model is

$$\Xi = \sum_{\{n_i\}} \int \mathcal{D}\delta\rho(r)\, C[\{n_k\}, \delta\rho(r)] \exp(-\beta H[\{n_k\}, \delta\rho(r)]), \tag{3}$$

where $\int \mathcal{D}\delta\rho(r) = \int \prod_i \mathcal{D}\delta\rho(r_i)$ denotes the functional integration over the small-length-scale field, $H[\{n_k\}, \delta\rho(r)]$ is the Hamiltonian as a functional of both n_i and $\delta\rho(r)$, and β^{-1} is Boltzmann's constant times temperature, $k_B T$. The quantity $C[\{n_k\}, \delta\rho(r)]$ is a constraint functional. It has unit weight when the field $\delta\rho(r)$ together with $\{n_i\}$ satisfy whatever constraints are imposed by strong forces, and it is zero otherwise. Since $\{n_i\}$ and $\delta\rho(r)$ have greatly different character, the summation and integration in equation (3) do not redundantly count configuration space to any significant degree.

The Hamiltonian of our model is given by the following expression:

$$\begin{aligned} H[\{n_k\}, \delta\rho(r)] = {} & H_L[\{n_k\}] - \epsilon' \sum_{i,j\,(\mathrm{nn}i)} \int \mathrm{d}r_i\, \delta\rho(r_i) \frac{n_j - 1}{\rho_l l^3} \\ & + \frac{k_B T}{2} \sum_{i,j} \int \mathrm{d}r_i \int \mathrm{d}r'_j\, \delta\rho(r_i)\, \chi^{-1}[r_i, r'_j; \{n_k\}]\, \delta\rho(r'_j) + H_{\mathrm{norm}}[\{n_k\}]. \end{aligned} \tag{4}$$

In the above expression, H_L is the Hamiltonian of a lattice-gas model. It is given by

$$H_L[\{n_k\}] = -\mu \sum_i n_i - \epsilon \sum_{\langle i,j \rangle} n_i n_j, \tag{5}$$

with μ the imposed chemical potential, and ϵ the interaction parameter between nearest-neighbour cells. The lattice-gas model describes density fluctuations at length scales larger than the grid spacing and it can support phase transitions. Importantly, the energetic cost of creating an interface is determined by the value of ϵ.

The term in equation (4) that is quadratic in $\delta\rho$ ensures the Gaussian weight of the small-length-scale field. In particular, the response function $\chi^{-1}[r_i, r'_j; \{n_k\}]$ determines the variance of this field. It depends upon the state of the large-length-scale field. If $n_i = 1$ for all i, then $\chi^{-1}[r_i, r'_j; \{n_k\}]$ reduces to the response function of the bulk liquid $\chi^{-1}[r_i, r'_j; \rho_l]$. In general, we employ the approximation that $\delta\rho(r_i) = 0$ whenever $n_i = 0$. Thus, $\delta\rho(r)$ is a Gaussian field, with its weight functional being that of the bulk liquid, but constrained to be zero whenever $n_i = 0$ [20].

The second term on the right-hand side of equation (4) is the term that couples the fluctuations of the field n_i with those of the field $\delta\rho(r)$. As discussed by Lum *et al* and Weeks and co-workers [15, 24], for simple fluids the parameter ϵ' is related to the energy density of the bulk liquid. The last term on the right-hand side of equation (4), H_{norm}, is a renormalization term, which ensures that, if the field $\delta\rho$ is integrated out in the *absence* of solutes, the Hamiltonian $H[\{n_k\}; \delta\rho]$ reduces to that of the lattice-gas model, $H_L[\{n_k\}]$.

In the *presence* of a hard sphere, the constraint functional in equation (3) is given by

$$C[\{n_k\}; \delta\rho(r)] = \prod_{r_i \in v_{\mathrm{ex}}} \delta[n_i \rho_l + \delta\rho(r_i)]. \tag{6}$$

By exploiting the Fourier representation of delta functions, we can integrate out the field $\delta\rho(\boldsymbol{r})$. As discussed in detail in [20], we then arrive at the following expression for the Hamiltonian:

$$\begin{aligned} H[\{n_k\}] &= H_L[\{n_k\}] + k_B T \sum_{i,j\,(\mathrm{occ})} \frac{n_i[\rho_l v_i + f_i][\rho_l v_j + f_j]n_j}{2\sigma_{v_{\mathrm{ex}}}} + k_B T \ln\sqrt{2\pi\sigma_{v_{\mathrm{ex}}}} \\ &\equiv H_L[\{n_k\}] + H_S[v_{\mathrm{ex}};\{n_k\}]. \end{aligned} \tag{7}$$

Here,

$$f_i \equiv \int_{\boldsymbol{r}_i \in v_{\mathrm{ex}}} \mathrm{d}\boldsymbol{r}_i\, f(\boldsymbol{r}_i) = n_i v_i \epsilon' \kappa \frac{\rho_l}{l^3} \sum_{k\,(\mathrm{nn}\,i)} (n_k - 1), \tag{8}$$

with κ the isothermal compressibility, and

$$\sigma_{v_{\mathrm{ex}}} = \int_{v_{\mathrm{ex}}} \mathrm{d}\boldsymbol{r} \int_{v_{\mathrm{ex}}} \mathrm{d}\boldsymbol{r}'\, \chi(\boldsymbol{r}, \boldsymbol{r}'; \rho_l). \tag{9}$$

The sum over i, j (occ) is over cells i and j that are occupied by the solute, and v_i is the volume occupied by the solute in cell i [20].

