Dynamics of Protein Domain Coalescence. 2

Gary P. Zientara, Janice A. Nagy, and Jack H. Freed

Baker Laboratory, Department of Chemistry, Cornell University, Ithaca, New York 14853 (Received: July 26, 1981; In Final Form: September 21, 1981)

Dynamic effects of a model pair correlation function and Oseen's tensor hydrodynamic interactions are included in the study of the kinetics of protein domain coalescence using the numerical approach of Zientara, Nagy, and Freed [J. Chem. Phys., 73, 5092 (1980)]. Modifications to the previously reported results due to hydrodynamic drag and either Debye–Hückel or Coulombic electrostatic forces are presented. A model domain pair correlation function is also incorporated to more accurately simulate the spatial and energetic aspects of hydrophobic bonding. Applying this extended model, the variation of coalescence lifetimes with ionic strength and temperature is then calculated and discussed with reference to published experimental data. A frequently observed but anomalous temperature variation in a renaturation rate constant is explained by our results.

I. Introduction

An analysis of the kinetics of protein domain coalescence, an elementary step in the complex protein folding process, has previously been introduced by Karplus and Weaver1 in terms of a diffusion–reaction model. The analytic results of their model for the mean lifetime of uncoalesced domains in the limit of low coalescence probability16 (i.e., reactivity) complement the study of Adam and Delbrück,2 who discussed the mean lifetime of reactive species confined to finite spatial domains in the limit of infinite reactivity. Recently, the first passage time approach of Szabo et al.3,4 has unified previous theories by consideration of the complete range of reactivities. These studies provide an excellent mathematical basis for the study of protein domain coalescence assuming simple domain interactions. The restriction to simple interactions is due to the analytical mathematical difficulties in solving the Smoluchowski equation with a general interaction potential, \( U(r) \).

Model systems of protein dynamics, however, must include protein–solvent, protein–ion, and protein–protein interactions so that model results can be useful in predicting and interpreting experimental data. In order to present a method which allows an unrestricted choice of domain–domain and domain–solvent interactions, Zientara et al.5 (hereafter referred to as I) have employed numerical solutions of the Smoluchowski equation to calculate the mean lifetime of coalesced domains. The model utilized in I is based upon that of Karplus and Weaver5 modified to include the orientation dependence of domain reactivities, interdomain electrostatic forces mediated by a solution with a finite ionic strength, and the effects of domain hydration shells.

In this study we first discuss hydrodynamic effects on the simple Brownian diffusive motion of the protein domains. This modification is included through Oseen's tensor correction to the diffusion tensor and is applied in the study of domains interacting through Debye–Hückel or Coulombic potentials. Also, the hydration shell structure that provides the energy barrier involved in hydrophobic bonding in protein systems is simulated by employing a model domain–domain pair correlation function within the mathematical framework of I. The ionic strength dependence of the coalescence lifetimes is then described by a Debye–Hückel interaction including hydrodynamic effects. Finally, the variation in the rate of coalescence with temperature is examined for different cases of electrostatic interactions.

The discussion of the theoretical and numerical aspects of the calculations is contained in section II. The general effects upon predicted values of mean coalescence lifetimes of hydrodynamic terms and a pair correlation function are presented in section III. In section IV the ionic strength and temperature dependence of coalescence rates are discussed and compared with experiments. A summary of the results of this study appears in section V.

II. Theory

Hydrodynamics. The calculation of the mean lifetime of uncoalesced protein domains has been developed in I through numerical solutions of the Smoluchowski equation

---

1Supported by NIH Grant GM-25862 and NSF Grant CHE-80-24124.

Acknowledgment. This work was supported in part by a grant from the National Science Foundation. After the experimental work had been completed, Professor Stanley D. Furrow of the Berks Campus of The Pennsylvania State University pointed out the Liebhafsky6 observations that made necessary the dashed extension of the curve in Figure 6.

---

(1) (a) M. Karplus and D. L. Weaver, Nature (London), 260, 404 (1976); (b) M. Karplus and D. L. Weaver, Biopolymers, 18, 1421 (1979); (c) D. L. Weaver, Biophys. Chem., 19, 245 (1979); (d) D. L. Weaver, J. Chem. Phys., 72, 3483 (1980).
with a "sink" or reaction term
\[ \frac{\partial P(r,t)}{\partial t} = \nabla_r \cdot D(r) \left[ \nabla_r P(r,t) + \frac{1}{k_B T} P(r,t) \nabla_r U(r) \right] + K(r) P(r,t) \]  

(1)

where \( P(r,t) \) is the probability distribution of domain pairs at separation \( r \) at time \( t \) influenced by a potential \( U(r) \). The spatial dependence of irreversible domain reaction is described through \( K(r) \). Complete details of the theoretical treatment can be found in I. The probability distribution of uncoalesced domains can be calculated using finite difference methods as applied earlier by Zientara and Freed.\(^6\) The spatial integral of the probability distribution can be related directly to a dimensionless mean lifetime, \( \tau^* \), as follows. One first Laplace transforms \( P(r,t) \)

\[ \hat{P}(r,\sigma) = \int_0^\infty e^{-\sigma t'} P(r,t') \, dt' \]  

(2a)

then we obtain

\[ P(\sigma) = \int \hat{P}(r,\sigma) \, dr \]  

(2b)

which then yields

\[ \tau^* = \frac{P(\sigma)}{1 - \sigma P(\sigma)} \]  

(2c)

where \( \tau^* = \tau D/d^2 \) and \( \sigma P(\sigma) \) is the probability the domains have not coalesced irreversibly for given \( \sigma \). The distance of closest approach of the domains is \( d = r_+ + r_- \), with \( r_+ \) and \( r_- \) the domain radii. (The constant \( D \) is discussed below.) For \( \sigma \ll 1 \), eq 2c yields a \( \sigma \)-independent \( \tau^* \) value, consistent with the first passage time, i.e., single mode, approximation.