The term $H_S[v_{\mathrm{ex}};\{n_k\}]$ contains all the effects of the interaction between the ideal hydrophobic solute and the solvent. We stress that the interaction term solely arises from the constraint that is imposed upon the allowed density fluctuations in the solvent. The excess chemical potential of the solute may be obtained by averaging this interaction free energy:

$$\beta\,\Delta\mu(v_{\mathrm{ex}}) = -\ln\langle\exp(-\beta H_S[v_{\mathrm{ex}};\{n_k\}])\rangle_L, \tag{10}$$

where $\langle\cdots\rangle_L$ denotes an ensemble average with the Hamiltonian of the lattice gas, $H_L[\{n_k\}]$ (see equation (5)).

4. Scaling behaviour of solvation free energies

Figure 1 shows the excess chemical potential of a hard sphere as a function of its radius, R, in water at ambient conditions, as predicted by equations (7) and (10). For comparison, we also show the results of a Gaussian model and the results of a molecular simulation of a hard sphere in SPC/E water [22].

4.1. *Small-length-scale regime*

In the small-length-scale regime, our model reduces to a Gaussian model. Small solutes only affect density fluctuations in the solvent at small length scales. As a consequence, $n_i = 1$ for nearly all i. With $n_i = 1$ for all i, $H_L[\{n_k\}]$ and $H_{\mathrm{norm}}[\{n_k\}]$ in equation (4) become constants and thus irrelevant. Further, the response function reduces to that of the uniform fluid and the term that couples the density fluctuations of the small- and large-length-scale fields becomes zero. As such, the Hamiltonian reduces to that of a Gaussian model [5]:

$$H_G[\delta\rho(\boldsymbol{r})] = \frac{k_B T}{2} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \delta\rho(\boldsymbol{r})\, \chi^{-1}(\boldsymbol{r}, \boldsymbol{r}'; \rho_l)\, \delta\rho(\boldsymbol{r}'), \tag{11}$$

with $\delta\rho(\boldsymbol{r}) = \rho(\boldsymbol{r}) - \rho_l$, and the response function $\chi^{-1}(\boldsymbol{r}, \boldsymbol{r}'; \rho_l)$ being the response function of the uniform fluid. The Pratt–Chandler theory of hydrophobicity is based upon such a Gaussian model [3]. Similarly, we can directly obtain the solvation free energy of a Gaussian model from equation (10). It is given by

$$\beta\,\Delta\mu(v_{\mathrm{ex}}) \simeq \rho_l^2 v_{\mathrm{ex}}^2 / 2\sigma_{v_{\mathrm{ex}}} + \ln\sqrt{2\pi\sigma_{v_{\mathrm{ex}}}}. \tag{12}$$

This is also the result for the solvation free energy of the theory of Hummer *et al* [6, 23].

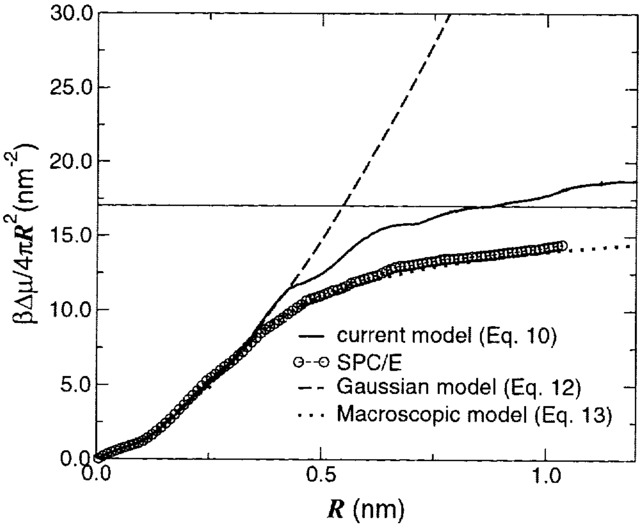

Figure 1. Comparison of the results of the full model, based on equations (7) and (10), with the predictions of the Gaussian model, equation (12), the macroscopic model, equation (13), and the results of a molecular simulation of a cavity in SPC/E water [17] for the excess chemical potential per unit area of a hard sphere in water as a function of its size; the hard sphere excludes water from a spherical volume of radius R. The horizontal line lies at the value of the surface tension γ of the vapour–liquid interface of water. The parameters of the full model are: $l = 4.2$ Å, $\epsilon = 6.02k_BT$, $\epsilon' = 15.2k_BT$, and $\mu - \mu_{\rm coex} = 5.5 \times 10^{-4}k_BT$. These parameters are chosen such that the surface tension, the energy density, and the imposed chemical potential of the lattice-gas model match the corresponding values for water (see [20]).

Figure 1 shows that the full model and (thus) the Gaussian model give an excellent prediction for the solvation free energy of small solutes. The reason, as discussed in more detail in section 2, is that at small length scales, density fluctuations are highly Gaussian. A second point to note is that for small solutes, the excess chemical potential scales with the excluded volume of the solute. In this regime, the excess chemical potential is dominated by the entropic cost of constraining small-length-scale fluctuations, which is proportional to the volume of the space over which the solvent density fluctuations are constrained.

4.2. *Large-length-scale regime*

For solutes with $R > 4$ Å, the prediction of the Gaussian model diverges from both the full model and the molecular simulation results. The divergence is due to a cavitation transition in the solvent. This cavitation transition can be understood as a microscopic liquid–gas phase transition in the solvent that is induced by the solute. In order to describe the solvation behaviour of large hydrophobic objects in this 'drying' regime, a theory of solvation should be able to support a liquid–gas phase transition and capture the enthalpic effects of creating a gas–liquid interface. The full model discussed in the previous section is such a model. A Gaussian model, on the other hand, cannot support a gas–liquid phase transition, nor sustain a gas–liquid interface, as it is based upon a density expansion around the uniform fluid.

Figure 1 shows that in the large-length-scale regime, the molecular simulation results for the excess potential are well described by [17]

$$\frac{\Delta\mu(R)}{4\pi R^2} \approx \frac{pR}{3} + \tilde{\gamma}\left(1 - \frac{2\delta}{R}\right). \tag{13}$$

The first term on the right-hand side is the work needed to expand a cavity against the external pressure, p. The second term on the right-hand side of the above equation describes the work needed to form an interface between the solute and the solvent, including a first-order 'Tolman' correction due to the curvature of the interface.

At ambient conditions, the pressure is very small, and the pV contribution to the free energy is only significant for extremely large solute volumes. Moreover, as discussed in detail in [17], $\tilde{\gamma}$ is well approximated by the surface tension of the vapour–liquid interface at coexistence. Indeed, for large (but not unphysically large) solutes, the solvation free energy is dominated by the work needed to create a gas–liquid interface.

It is often assumed that the solvation free energy of a hydrophobic species is proportional to the exposed surface area [25, 26]. The molecular simulations, as well as the theoretical analyses, indicate that for solutes of biologically relevant size, this is a reasonable assumption in the drying regime. For small solutes, however, the excess chemical potential of a hydrophobic species does not scale with the area of the interface between the solute and the solvent. Rather, in the small-length-scale regime, the excess chemical potential scales with the excluded volume of the solute. Finally, we point out that while macroscopic thermodynamic descriptions such as that underlying equation (13) are useful for understanding the solvation behaviour of large solutes ($R > 8$–10 Å), they are of limited use for describing the solvation behaviour of small solutes. In the small-length-scale regime, microscopic models, such as the Gaussian model discussed in the previous section, are required.

5. Langevin dynamics in Ising solvent

The model discussed in section 3 can be used to develop a scheme in which only the solutes are treated at the atomic level; the solvent is described in terms of the binary, large-length-scale density field n_i. In this way, the solvent can be simulated much more efficiently than using an explicit atomistic solvent model. The scheme is generally applicable to describe the motion of a collection of solutes, but here we will confine our attention to the motion of an ideal hydrophobic polymer consisting of N_s hard spheres.

Before we discuss the scheme in detail, we point out that, in practice, it is reasonable to simplify the Hamiltonian in equation (4) by neglecting the relatively small off-diagonal elements [18]. This yields the following Hamiltonian for the coarse-grained density field:

$$H(\boldsymbol{r}^{N_s}; \{n_k\}) = \sum_i (-\mu + \Delta\mu_{\mathrm{ex}}(v_i))n_i - \epsilon \sum_{\langle i,j \rangle} n_i n_j. \tag{14}$$

In the above expression, $v_i = \sum_{s=1}^{N_s} v_{s,i}$, where $v_{s,i}$ is the volume that monomer s occupies in cell i and N_s is indeed the number of monomers with coordinates $\boldsymbol{r}^{N_s}$. Further, $\Delta\mu_{\mathrm{ex}}(v_i)$ is the reversible work needed to accommodate a volume v_i in the solvent. We have taken $\Delta\mu_{\mathrm{ex}}$ to be proportional to the excluded volume:

$$\Delta\mu_{\mathrm{ex}}(v_i) \approx c v_i, \tag{15}$$

with $c = 2.67 \times 10^8$ J m^{-3} $= 65 k_B T$ nm^{-3} at room temperature.

From $H(\boldsymbol{r}^{N_s}; \{n_k\})$, we can construct a free-energy functional, Ω, for the field $\{n_k\}$ in the mean-field approximation:

$$\Omega(\boldsymbol{r}^{N_s}; \{n_k\}) = \sum_i \{k_B T(\langle n_i \rangle \ln\langle n_i \rangle + (1 - \langle n_i \rangle)\ln(1 - \langle n_i \rangle)) - (\mu - \Delta\mu_{\mathrm{ex}}(v_i))\langle n_i \rangle\} - \frac{\epsilon}{2} \sum_{\langle i,j \rangle} \langle n_i \rangle \langle n_j \rangle. \tag{16}$$

We can now construct a Car–Parrinello scheme [27, 28] to propagate the solvent. To this end, we define the Lagrangian:

$$\mathcal{L}(r^{N_s}; \{n_k\}) = \frac{1}{2}\sum_{\alpha=1}^{N_s} M\dot{r}_\alpha^2 - V_{SS}(r^{N_s}) + \frac{1}{2}\sum_i m\langle\dot{n}_i\rangle^2 - \Omega(r^{N_s}; \{\langle n_k\rangle\}). \tag{17}$$

Here $V_{SS}(r^{N_s})$ is the intra-chain potential that describes the direction interaction between the monomers, M is the mass of the monomers, and m is a fictitious mass associated with the dynamical variable $\langle n_i\rangle$. From the Lagrangian, the following equations of motion for the monomers and the variables $\langle n_i\rangle$ are obtained:

$$M\ddot{r}_\alpha = -\nabla_\alpha\mathcal{L} = -\nabla_\alpha[V_{SS}(r^{N_s}) + \Omega((r^{N_s}); \{\langle n_k\rangle\})] \tag{18}$$

$$m\langle\ddot{n}_i\rangle \equiv \frac{\partial\mathcal{L}}{\partial\langle n_i\rangle} = -\frac{\partial\Omega(r^{N_s}; \{n_k\})}{\partial\langle n_i\rangle}. \tag{19}$$

Propagating the solvent field $\{n_k\}$ according to the above equations of motion, with a proper choice of both the fictitious mass m and the kinetic energy associated with $\{\langle n_k\rangle\}$, ensures that the free energy of the solvent is close to its minimum for each configuration of the polymer [28]. This approach is similar in spirit to the method employed by Löwen *et al* [29] to simulate counterion screening in colloidal suspensions of polyelectrolytes. However, as we will demonstrate below, we cannot assume that the chain moves slowly on timescales for which disturbances in the solvent may relax. As we will see, the collapse of the polymer is driven by a cavitation transition in the solvent. Such a rare event cannot be captured by the above Car–Parrinello scheme as nucleation barriers cannot be crossed. This means that, using a Car–Parrinello scheme, the solvent would remain in its metastable state until it would become unstable. We therefore developed a novel scheme in which the dynamics of the solvent and the chain are treated together.