In the absence of interdomain forces, i.e., \( U(y) = 0 \), the integral formulas of Szabo\(^3\) may be solved exactly for \( \tau^* \), that arises from Oseen's tensor. The method of I enables us to extend this hydrodynamic treatment in the presence of interdomain forces.

Using the general form of the Smoluchowski equation, eq 1, which retains the full properties of the diffusion tensor \( D(r) \), we can include the effects of hydrodynamic drag in our calculation. A hydrodynamic treatment of the motion of solute particles includes the consideration of either a slip, stick, or mixed boundary condition at the surface of an impenetrable sphere considered to be moving in an incompressible fluid which is at rest at infinity.\(^7\) The motion of the solute particles sets up a velocity gradient in the surrounding fluid medium so as to oppose their relative diffusion, and this velocity gradient does depend on the boundary conditions between solute and solvent medium. As in the analysis of diffusion-controlled reactions by Deutch and Felderhof,\(^8\) we will study the hydrodynamic effects on the dynamics of our system by the use of a spatially dependent diffusion tensor

\[ D(r) = DI - 2(k_BT)T(r) \]  

(3a)

where Oseen's tensor is introduced

\[ T(r) = (8\pi \eta r)^{-1} [I + \gamma r(r^{-2})] \]  

(3b)

Here \( D = D_0 + D_\eta \) is the relative diffusion coefficient of unconnected domain centers at infinite separation, which is equal to the sum of their individual Stokes–Einstein values assuming stick boundary conditions at the domain surfaces.\(^5\) \( \eta \) is the solvent viscosity, \( k_B \) is the Boltzmann constant, and \( T \) is the absolute temperature.

Earlier studies\(^6\) of diffusion-controlled reaction dynamics in infinite spatial domains have discussed the adequacy of Oseen's tensor as an approximation to the general Stokes' tensor. Also, others\(^10\) have studied the appropriateness of a slip rather than stick boundary condition in an analysis of macromolecular diffusion and a hydrodynamic description which accounts for the structural nature of protein subunits.\(^11-13\) which are idealized as impenetrable hard spheres in the derivation of eq 3, yet are known to be flexible structures, in some cases penetrable by solvent molecules. After considering these analyses,\(^6\) we have employed Oseen's tensor with stick boundary conditions; however, we recognize that these assumptions can overemphasize the effect of the hydrodynamic interaction\(^10\) by a factor of \( \sim 1-1.5 \).

In I and this work we explain the diffusion–reaction approach to domain coalescence through the motion of two spherical domains which are assumed to be connected by a flexible, noninteracting chain. This simplifying assumption neglects the various dynamical (including excluded volume) and electrostatic contributions to domain motion due to the intervening chain and solvent–chain interactions.

To include chain effects in segmental motion without direct modeling of chain dynamics, one may utilize the approximation of an "internal viscosity"\(^16\) applicable to chain elements. An internal viscosity is employed to describe the motion of a particular chain element independent of all individual intrachain interactions. The concept of an internal viscosity was used recently to qualitatively explain fluorescence decay experimental data by Haas et al.\(^17\). Their experimental observations are also supported by the polymer chain molecular dynamics simulations of Gottleib et al.\(^18\). The internal viscosity, \( \eta_{int} \), is taken to vary as \( \eta_{int} = A + B\eta_{rel} \), where \( \eta_{rel} \) is the solvent viscosity and \( A \) and \( B \) are functions of temperature, ionic strength, and molecular properties of the chain and solvent. The effects of \( \eta_{int} \) have been considered in this study with reference to the temperature dependence of the mean coalescence lifetime, through an ad hoc modification of \( \eta \) appearing in eq 3b.

**Pair Correlation Functions.** The use of Oseen's tensor to describe deviations from simple Brownian diffusion is clearly more appropriate for dynamical modeling in the cases of large separation of domains. The short range
dynamical characteristics of two reacting particles in solution are included employing a reacting pair correlation function which, theoretically, contains statistical information about all solvent–reactant interactions, solvent–solvent interactions, and reactant and solvent molecular structure that will possibly moderate ideal Brownian dynamics. 9,15,19,20

We have shown in I that the pair correlation function is related to an effective interdomain potential of interaction. The extreme importance of solvation effects in hydrophobic bonding21,22 necessitates the incorporation of either an interaction potential or pair correlation function in modeling. In I we chose an approximate radially dependent gaussian form of the domain pair mean potential energy due to solvent–domain interactions. To a first approximation this simulated a single solvent shell or solvent cage barrier to diffusion with facilitated motion upon penetration of the solvent cage (see also ref 23).

Within our approach, we assume spherical hydrodynamic properties of globular domains whose equilibrium properties are sufficiently described through the use of a single radial pair correlation function.24 This assumption will not be valid when considering proteins in solutions containing denaturing solutes at concentrations which may appreciably alter the domain structure, or at very high temperatures. Within the bounds of the assumption of spherically symmetric domain properties, information regarding the solvation of macromolecules can be extracted from published studies25–29 of solutes in water and applied to protein domains existing during intermediate protein folding stages.

In this study we employ a model domain pair correlation function in the Smoluchowski equation. The domain pair correlation function written in terms of \( y \), the dimensionless radial domain separation, is \( g(y) \) (cf. I), which is related to the dimensionless interdomain potential, \( U_m(y) = U_m(r)/kT \), by

\[
\ln g'(y) = \ln g'(y) - \ln U_m(y) \tag{4a}
\]

which can be rewritten as

\[
g(y) = g'(y) \exp[-U_m(y)] \tag{4b}
\]

where \( g'(y) \) represents the effects on the pair correlation function from domain–solvent and solvent–solvent interactions.25,27 and \( U(y) \) is the separable portion of the domain–pair potential (assuming pairwise interactions) used to express the shielded electrostatic interaction between the domains. Since the potential of mean force \( U_m(y) \) enters the Smoluchowski equation,5 we rewrite eq 4 as

\[
U_m(y) = -\ln g'(y) + U(y) \tag{5}
\]