First, it should be realized that the small-length-scale field, $\delta\rho$, has been integrated out. This means that trajectories are only true to nature on timescales larger than those required to relax the small-length-scale field $\delta\rho$. This relaxation time is of the order of picoseconds [30], which implies that the dynamics is diffusive. We thus constructed a stochastic dynamics scheme. The elementary step of the algorithm consists of propagating the polymer for M_S Langevin steps in the field of constant $\{n_k\}$, followed by a full Glauber sweep [31] over the solvent field $\{n_k\}$. One Langevin step corresponds to

$$r_\alpha(t + \delta t_s) = r_\alpha(t) + \frac{\delta t_s}{\gamma}(-\nabla_\alpha[V_{SS}(r^{N_s}) + H(r^{N_s}; \{n_k\})] + \delta F) \tag{20}$$

where γ is a friction coefficient, the value of which can be obtained from the diffusion constant of a single monomer in water. The beads also experience a random force δF; it is the dynamical remnant of the small-length-scale field. In order to obtain a physically meaningful value for M_s, we have to compare the time step for the propagation of the polymer, δt_s, to the timescale that corresponds to a Glauber sweep over the solvent variables. An estimate for the latter can be obtained by estimating the correlation time for a density fluctuation of length scale l: $\delta t_l = 1/[D(2\pi/l)^2]$, where D is the self-diffusion constant of liquid water. The value of M_s is thus given by $M_s = \delta t_l/\delta t_s = 36$.

6. Attractions

So far, we have discussed the solvation of ideal hydrophobic solutes—objects that have no attractive interactions with the solvent molecules. Naturally, real hydrophobic molecules have some affinity for water, and vice versa, because of the ubiquitous van der Waals interactions.

Huang and Chandler [32] have studied in detail the effect of weak solute–solvent attractions on the solvation of non-polar molecules in water. Their analysis was performed using an extension and improved parametrization of the theory of Lum *et al* [15].

Huang and Chandler showed that for small solutes, the presence of attractions has little effect on the solvent density around the solutes, as has previously been appreciated theoretically [3] and as observed in simulations [33]. The solvation behaviour can be understood from the observation that the attractive forces are very small in comparison with hydrogen bond forces, and their effects can be estimated by assuming that the water structure around a hard sphere is unaltered on adding an attraction to water. In particular, attractive interactions should produce a simple additive contribution to the solvation enthalpy, but no significant effect on the solvation entropy.

For large hydrophobic species, attractions do have a notable effect. In the case of hard spheres, i.e. without attractions, the solvent density near the surface of the solute is strongly depleted relative to that in the bulk liquid. In contrast, in the presence of attractions, the solvent density near the surface of the solute is close to that in the bulk liquid ($g(R^{+}) \approx 1.24$, where $g(r)$ is the solvent radial distribution function and R is the radius of the hydrophobic sphere). This, however, is a result of the fact that the drying layer is very compressible. In particular, there is little free-energy cost to moving the vapour–liquid interface. Due to this small energetic cost, the addition of an attractive potential as weak as that between alkanes and water is sufficient to draw the drying layer into contact with the hydrophobic surface.

While drawn into contact with the hydrophobic surface, this interface is distinct from the interface that surrounds a small hydrophobic solute. The contact values for $g(r)$ for the large hydrophobic solutes are close to one, while the small hydrophobic solutes have contact values larger than two. More importantly, as for ideal hydrophobic objects, the solvation free energy scales with the size of the excluded volume for smaller solutes, whereas the solvation free energy scales with the area of the excluded volume for larger solutes. In addition, the crossover from the small- to the large-length-scale regime is around 8–10 Å for both the ideal hydrophobic objects and the hydrophobic solutes that have attractive interactions with the solvent. Thus, the scaling behaviour of the solvation free energy is not affected by the presence of weak attractions between the solute and the solvent. We therefore expect weak dispersion forces not to have a strong effect on the role of hydrophobic forces in biological self-assembly.

7. Hydrophobic polymer collapse

Hydrophobic interactions have long been considered to play an important role in protein folding. We have performed computer simulations to study the collapse of a polymer consisting of $N_s = 12$ hard spheres in water. Even though this allows us to study the effect of hydrophobic interactions on protein folding in perhaps its most basic form, the conventional approaches will not work. The theoretical analyses, as discussed above, suggest that the hydrophobic effect arises from a collective effect in the solvent. Implicit solvent models cannot conveniently capture this effect. It thus appears that explicit solvent models are required. Explicit solvent models, however, are computationally demanding. The analysis is further complicated by the fact that the crossover from solvation in the small-length-scale regime to solvation in the large-length-scale regime occurs at around a nanometre. This implies that in order to study the hydrophobic effect, large system sizes are required. For the required system sizes, an explicit atomistic model of water would not be tractable in the simulations. In contrast, the scheme discussed in section 5 makes it possible to study the polymer collapse in water.

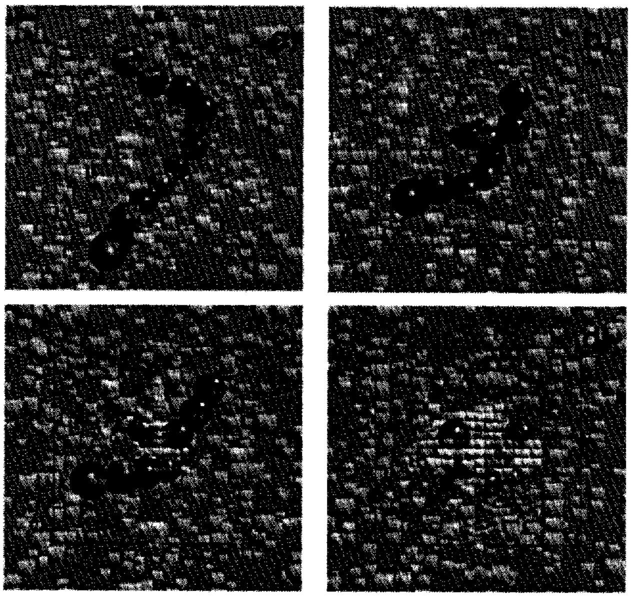

Figure 2. Four configurations from a trajectory where a 12-unit hydrophobic chain in water goes from the coil to the globule state. (a) shows a configuration from the equilibrated coil. The chain remained in configurations like that throughout a 10 ns run at room temperature ($T = 0.663\epsilon$). On a much longer timescale, about 10^{-5} s, the chain typically does exhibit a transition from coil to globule. Such events are found with transition path sampling, equilibrating from an initial high-temperature ($T = 0.74\epsilon$) 10 ns trajectory that exhibited the transition spontaneously. Three configurations from an equilibrated 1.5 ns trajectory that exhibits the collapse transition at room temperature are shown in (b)–(d), with that in (c) being a configuration from the transition state surface. The transparent cubes denote the vapour cells. Those seen far from the chain are typical spontaneous density fluctuations in bulk liquid water. The size of the simulation box is 397 nm^3, corresponding to 42 875 cells. The potential energy function is given by $V_{SS}(r^{N_s}) + H(r^{N_s}; \{n_k\})$. The latter term describes the solvent and the interaction between the polymer and the solvent; it is given by equation (14). The parameters are $l = 0.21$ nm, $\mu - \mu_{\text{coex}} = 2.25 \times 10^{-4} k_B T$, and $\epsilon = 1.51 k_B T$. The intra-chain potential, $V_{SS}(r^{N_s})$, is a function of the positions of the centres of each of the 12 hydrophobic spheres (the dark (red) particles in the figure). It contains three parts: (1) steep (essentially hard-sphere) repulsions between solute particles such that their interparticle separations are larger than $\sigma = 0.72$ nm; (2) stiff harmonic potentials bonding adjacent particles in the hydrophobic chain, $\frac{1}{2}k_s(\sigma - |r_{\alpha+1} - r_\alpha|)^2$, with $k_s = 14.1$ J m^{-2}; (3) a bending potential favouring an extended chain, $\frac{1}{2}k_\theta\theta_\alpha^2$, where θ_α is the angle between $(r_{\alpha+2} - r_{\alpha+1})$ and $(r_{\alpha+1} - r_\alpha)$, and $k_\theta = 1.85 \times 10^{-20}$ J rad^{-2}. The volumes v_i excluded from water by the chain are dynamic as they change with changing chain configuration, i.e., $v_i = v_i(\{r_\alpha\})$. Specifically, these volumes are computed by assuming water molecules have van der Waals radii equal to 0.14 nm, and that the diameter of each hydrophobic unit is $\sigma = 0.72$ nm. That is, points in the excluded volume, r, are those in the union of all volumes inscribed by $|r - r_\alpha| < 0.5$ nm, $\alpha = 1, 2, \ldots, N_s$.

Brute-force simulations, confirmed by the analysis discussed below, suggest that the collapse transition is a rare event. We therefore performed transition-path-sampling simulations [34] to harvest the rare, but representative, trajectories from the extended coil to the collapsed globule state. Figure 2 shows an example of such a trajectory at room temperature. Initially, the polymer is in the extended coil state. Then, by some spontaneous fluctuation, the polymer partly collapses. The collapsed section forms a sufficiently large hydrophobic cluster that a vapour bubble is nucleated. Finally, the vapour bubble drives the beads of the polymer together.