In order to attempt to model the domain pair correlation function, we will concentrate on the empirical form given by the damped oscillatory function

\[
g'(y) \approx 1 + g_0 \cos[\lambda_0(y - 1)] \exp[-\lambda_1(y - 1)] \tag{6a}
\]

with

\[
g'(y) = 0 \quad \text{if } y < 1 \text{ or } y > y_N \tag{6b}
\]

The boundary conditions of eq 6b are a restatement of the domain hard-sphere assumption2 and the extent of the finite spatial domain. Since protein domains are not uniformly hydrophobic, their interactions with the solvent will be nonnegligible. Therefore, \( g'(y) \) does not exactly represent the pair correlation arising solely from repulsive interactions, \( g^{sp}(y) \), as in the theory of Pratt and Chandler,27 yet the approximation is consistent within our approach which considers an orientational average of reactivity or hydrophobicity. Thus, the hydrophobic patches on the surface of a domain that eventually will compose the reactive site in coalescence will necessarily have a solvation character closely resembling the results of the recent computational studies26,27. An alternate description of solvation effects on the coalescence process would be necessary if it were completely dominated by the interactions of ionizable groups.

We can relate the thermodynamic quantities discussed in I to the modeling parameters using eq 5 and 6. First we have the coalescence energy5

\[
U_c = U_m(1) = -\ln [1 + g_0] + U(1) \tag{7a}
\]

and next the magnitude of the activation barrier encountered prior to domain contact

\[
U_{act} = \max[U_m(y)] = U_m(y_N) \tag{7b}
\]

1 \( \leq y \leq y_N \)

\( U_c \) is equal to the difference in the standard molar Gibbs free energies of coalesced and uncoalesced domains through eq 50–52 of I. Explicitly, \( U_{act} \) is the activation energy for the primary hydration shell (located at \( r-d \approx 1.4 \) \( \AA \)) penetration process while \( U_{act} - U_c \) is the activation energy of domain dissociation.

The radial distance from a domain center to the primary solvation shell, \( y_N \), is just the location of the global minimum in \( g(y) \) and is a constraint on the choice of the modeling parameters \( \lambda_0 \) and \( \lambda_1 \). Likewise, calculated spatial correlations for hydrophobic spheres in water significantly beyond a distance equal to \( \approx 2–3 \) times the diameter of a water molecule. This limits the range of suitable choices of \( \lambda_1 \) (with \( \lambda_1^{-1} \) playing the role of a correlation length). Also, the location of the second maximum in \( g(y) \) farthest from \( y = 1 \) is strongly dependent on the choice of \( \lambda_0 \) and \( \lambda_1 \). For the calculations of section III B we have used these constraints to determine the parameters \( g_0, \lambda_0 \), and \( \lambda_1 \). The decay of spatial correlations by \( \approx 1.4 \) \( \AA \) then requires \( \lambda_0 = 15 \). This produces a second maximum in \( g(y) \) at separation \( r-d \approx 3 \) \( \AA \) as desired. We have also studied the case where the minimum of \( g(y) \) occurs at a greater separation, \( r-d \approx 2.5 \) \( \AA \), because of the uncertainty in values to be used when the spatial coordinates of domain structures have been averaged over orientations. Equation 7a was used to set the value of \( g_0 \) based on the coalescence energy studied. The activation energy, \( U_{act} \), is then regarded as a function of \( g_0, \lambda_0 \) and \( \lambda_1 \).
Application to Experimental Observations. The use of a pair correlation function to explicitly include solvent effects in diffusion-reaction models permits the numerical study of both domain association (i.e., coalescence) and dissociation processes. This follows from the application of the standard treatment of kinetics of reactions in liquids to our consideration of domain coalescence. Phenomenologically, the overall observed rate of formation of a folded protein can be described by considering the mechanism

\[ \text{domain}(A) + \text{domain}(B) \xrightarrow{k_1} \text{(AB)domain-pair} = \text{domain}(A) + \text{domain}(B) \xrightarrow{k_2} \text{domain}(A) + \text{domain}(B) \]

where nucleation steps form the spatially separated domains, prior to the coalescence step. Following coalescence, further folding steps add to the structure of a coalesced domain-pair till final formation of the native protein.

For \( k \gg k' \), this simple picture allows us to relate the calculated coalescence lifetime to the kinetic parameters

\[ \tau^{-1} = k_i/(1 + k_i/k) \]

where \( \tau \) has the units of seconds and \( k_i \) and \( k_r \) are the specific rate constants of domain association and dissociation, respectively. Also, \( k = \kappa \Delta D/d^2 \) is the rate constant describing the depletion of domain pairs in contact, so its value is an adjustable parameter in our calculations from which we consider the time dependence of the coalescence step only. Once the physical quantities that characterize the domain-pair system have been specified and inserted into the model, the predicted values for \( k_i \) and \( k_r \) can be determined numerically by two calculations. Using the two limiting cases of eq 8, we obtain

\[ \tau^{-1} \approx k_i \quad \text{for} \quad k \gg 1 \]

and

\[ \tau^{-1} \approx (k_i/k_r)k \quad \text{for} \quad k \ll 1 \]

It is important to note that since few experimental protein studies conclusively distinguish both the kinetic and structural features of folding intermediates, the assignment of kinetic constants to a particular folding step, whether a fast step or slow step, is often based only on qualitative arguments. Thus, assignment of experimentally measured reaction rates to domain association and dissociation is most frequently conjecture. Yet, the mechanism of the rate-limiting step in protein folding remains particularly important in protein studies. There exists no clear evidence that domain coalescence is the event which generally constitutes the slowest step in protein folding. For example, in an investigation of ribonuclease A folding in solvents of varying composition, Tsong and Baldwin did not observe changes in the rate of the slowest step of folding, as would be predicted from a mechanism described by diffusion-controlled domain coalescence. The ion strength and temperature dependence of folding rates from other published experimental works are, however, correlated with the predictions of our model calculations in section IV.