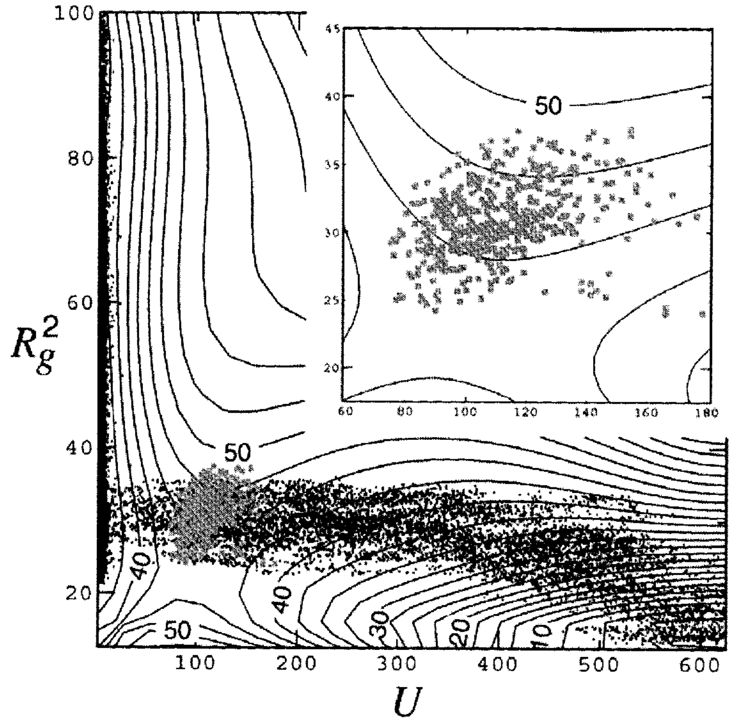

Figure 3. A contour plot of the free-energy landscape for the collapse of our hydrophobic polymer, computed by Monte Carlo umbrella sampling [28], using the weight functional $\exp[-\beta(V_{SS}(r^{N_s})+H(r^{N_s};\{n_k\}))]$. The curves of constant free energy are drawn as a function of the squared radius of gyration of the polymer, R_g^2, and the size of the largest bubble in the system, U. Here we have used the 'cluster' criterion that two vapour cells that are nearest neighbours belong to the same bubble; the size of the bubble is given by the number of vapour cells. Neighbouring lines of the free-energy landscape are separated by $2.5k_BT$. Superimposed is a scatter plot (in black) of the harvested 150 ps trajectories going from the coil to the globule state. The transition states are indicated in grey (green). The harvesting was performed with transition path sampling, making 8400 moves in trajectory space, of which 75% were shooting and 25% were shifting [34]. We find that the plateau regime of the flux correlation function is reached after 50–70 ps [34, 40], implying that the typical commitment time for trajectories to pass over the barrier is of the order of 0.1 ns. Given this time and the fact that the figure shows the free-energy barrier separating the extended coil and compact globule states to be about $9k_BT$, the half-life of the extended chain is about 0.1 ns $\times \exp(9) \approx 10^{-5}$ s.

7.1. Free energy of collapse

Visual inspections of this and similar trajectories suggest that the collapse arises from an interplay between the size of the polymer and the formation of a vapour bubble. We therefore mapped out the free-energy landscape as a function of the squared radius of gyration of the polymer, R_g^2, and the size of the largest vapour bubble present in the system, U. A contour plot of the free-energy landscape is shown in figure 3. It is seen that the path from the coil to the collapsed globule is one where, initially, the radius of gyration decreases, while the size of the largest vapour bubble is still essentially zero. In this regime, the solvent still wets the polymer and the free energy hardly changes. When the radius of gyration becomes small enough, however, a vapour bubble is formed. The free energy now sharply increases by some $9k_BT$, until it reaches a saddle point at $(U, R_g^2) = (98, 23.5l^2)$. From here on, the bubble grows spontaneously and drives all the beads of the polymer together.

A close examination of the free-energy landscape reveals that there is a small barrier of some $2k_BT$ at $(U, R_g^2) = (6.5, 70.0l^2)$, which separates the coil from a metastable intermediate. The presence of these two states arises from a competition between the entropy of the chain, which favours the fully extended coil state, and depletion forces, which favour the intermediate state. Depletion forces are caused by the reduction in the volume from which the solutes exclude the solvent, when the solutes come together. The attraction between two small hydrophobic objects, which cannot induce a drying transition in the solvent, predominantly arises from this effect [3, 4, 6]. However, the free-energy landscape shows that this driving force is relatively small. The large driving force for the collapse of the polymer comes from the drying transition. The free-energy difference between the fully collapsed globule and the intermediate is $30k_BT$, whereas the free-energy difference between the intermediate and the coil is only a few k_BT.

The collapse transition arises from an interplay between density fluctuations of the solvent at small and large length scales. In the coil state, the monomers are well separated and the solvation free energy is dominated by the entropic cost of constraining small-length-scale fluctuations. This cost scales with the size of the excluded volume and, to a good approximation, is given by $\sum_i \Delta\mu_{\mathrm{ex}}(v_i)$. But when the monomers come together, this entropic cost is larger than the energetic cost of forming a vapour bubble that envelopes the monomers. As water is close to phase coexistence, this cost is dominated by the work needed to form an interface, which scales with the area of the excluded volume. Indeed, it is the crossover in the scaling behaviour of the solvation free energy, as shown in figure 1, which is the origin of the collapse transition.

In order to make the above analysis more quantitative, we write the free energy of 'folding' as

$$\Delta G_{\mathrm{fold}} = \Delta G_{\mathrm{solv}} + \Delta G_{\mathrm{intra}}. \tag{21}$$

Here ΔG_{solv} is the reversible work needed to transfer the chain from its 'wet' solvated state to the 'dry' state in the vapour bubble, and $\Delta G_{\mathrm{intra}}$ denotes the free-energy change associated with the internal degrees of freedom of the polymer (see figure 4). In order to estimate ΔG_{solv}, we make the following assumptions:

(1) the fully collapsed globule is confined in a spherical bubble of radius R;
(2) the volume from which the polymer excludes solvent in the extended coil state is equal to the volume of the bubble that contains the globule;
(3) in the extended coil state, the solvent wets the polymer and $n_i = 1$ for all i.

We then arrive at the following expression for ΔG_{solv}:

$$\Delta G_{\mathrm{solv}} = 4\pi R^2\gamma + \tfrac{4}{3}\pi R^3 p - \sum_i \Delta\mu_{\mathrm{ex}}(v_i) \tag{22}$$

$$= 4\pi R^2\gamma + \tfrac{4}{3}\pi R^3(p - c), \tag{23}$$

where γ is the surface tension of the vapour–liquid interface, p is the external pressure, and c is given by equation (15). Let us now compare the resulting values. First, from the measured free-energy difference between the coil and the globule, $\Delta G_{\mathrm{fold}} \approx -30k_BT$, the radius of the bubble, $R \approx 1.1$ nm, and equations (21) and (23), we can obtain $\Delta G_{\mathrm{intra}}$. This yields $\Delta G_{\mathrm{intra}} \approx 73k_BT$. This should be compared with the respective contributions to ΔG_{solv}, $4\pi R^2\gamma \approx 258k_BT$ and $\frac{4\pi R^3}{3}c \approx 361k_BT$, showing that the contribution of the internal degrees of freedom of the polymer to the free energy of collapse is relatively small. Indeed, the free energy of collapse of the hydrophobic polymer is dominated by the solvation free energy. Further, at ambient conditions $p \ll c$, which shows that the collapse transition is dominated by a competition between the energetic cost of creating a vapour–liquid interface in the globule and the entropic cost of maintaining the polymer in the wet state in the coil.

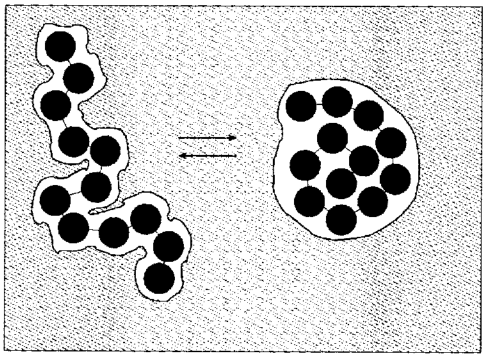

Figure 4. A sketch of the collapse transition of the hydrophobic polymer. The patterned region depicts the solvent, water. The white region indicates the volume from which the polymer excludes solvent.

This simple model also allows us to address the experimental observation that a wide range of proteins denature under pressure. If we assume that (1) the specific volume of folding is determined primarily by ΔG_{solv} and that (2) the dependence of the surface free energy upon pressure can be neglected, then we arrive at

$$\Delta v_{\mathrm{fold}} \approx \frac{\partial \Delta G_{\mathrm{solv}}}{\partial p} \approx V\left(1 - \frac{\partial c}{\partial p}\right). \tag{24}$$

To our knowledge, no values for the dependence of the solvation free energy upon pressure for hard spheres of the size used here have been reported. We therefore make some drastic assumptions. In order to compute $\partial c/\partial p$ we assume that $c \approx k/\kappa_T$, where k is a constant, and κ is the isothermal compressibility. With $\kappa_T \approx 4.6 \times 10^{-10}$ Pa^{-1} at 1 atm and $\kappa_T \approx 3.5 \times 10^{-10}$ Pa^{-1} at 1000 atm, we arrive at $\partial c/\partial p \approx 0.82$. This yields $\Delta v_{\mathrm{fold}} = V(1.0 - 0.82) \approx 600$ cm^3 mol^{-1}. This is in agreement with the experimental observation that proteins denature under pressure, although values reported in the literature for proteins of similar size are significantly smaller, by about a factor of six [35, 36].

7.2. Dynamics of collapse

The free-energy landscape shows that the rate-limiting step is the formation of a vapour bubble of critical size. To investigate the extent to which R_g^2 and U also correctly describe the *dynamics* of this nucleation event, we performed extensive transition-path-sampling simulations on an ensemble of trajectories, each of length 150 ps [18]. A scatter plot of the trajectories is shown in figure 3. To identify the transition state surface, we have identified the configurations on each trajectory from which newly initiated trajectories have equal probability of landing in either the coil or globule [34, 37]. These configurations are members of the transition state ensemble. We project them onto figure 3. First, it is seen that the ensemble is located close to the saddle point in the free-energy landscape. It should also be noted that the transition state ensemble is slightly tilted in the (U, R_g^2) plane. This shows that the larger the polymer, the larger the size of the critical vapour bubble that has to be nucleated. However, the scatter of the transition state ensemble from a line in the (U, R_g^2) plane is notable, which indicates that at least one other variable in addition to U and R_g^2 plays a pertinent role in the collapse transition.

It is also noteworthy that the transition paths do not follow the lowest free-energy path from the metastable intermediate to the fully collapsed globule (figure 3). In particular, the transition states are not centred around the saddle point and the dynamical paths do not follow the steepest-descent path from the saddle point to the fully collapsed globule. The reason is that the polymer and the solvent move on different timescales. The polymer moves on timescales of nanoseconds, while the solvent moves on timescales of picoseconds. As a result, when a vapour bubble is nucleated, the chain cannot respond on the timescale at which the bubble is formed. Only when the bubble has reached a size of around 400 does the polymer fully collapse into the globule state.

The above observation indicates that free-energy surfaces should be interpreted with care. The dynamical pathways for a transition are not always fully determined by the underlying free-energy surface. Figure 3 shows that other aspects can be important. In fact, when we artificially force the polymer to move a hundred times faster with respect to the solvent (i.e. $M_s = 3600$ instead of 36—see section 5), then the transition paths do follow the lowest free-energy path from coil to globule. This clearly demonstrates that the dynamical trajectories are not only determined by the underlying free-energy surface, but that the natural dynamics of the system, or the dynamics of the algorithm, can be important as well. It also means that the transition state ensemble does not need to coincide with the dividing surface in the free-energy landscape, which is the surface that yields the highest transmission coefficient [34].

8. Conclusions

In this paper, we have given an overview of the solvation of hydrophobic solutes. Its character is very different at small and large length scales. At small length scales, solvation is dominated by entropic effects and the solvent still wets the surface of the solute, even when the solute is highly hydrophobic [15, 22]. In contrast, at large length scales, solvation is dominated by enthalpic effects. In this regime, large apolar species can induce a cavitation transition in the solvent. More importantly, in the small-length-scale regime, the solvation free energy scales with the volume of the solute, whereas in the drying regime, the solvation free energy scales with the exposed area of the solute. This behaviour is not only observed for ideal hydrophobic objects, but also for solutes that have weak dispersive interactions with the solvent. The crossover behaviour of the solvation free energy from the wetting regime to the drying regime is important, because it could be of significance to the formation of biological structures. In most biological systems, the size of the hydrophobic species is such that individual species are in the wetting regime, while assemblies of such species are in the drying regime. In the wetting regime, water can only induce a relatively weak attraction between two small apolar species. When several of these species come together, however, water can induce a strong attraction between them.