In the limit that folding steps prior to and following coalescence occur on a shorter time scale than coalescence,
TABLE I: Mean Coalescence Lifetimes Including the Effects of a Model Pair Correlation Function and Hydrodynamic Interactiona

<table>
<thead>
<tr>
<th>$U_{act}$</th>
<th>$T_{FP,OS}^{*}$</th>
<th>$T_{FP,OS}^{*}$</th>
<th>$T_{FP,OS}^{*}$</th>
<th>$T_{FP,OS}^{*}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>kcal/mol</td>
<td>kcal/mol</td>
<td>kcal/mol</td>
<td>kcal/mol</td>
<td>kcal/mol</td>
</tr>
<tr>
<td>2.5</td>
<td>25.4 (26.1)b</td>
<td>25.5 (26.2)</td>
<td>24.1 (23.4)</td>
<td>18.6 (22.8)</td>
</tr>
<tr>
<td>2.0</td>
<td>16.2 (17.9)</td>
<td>15.5 (17.3)</td>
<td>14.6 (15.5)</td>
<td>13.1 (14.9)</td>
</tr>
<tr>
<td>1.5</td>
<td>12.0 (12.8)</td>
<td>11.6 (13.1)</td>
<td>10.7 (11.2)</td>
<td>10.0 (10.4)</td>
</tr>
<tr>
<td>1.0</td>
<td>9.4 (9.7)</td>
<td>9.0 (9.1)</td>
<td>8.6 (8.6)</td>
<td>8.2 (7.9)</td>
</tr>
<tr>
<td>0.5</td>
<td>7.8 (7.4)</td>
<td>7.5 (6.9)</td>
<td>7.2 (6.5)</td>
<td>7.0 (6.1)</td>
</tr>
<tr>
<td>0.1</td>
<td>7.0 (5.9)</td>
<td>6.7 (5.5)</td>
<td>6.6 (5.2)</td>
<td>6.5 (5.0)</td>
</tr>
</tbody>
</table>

a No constant input includes $y_1 = 11$, $T = 293$ K, $\kappa = 10^{18}$, $\Delta_1 = 0.01$, $r_n = r_o, N = 800$. The pair correlation function used is given by eq 6 with $(y_1 - 1)d = 1.4 A$ for $d = 10 A$. Modeling parameters lie in the ranges $14.5 < \lambda_n < 18$ and $9.5 < \lambda_o < 38.2$. The value of $T_{FP,OS}$ for each $U_{act}$ can be obtained from eq 10, noting $U(1) = 0$ in this example. An isotropic initial condition with $y_1 = 11$ was used for each calculation. No electrostatic forces were included. For the input given, the corresponding force free result is $T_{FP,OS}^{*} = 7.6 \times 10^3$.

b The results in parentheses were calculated with input identical with that in footnote a except that the primary activation barrier was moved outward, $(y_1 - 1)d = 2.5 A$ for $d = 10 A$. Modeling parameters then ranged from $4.6 < \lambda_n < 11.3$ and $6 < \lambda_o < 12$.

and are presented for the cases of Debye–Hückel and Coulombic interactions. The Coulombic case is, of course, the limit of the Debye–Hückel calculations for zero ionic strength. In the range of electrostatic repulsions between domains $0 < Z_1 Z_2 \lambda_s < 20$ it is observed from Figure 1 that hydrodynamic interactions cause an increase by a factor of $\sim 0.5$ in the coalescence lifetimes compared to those predicted in I. Domains experiencing attractive electrostatic forces are found to be generally unaffected by hydrodynamic effects. This corresponds to the results for solute attraction obtained in studies of diffusion in infinite spatial regions.8,15 The results in Figure 1 may be regarded as an upper bound for hydrodynamic effects since we have used stick boundary conditions (see previous section), while the results for $T_{FP,OS}^{*}$ give the lower bound. The ratios of Figure 1 can be used to obtain numerical values of $T_{FP,OS}^{*}$ by reference to Figure 5 of I.

Deutsch and Felderhof6 observed that hydrodynamic effects become most significant in the case of equal reactant size, causing a decrease in the reaction rate by a factor of $\sim 1/2$ depending on reactant electrostatic charge. Therefore, in this study we have restricted our calculations to the case $r_n = r_o = d/2$, where $r_n$ and $r_o$ are the domain pair radii. The importance of hydrodynamic terms in diffusive dynamics including Debye–Hückel electrostatic reactant interactions has also been studied numerically by Pedersen and Fred6 and Hwang,15 who found similar decreases in predicted reaction rates.

Our numerical studies show the insensitivity of $T_{FP,OS}^{*}$ to hydrodynamic interactions in the slow reaction limit that may also be seen from Szabo’s first-passage time approach.5 In the diffusion-controlled limit Szabo’s formulas assert the importance of hydrodynamic corrections and interdomain interactions, since $T_{FP,OS}^{*}$ is highly dependent on the reactant dynamics, also in agreement with our results.

B. Model Domain Pair Correlation Function. Table I contains results for coalescence lifetimes obtained when a model pair correlation function and hydrodynamic effects are included. The values in Table I were obtained for the $k \to \infty$ limit where $T_{FP,OS}^{*} = k_{FP,OS}$ (cf. eq 8 and 9a).

As expected, the mean coalescence lifetimes reflect the primary activation barrier height, but the rate of coalescence assumes a partial diffusion-controlled character. In the

Figure 2. $g'(y)$, the model pair correlation function vs. interdomain separation, $y$. The solid curve corresponds to a case in which $U_c = -1$ kcal/mol and $U_{act} = 1$ kcal/mol, modeled with input $g_d = 4.6, \lambda_n = 10$ and $\lambda_o = 6$. The broken curve corresponds to a case in which $U_c = -0.5$ kcal/mol and $U_{act} = 0.2$ kcal/mol, modeled with input $g_d = 1.4, \lambda_n = 10$, and $\lambda_o = 8$. The location of the primary activation barrier is $y_{FP,OS} = 2.5 A$ as shown, if $d = 10 A$.

Figure 3. $U_{act}(y)$, the domain–domain potential of mean force vs. interdomain separation, $y$. Both the solid and broken curves refer to the respective curves of $g'(y)$ seen in Figure 2: (solid curve) $U_c = -1$ kcal/mol and $U_{act} = 0.2$ kcal/mol; (broken curve) $U_c = -0.5$ kcal/mol and $U_{act} = 0.2$ kcal/mol.

range of $U_{act}$ between 1 and 2 kcal/mol, $T_{FP,OS}^{*}$ varies approximately linearly with $U_{act}$ for constant $U_c$. In contrast, $T_{FP,OS}^{*}$ is slowly varying as $U_c$ is progressively made large and negative, indicating that the effective domain attraction experienced after penetration of the hydration shell does not significantly increase the forward rate of coalescence, although it is important.

Figure 2 shows the form of $g'(y)$ used in this study, with the corresponding potential of mean force, $U_{act}(y)$, illustrated in Figure 3. Recall that the pair correlation function includes the effects of solvent-solvent and solvent–domain interactions. The thermodynamically relevant quantities from eq 7, $U_c$ and $U_{act}$, therefore are measures of the solvent structure surrounding a hydrophobic solute or, here, a domain with an averaged hydrophobic character. Values for the thermodynamic quantities describing the formation of hydrophobic bonds between amino acid side chains have been given by Nemethy and Scheraga21 and others.21 Values of molar $\Delta G^\circ$ for amino acids reported for 25 °C range from ~0.3 to ~1.6 kcal. The total hydrophobic bonding energy of domain coalescence, $U_{act}$, is then a sum of the contributions of side chains participating in bonding. The value of $U_{act}$ includes the effects of possible hydrogen bonding between protein segments and water and between individual water molecules in the surrounding hydration shell. Restructuring of the hydrogen bonds that exist in the solvent cage sur-
ranging the domains can produce a marked variation in \( U_{act} \). Though the rate of domain coalescence and the rate of dissociation are functions of \( U_{act} \), the rate of dissociation exhibits the stronger dependence with \( \mu \). An input includes the magnitude of the activation barrier, where \( \Delta \mu_{act} = \exp(-U_{act} - U_0) \). (Note: \( U_{act} \) and \( U_0 \) are dimensionless here.)

Experimental studies of protein denaturation commonly take advantage of solvent structure perturbation techniques. An important solvent effect of a denaturant is the disruption or reformation of the hydrogen bonding networks in water, which influences the clathrate structure around protein side chains. Denaturants generally reduce the magnitude of the activation energy governing renaturation (i.e., \( U_{act} \)). Additional solution components that hydrogen bond with protein segments can also sterically interfere with folding events, a different mechanism than solvent structure perturbation.

The sensitivity of the rate constant of coalescence to the hydration shell parameters is shown in the results of Table I, through the change in \( \tau_{FF,On} \) as the magnitudes of both \( U_{act} \) and \( U_0 \) vary. Though the dependence of \( \tau_{FF,On} \) on \( U_{act} \) and \( U_0 \) differs, a net decrease in \( \tau_{FF,On} \) is observed as the magnitudes of both \( U_0 \) and \( U_{act} \) decrease. In cases of \( |U_0| > U_{act} \) and \( U_{act} \leq 1 \text{kcal/mol} \), the predicted lifetime is actually shorter than the result in the absence of any activation barrier, where \( U_{act} = U_0 = 0 \) (cf. footnote a, Table I). The results of Table I demonstrate that \( \tau_{FF,On} \) is dependent on the energetics of coalescence such that 0.5–1.0 kcal/mol variations in the thermodynamic quantities associated with hydrophobic bonding, \( U_0 \) and \( U_{act} \), would produce experimentally observable variations in the time dependence of coalescence. From the data of ref 21 and 22 we note that a variation in \( U_0 \) of 0.5–1.0 kcal/mol could be attributed to the formation or lack of formation of as few as \( \sim 1–3 \) hydrophobic bonds between amino acid side chains.

**IV. Comparison with Experimental Results**

**A. Dependence of \( k_\ell \) on Ionic Strength.** In Figure 4 we exhibit the ionic strength (i.e., \( \mu \)) dependence of \( \tau_{FF,On} \) for several cases of charged domains including both electrostatic and hydrodynamic interactions. We have used \( \tau_{On} \) to represent the coalescence lifetime for uncharged domains including Oseen’s tensor hydrodynamic interactions. Computed in the limit \( \kappa \to \infty \), our calculations yield values of \( \tau_{FF,On} = k_\ell^{-1} \). The results for \( \mu \geq 0.01 \) reveal an ionic strength variation similar to that of freely diffusing charged reactants in solution, i.e., \( \ln(k_\ell/k_\ell^0) = Z_IZ_Z\mu/(1 + \kappa^{2/3}) \), with the inverse of the Debye length \( \kappa_D \approx \mu^{-1/2} \). All results in Figures 1 and 4 were obtained within the range of validity of Debye–Hückel ionic strength, i.e., \( \kappa_D \approx 3 \). In our calculations we assumed that the dielectric constant of water is independent of the ionic strength. This is known to be a good approximation even for the maximum mole fraction of ions considered \( \kappa_D \approx 0.02 \) which corresponds to \( \mu = 0.4 \text{ M} \) to \( \mu = 0.15 \text{ M} \).

For charged domains with \( |Z_IZ_Z| > 1 \), Figure 4 shows a strong ionic strength dependence of \( \tau_{FF,On} \), displaying approximately an order of magnitude variation as \( \mu \) decreases from 0.15 M to 0.001 M. A weaker ionic strength dependence is obtained for domains where the magnitude of the net charge product is \( \leq 10 \).