A clear example of this process is given by the collapse of the hydrophobic polymer. The strong driving force for the collapse of the polymer is provided by the cavitation transition in the solvent. Importantly, the cavitation transition is only induced when a sufficiently large number of monomers come together. This means that implicit solvent models, in which the interaction between the solutes is described by a sum of two body terms, cannot describe the cavitation transition. Equally important, the dynamics of the hydrophobic collapse is dominated by the dynamics of the solvent, especially the formation of a vapour bubble near the surface of a nucleating cluster of hydrophobic monomers. As implicit solvent models cannot capture the dynamics of the solvent, they are of limited use in studying the effect of hydrophobic interactions on the *dynamics* of biological self-assembly.

The connection between hydrophobic collapse and the cavitation/drying transition provides a simple explanation for both cold denaturation and pressure denaturation of proteins. The range and the strength of the interactions between the hydrophobic objects in the drying regime are smaller when the solvent is moved away from phase coexistence. Thus, the lowering of temperature and the increase of pressure destabilize hydrophobic collapse, because both actions move the solvent away from phase coexistence. Finally, we believe that fluorescence resonance energy transfer (FRET) experiments [38] and nuclear magnetic resonance (NMR) spectroscopy [39] should make it possible to observe the drying transition in protein folding experimentally.

Acknowledgments

It is a pleasure to thank David Chandler and Paul Wessels for carefully reading the manuscript. This work was supported in its initial stages by the National Science Foundation (grant nos 9508336 and 0078458) and in its final stages by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy (grant no DE-AC03-76SF00098).

References

[1] Kauzmann A 1959 *Adv. Prot. Chem.* **14** 1
[2] Tanford C 1973 *The Hydrophobic Effect—Formation of Micelles and Biological Membranes* (New York: Wiley–Interscience)
[3] Pratt L R and Chandler D 1977 *J. Chem. Phys.* **67** 3683
[4] Pangali C, Rao M and Berne B J 1979 *J. Chem. Phys.* **71** 2975
[5] Chandler D 1993 *Phys. Rev.* E **48** 2898
[6] Hummer G, Garde S, García A E, Pohorille A and Pratt L R 1996 *Proc. Natl Acad. Sci. USA* **93** 8951
[7] Carambassis A, Jonker L C, Attard P and Ruthland M W 1998 *Phys. Rev. Lett.* **80** 5357
[8] Tyrrell J W G and Attard P 2001 *Phys. Rev. Lett.* **87** 176104
[9] Attard P 1989 *J. Phys. Chem.* **93** 6441
[10] Eriksson J C, Ljunggren S and Claesson P M 1989 *J. Chem. Soc. Faraday Trans.* II **85** 163
[11] Parker J L, Claesson P M and Attard P 1994 *J. Phys. Chem.* **98** 8468
[12] Attard P 2000 *Langmuir* **16** 4455
[13] Evans R 1990 *J. Phys.: Condens. Matter* **2** 8989
[14] Attard P, Ursenbach C P and Patey G N 1992 *Phys. Rev.* A **45** 7621
[15] Lum K, Chandler D and Weeks J D 1999 *J. Phys. Chem.* B **103** 4570
[16] Stillinger F H 1973 *J. Sol. Chem.* **2** 141
[17] Huang D M, Geissler P L and Chandler D 2001 *J. Phys. Chem.* B **105** 6704
[18] ten Wolde P R and Chandler D 2002 *Proc. Natl Acad. Sci. USA* **99** 6539
[19] Huang D M and Chandler D 2000 *Proc. Natl Acad. Sci. USA* **97** 8324
[20] ten Wolde P R, Sun S X and Chandler D 2002 *Phys. Rev.* E **65** 011201
[21] Crooks G E and Chandler D 1997 *Phys. Rev.* E **56** 4217
[22] Huang D M and Chandler D 2000 *Phys. Rev.* E **61** 1501
[23] Hummer G, Garde S, García A E, Paulaitis M E and Pratt L R 1998 *J. Phys. Chem.* B **102** 10469
[24] Weeks J D, Katsov K and Vollmayr K 1998 *Phys. Rev. Lett.* **81** 4400
[25] Eisenberg D and McLachlan A D 1986 *Nature* **319** 199
[26] Vallone B, Miele A E, Vecchini P, Chiancone E and Brunori M 1998 *Proc. Natl Acad. Sci. USA* **95** 6103
[27] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
[28] Frenkel D and Smit B 1996 *Understanding Molecular Simulation, from Algorithms to Applications* (San Diego, CA: Academic)
[29] Löwen H, Madden P A and Hansen J-P 1992 *Phys. Rev. Lett.* **68** 1081
[30] Stillinger F H 1975 *Adv. Chem. Phys.* **31** 1
[31] Landau D P and Binder K 2000 *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge: Cambridge University Press)
[32] Huang D M and Chandler D 2002 *J. Phys. Chem.* B **106** 2047
[33] Guillot B and Guissani Y J 1993 *J. Chem. Phys.* **99** 8075
[34] Bolhuis P, Chandler D, Dellago C and Geissler P L 2002 *Annu. Rev. Phys. Chem.* **59** 291
[35] Vidugiris G J A, Markley J L and Royer C A 1995 *Biochemistry* **34** 4909
[36] Panick G, Vidugiris G J A, Malessa R, Rapp G, Winter R and Royer C A 1999 *Biochemistry* **38** 4157

[37] Northrup S, Pear M R, Lee C-Y, McCammon A and Karplus M 1982 *Proc. Natl Acad. Sci. USA* **79** 4035
[38] Deniz A A, Laurence T A, Beligere G S, Dahan M, Martin A B, Chemla D S, Dawson P E, Schultz P G and Weiss S 2000 *Proc. Natl Acad. Sci. USA* **97** 5719
[39] Klein-Seetharaman J, Oikawa M, Grimshaw S B, Wirmer J, Duchardt E, Ueda T, Imoto T, Smith L J, Dobson C M and Schwalbe H 2002 *Science* **295** 1719
[40] Chandler D 1978 *J. Chem. Phys.* **68** 2959