Strongly acidic (or basic) experimental conditions will generally protonate (or deprotonate) all solvent-exposed ionizable protein groups. In both these pH limits domains will typically carry like charges and \( Z_IZ_Z \) will be positive. Predictions for the ionic strength dependence of the coalescence rate constant, \( \k_\ell \), are shown in Figure 5 for several values of \( Z_IZ_Z \). The variation in \( k_\ell(k_\mu) \) is shown relative to the low ion strength \((\mu = 10^{-4} \text{ M}) \) limiting value. The predicted rate of coalescence increases significantly, for fixed \( Z_IZ_Z \), reflecting the effect of ionic shielding of the repelling domains. Temperature-jump data of trypsin folding from Pohl recorded under strongly acid conditions are compared to the predictions in Figure 5. However, the trypsin results of Pohl refer to a single rate-limiting folding step whose mechanism has not yet been determined experimentally.

A complication in the analysis of experimental data is encountered when attempting to use the results shown in Figure 4 for predictive purposes due to the variation of protein charge with the solvent ionic strength. Hammes

---

and Alberty\textsuperscript{41} have discussed the use of an average protein charge to simplify this problem, recognizing that the average value represents an approximation appropriate over a small range of both pH and ionic strength. The results of our calculations can be used to analyze experimental observations if the protein charge (i.e., value of $Z_bZ_b$) is changed as $\mu$ is changed. We first note that each curve in Figure 4 was calculated for a specific value of $Z_bZ_b$. As $\mu$ is increased, the mean coalescence time for domains with fixed charge will increase (for $Z_bZ_b < 0$) or decrease (for $Z_bZ_b > 0$), to the limit $\tau_{FP,\text{O}}(\mu \gg 1 \text{ M}) = \tau_{OP,\text{O}}$. The opposite variation of $\tau_{FP,\text{O}}$ with $\mu$ can be predicted, however, in cases where the magnitude of $Z_bZ_b$ increases with $\mu$. An example of this effect is demonstrated by comparing the values of $\tau_{FP,\text{O}}(Z_bZ_b = 4, \mu = 0.1 \text{ M})$ and $\tau_{FP,\text{O}}(Z_bZ_b = 16, \mu = 0.4 \text{ M})$ from Figure 4. Therefore, it is possible to predict a constancy in $\tau$ or $k_{\tau}$ with ionic strength under certain conditions which, however, are in the regime where the time dependence of domain coalescence is strongly affected by electrostatic interactions. Unfortunately, a folding step proceeding through an entirely different mechanism can also display insensitivity to ionic strength variation.

Several processes can account for the variation in domain charges as $\mu$ is increased: (1) the binding of ions by a protein\textsuperscript{40} and (2) intradomain conformational changes which expose or "bury" charged groups, or allow side chains to orient forming salt bridges. As a further complication, a variation in $\tau$ with $\mu$ can also be produced by the effects of ions in solution on chain segment motion. The moderation of intrachain electrostatic interactions due to ionic shielding can: (1) increase the domains' available diffusive space, which alone would tend to lengthen the mean coalescence lifetime (cf. eq 45 of I, where $\tau_{\text{FP}} = \gamma Z_b^2$), and (2) change the chain segment diffusivity described through the ionic strength dependence of an apparent internal viscosity.

Because of these analytical difficulties, we conclude that the information obtained in ionic strength dependent measurements of protein folding times need not always disclose the diffusive nature of folding steps nor alone reveal unambiguously the electrostatic effects on folding.

**B. Dependence of $k_\tau$ on Temperature. Predictions.**

The temperature dependence of the mean coalescence lifetime $\tau (= \tau_{\text{FP}}^2/D)$ enters implicitly through the diffusion coefficient (via $\eta_{\text{FP}} = A + B\eta_{\text{FP}}$), the dielectric constant of water $\epsilon(T)$, and the pair correlation function $g(y)$. The dependence of $\eta_{\text{FP}}$\textsuperscript{42} and $\epsilon$\textsuperscript{43} on temperature has been taken from published formulas based on experimental data. At present, the temperature variation of $A(T)$ and $B(T)$ is not adequately understood,\textsuperscript{16} complicating any attempt to describe $\eta_{\text{FP}}(T)$ rigorously. Similarly, an empirical description of the change of $g(y)$ is lacking. However, the graphical results of Pratt and Chandler\textsuperscript{22} for free nonpolar particles in water reveal a variation in $g(y)$ as temperature increases from 4°C to 50°C. This change corresponds to an increase in the primary activation barrier at high temperatures as water takes on the characteristics of a hard-sphere solvent. A much weaker increase in the magnitude of $U_*; is noted over this temperature range.

\(\text{Figure 6. Plots of } \ln(1/\tau), \text{ the logarithm of the coalescence rate constant, vs. inverse temperature for } Z_bZ_b = -16, 0, 16, \text{ as shown. Results correspond to } (a) U_* = -1.7kT, U_{\text{act}} = 3.4kT, \gamma y = 1.26; (curve b) U_* = -1.7kT, U_{\text{act}} = 1.7kT, \gamma y = 1.25; \text{ and } (curve c) \text{ no barrier function employed. Other modeling input includes } N = 800, \gamma y = 11, r_b = 10^5, \Delta = 0.01, \text{ and an isotropic initial condition. Included are hydrodynamic effects and Debye-Hückel electrostatic interactions assuming } \mu = 0.15 \text{ M, } r_b = r_b, \text{ and } d = 10 \AA.\)

In Figure 6 we show the predicted Arrhenius plots of $\ln(1/\tau)$ or $\ln(k_\tau)$ for a negative, zero, and positive value of $Z_bZ_b$. In order to obtain these results it was necessary to make the simplifying assumptions for the reasons noted above: (1) $\eta_{\text{FP}} \approx \eta_{\text{O}},$ and (2) $U_{\text{FP}}(r)/kT = f(r)$. Curves a, b, and c represent cases modeled with activation energy barriers of (a) $3.4kT$, (b) $1.7kT$, and (c) no barrier. These results then form the bounds on results that would be obtained employing $U_{\text{FP}}(r)/kT = f(r, T)$. This more realistic use of a temperature dependent $U_{\text{FP}}(r)/kT$ is discussed below. These approximate results, however, simplify the discussion of the effects of bulk solvent properties on coalescence.

From Figure 6 we observe that the various fixed $Z_bZ_b$ cases predict approximately an inverse temperature dependence of the log of the forward coalescence rate constant, common to an Arrhenius description. The slope of a linear fit of the results in Figure 6 yields an apparent activation energy of $\sim 4.5$ kcal/mol in all cases of $Z_bZ_b$, although no activation barrier of this magnitude was included in the model, as is evidenced most clearly from the results obtained without use of a pair correlation function (curves c in Figure 6). The temperature dependence of the predictions therefore is a measure of the viscous and energetic aspects of the diffusive dynamics governed particularly by solvent properties. In our computations we have concentrated on the range of temperature from 20 to 70°C, common to thermal denaturation-renaturation experiments. As the temperature is increased from 20 to 70°C the dielectric constant of water decreases\textsuperscript{43} ~20%, and the viscosity of water decreases\textsuperscript{42} ~60%. The change in the dielectric properties of water is effective in slightly increasing the electrostatic force of repulsion between domains with like charges. In I (cf. Figure 5 of I) computed predictions for $\tau_{\text{FP}}$ or $k_\tau$ revealed a significant dependence on electrostatic forces, particularly for charged domains where $Z_bZ_b > 4$. From the results in I for $Z_bZ_b > 4$ we conclude that interdomain repulsion is able to measurably reduce $k_\tau$ compared to the cases of uncharged or oppositely charge domains. Hydrodynamic interactions also enhance the dependence of $k_\tau$ on electrostatic interactions for cases of positive $Z_bZ_b$, as seen in Figure 1.

The corrections due to the change in the primary activation energy barrier of the mean domain interaction at higher temperatures\textsuperscript{27} can be determined qualitatively from Figure 6. If, for example, the primary activation barrier energy approximately doubled as the temperature
increased from 20 °C, then the predictions of ln $k_f$ would be best approximated by the results in curve b at $T \approx 20$ °C and the results in curve a at higher temperatures. This qualitative argument demonstrates two different effects: (1) for $Z_a Z_b$ negative, i.e., oppositely charged domains, the temperature dependent results would be relatively insensitive to short range solvent cage factors; (2) for domains of zero or like charges, $Z_a Z_b \approx 0$, changes in the solvent cage characteristics would emphasize the decrease in $k_f$ at higher temperatures.

Figure 7 shows the relative variation in $k_f$ when we explicitly include the change in the primary activation energy barrier of the mean domain interaction at temperatures above 4 °C. The increase in the barrier energy is based upon the results of Pratt and Chandler\(^{(27)}\) for nonpolar (spherical) solute pair correlation functions which vary due to the tendency of water to assume hard-sphere solvent qualities with increasing temperature. Hard-sphere molecules can be more closely packed about the surface of a nonpolar solute compared to the packing of water in a hydrogen bonded clathrate structure. Thus, the exact magnitude and description of the variation of $U_{\text{set}}$ is a function of solute (i.e., domain) composition and dimensions. To calculate the results in Figure 7 we have chosen a linear temperature variation in $U_{\text{set}}$, from 1.7$kT$ at 20 °C to 3.4$kT$ at 70 °C. The weak variation and magnitude of $U_{\text{set}}$ are used only to approximately include the changing solvent properties of water. From Figure 7 we observe the temperature-dependent competition between the effects of the reduced viscous forces, and the electrostatic forces and solvation shell barrier variation. This competition is most manifest in cases of large, positive $Z_a Z_b$, as expected.

An interpretation of the Figure 7 results found using $Z_a Z_b \approx -4$ yields a temperature-dependent apparent activation energy, if on straightforwardly applies an Arrhenius rate constant analysis. Calculated apparent activation energies obtained from these particular results by linearization over small temperature increments range from about +2 kcal/mol ($Z_a Z_b = 16$, 20 °C ≤ $T$ ≤ 45 °C) to about −3.5 kcal/mol ($Z_a Z_b = 16$, 60 °C ≤ $T$ ≤ 70 °C). The predictions of this section therefore reveal clearly that an Arrhenius analysis is inappropriate when applied to the results of such a diffusive process as we have modeled.

The results shown in Figure 7 found using $Z_a Z_b \approx -4$ and $\mu = 0.15$ M indicate a temperature at which there is predicted a maximum coalescence rate within the range 20 °C ≤ $T$ ≤ 70 °C. Similarly, this maximum rate is predicted for slightly higher values of $Z_a Z_b$ when $\mu$ increases, since ionic shielding has diminished domain interactions compared to the example of $\mu = 0.15$ M. For large values of $Z_a Z_b$ and $\mu \rightarrow 0$, one approaches the limit where $\beta \ln k_f / \beta (1/T)$ is always negative for 20 °C ≤ $T$ ≤ 70 °C.

Corrections to the results found in Figures 6 and 7 are necessary due to the internal viscosity approximation utilized, $\eta_{\text{int}} = \eta_{H,O}$. The form $\eta_{\text{int}} = C \exp[\mu / \theta T] + B \eta_{H,O}$ as suggested by Peterlin\(^{(46)}\) was employed in the recalculation of the results in Figure 7. Within this approximation of $\eta_{\text{int}}$, $C$, $B$, and $\theta$ are constants, independent of temperature and calculated from chain properties. A uniform and temperature-independent reduction in $k_f$ of an order of magnitude resulted from these computations. This result is consistent with the experimental data for a segmental diffusion coefficient obtained by Haas et al.\(^{(17)}\)

Experimental Comparison. The predicted temperature dependence of the rate of coalescence can be easily compared to temperature-jump experiment results and results from several other techniques. The temperature-jump technique is most suitable, since it does not depend on the use of additional chemical components to perturb the equilibrium between the folded and unfolded protein states. Therefore, it is a useful kinetic probe of a reaction process such as domain coalescence, which is expected to proceed at a rate that is highly sensitive to many solvent and protein interactions. Due to the complex multistep kinetics of the entire protein folding process, experiments may resolve information about a single kinetically dominant folding step\(^{(44,45)}\) or several folding steps.\(^{(42,50)}\)

The interpretation of these data is nontrivial when one attempts to recreate a folding scenario and to apply findings to a specific folding event. To apply such data to domain association is, then, supported only by qualitative arguments and comparisons with model predictions.

The results seen in Figure 7 for $Z_a Z_b \approx -4$ indeed demonstrate the qualitative behavior observed in many protein folding studies.\(^{(44-49)}\) The experimental data referred to describe the temperature dependence of a rate constant, $k_f^{0,0}$, of a single folding step. The predicted temperature dependence of $k_f$ (the domain coalescence rate constant) for $Z_a Z_b \approx -4$ reveals a marked deviation from an Arrhenius description with a temperature dependent maximum in $k_f$ obtained for intermediate values of $Z_a Z_b$ and a change in the sign of the slope $\beta \ln k_f / \beta (1/T)$ at large values of $Z_a Z_b$. Examples of proteins which exhibit one of these types of temperature dependence are (temperature-jump results unless noted): chymotrypsinogen A\(^{(44a)}\) and B,\(^{(39)}\) chymotrypsinogen,\(^{(460)}\) $\alpha$-chymotrypsin,\(^{(466)}\) trypsin,\(^{(36)}\) elastase,\(^{(44}\) ribonuclease A,\(^{(44,45)}\) $\alpha$-lactalbumin,\(^{(68)}\) lysozyme,\(^{(47)}\) deoxyribonuclease (difference spectroscopy),\(^{(48)}\) and ferricytochrome c.\(^{(49)}\) This observed temperature depen-
The properties of the domains have been utilized. We have described the diffusive model of protein domain coalescence which forms the basis of the protein folding event that is the focus of the present study.

A reduced variable model of the dynamics and interactions between protein domains and solvent components must, for practical reasons, include numerous approximations and be based upon many assumptions. Our approach assumes that folding events such as nucleation and domain formation have occurred prior to coalescence and that other folding events can take place after coalescence to stabilize a coalesced domain pair and conclude the formation of a native structure of the protein. Therefore, coalescence (i.e., domain association) has been depicted as the elementary reaction step in the multistep formation of correct or abortive intermediates in a multiple-domain protein.

In the calculations of this work, spatial and time averages of the discrete conformational and electrostatic interactions and changes in liquid water structure on the one hand and the combination of increased electrostatic interactions and changes in liquid water structure on the other.

The Smoluchowski equation solution of domain diffusion and association in has been extended in this work to include Smoluchowski's equation for the treatment of particles immersed in a continuum fluid. However, as a departure from this level of solvent modeling we have employed a pair correlation function with molecular scale detail to specifically include the domain solvation characteristics that are extremely important in describing hydrophobic bond formation. Hydrodynamic effects were observed to alter model predictions by increasing $k_{FP}$ for domains experiencing electrostatic repulsion. This case of positive domain charge product, that is, $Z_{a}Z_{b} > 0$, is most applicable to protein systems since ionizable groups contained in both protein domains will be similarly affected by pH and ionic strength conditions.

The sensitivity of $k_{h}$ to the solvent properties of water was examined by the use of a pair correlation function from which a mean interdomain potential can be derived. Variations of the order of $\sim 1$ kcal/mol in the energy of activation or coalescence energy were found to predict significant changes in $\tau$.

The effects of the model parameters on $\tau$ have been discussed in detail. The predictions of the complete model were compared with experiments. The ionic strength dependence of $\tau$ was discussed to illustrate that changes in protein domain charge with ionic strength can lead to apparently anomalous behavior. A simple analysis was presented to show that electrostatic effects on $\tau$ must be considered as a function of pH, $\mu$, and $Z_{a}Z_{b}$. Although this conclusion follows readily from Debye–Hückel theory, it is emphasized that in the case of coalescence, a single $\mu$-dependent folding experiment is ambiguous on the details of electrostatic effects.

Finally, the temperature dependence of the coalescence rate was discussed. The rate constant $k_{i}$ in cases of $Z_{a}Z_{b} \geq -4$ was predicted to exhibit a temperature dependence deviating greatly from an Arrhenius description. Within a certain range of $Z_{a}Z_{b}$ (related to ionic strength conditions), a temperature dependent maximum in $k_{i}$ is predicted. This characteristic feature in $k_{i}$ has been observed in folding experiments on many proteins and has been regarded as an anomaly. Our analysis has clearly and naturally shown that this result is an effect of the temperature-dependent competition between the variation in the domains' diffusivity (through $\eta_{im}$ and $\eta_{im}$) on the one hand and the combination of increased electrostatic interaction and changes in liquid water structure on the other.

It is clear that the stochastic modeling of domain coalescence requires a number of physical factors. That is, the domain diffusion–association process is not easily explained either in the simple context of free particle diffusion-controlled kinetics (cf. the basis of ref 31) nor just in chain end diffusion of polymers. Rather, domain coalescence, as our model reveals, relies on the structural properties of liquid water, the presence of ions in solution, protein chain segment motion, electrostatic domain–domain interactions, and domain hydrophobic character. These quantities also exhibit a temperature and composition variation. Most important, our predictions emphasize the cooperativity or competition between these factors in regulating the rate of coalescence. Although the coupling of these factors poses a difficult problem for theoretical analysis, their consideration has been shown to be critical in accurate modeling of domain coalescence. Ultimately, it is hoped that a careful analysis of this single event will aid in understanding the dynamics of the entire folding process.

Acknowledgment. J. A. N. wishes to acknowledge the guidance and support provided by Professor H. A. Scheraga.

References