INSTITUTE OF PHYSICS PUBLISHING PHYSICAL BIOLOGY

Phys. Biol. **2** (2005) R1–R25 doi:10.1088/1478-3975/2/3/R01

REVIEW ARTICLE

How do biomolecular systems speed up and regulate rates?

Huan-Xiang Zhou

Department of Physics and Institute of Molecular Biophysics and School of Computational Science, Florida State University, Tallahassee, FL 32306, USA

E-mail: zhou@sb.fsu.edu

Received 25 April 2005
Accepted for publication 2 August 2005
Published 24 August 2005
Online at stacks.iop.org/PhysBio/2/R1

Abstract

The viability of a biological system depends upon careful regulation of the rates of various processes. These rates have limits imposed by intrinsic chemical or physical steps (e.g., diffusion). These limits can be expanded by interactions and dynamics of the biomolecules. For example, (a) a chemical reaction is catalyzed when its transition state is preferentially bound to an enzyme; (b) the folding of a protein molecule is speeded up by specific interactions within the transition-state ensemble and may be assisted by molecular chaperones; (c) the rate of specific binding of a protein molecule to a cellular target can be enhanced by mechanisms such as long-range electrostatic interactions, nonspecific binding and folding upon binding; (d) directional movement of motor proteins is generated by capturing favorable Brownian motion through intermolecular binding energy; and (e) conduction and selectivity of ions through membrane channels are controlled by interactions and the dynamics of channel proteins. Simple physical models are presented here to illustrate these processes and provide a unifying framework for understanding speed attainment and regulation in biomolecular systems.

1. Introduction

A biological organism is a dynamic system consisting of many interacting molecules. Typically, biological functions are associated with specific molecular structures, such as those of a folded protein or a bound protein complex. It is, therefore, not surprising that great significance has been attached to the thermodynamic stabilities of folding and binding. However, biomolecules are never at thermodynamic equilibrium nor isolated. A biomolecule is constantly faced with competing pathways. For example, a protein undergoing the folding–unfolding transition may also aggregate. An enzyme may be bound by an inhibitor before it has a chance to bind to its substrate. The dynamic nature of biological systems is at the core of systems biology, which typically involves solving rate equations for networked chemical reactions. The focus of this review is biochemical reaction rates and other 'speeds' that measure changes in biological processes.

Plenty of examples illustrate the critical roles of speed in biological functions. For example, *Bacillus amyloliquefaciens* produces an extracellular ribonuclease, barnase, as a weapon against predators or competitors. However, the first target that barnase would find is the host cell's own RNA. The potentially lethal RNase activity is safeguarded by the intracellular inhibitor barstar rapidly forming a tight complex with the enzyme. Indeed, host cells expressing barnase cannot survive without co-expressing barstar [1]. The purple cone, *Conus purpurascens*, and other venomous animals capture prey with remarkable efficiency and speed through releasing toxins that rapidly bind to ion channels [2]. Rapid binding in these cases is afforded by electrostatic attraction. In other processes, the mechanisms of obtaining speed are more complex. These include directional movement of motor proteins and conduction of ions through membrane channels.

Here we will present simple physical models to illustrate the attainment and regulation of speed in biological processes, with a view toward unifying themes among the diverse

processes. The scope of the review will be limited to theories, and will not cover computational studies such as quantum calculations on enzyme catalysis, molecular dynamics simulations of protein folding, motor proteins, or ion channels, or Brownian dynamics simulations of protein binding.

2. Simple models for rates and fluxes

2.1. Unimolecular reactions

Conformational transitions as simple as the trans-to-gauche bond rotation of ethane and as complicated as protein folding share some common features. First, the transitions occur within a single molecule. Second, the molecules have well-defined states, corresponding to minima of the energy or free energy surface, separated by energy or free energy barriers. Often the transitions are modeled by a rate equation for the molecule concentration C:

$$\mathrm{d}C/\mathrm{d}t = -kC$$

which leads to an exponential decay of the concentration: $C(t) = C(0)\exp(-kt)$. To facilitate the calculation of the rate constant k, let us imagine the following bulk experiment. Molecules are prepared at the starting (i.e., reactant) state with a concentration C. Whenever a transition is made to the product state, the molecule is immediately removed and the reactant pool is replenished so that the reactant concentration is always maintained at C. In this experiment, after a brief transient period, a steady state will set in. The steady-state flux will be

$$J = kC. \tag{1a}$$

Re-arranging, we find

$$k = \frac{J}{C}. \tag{1b}$$

The rate constant is thus given by the ratio of the steady-state flux and the reactant concentration. The initial rate constant, observed when the system has just been prepared, is what is commonly referred to as the transition-state theory (TST) approximation, and is given by

$$k_{\mathrm{TST}} = \text{equilibrium probability at transition state} \times \text{rate of crossing transition state}. \tag{1c}$$

k_{TST} sets an upper limit for k. In general, the energy barrier, $\Delta G^{\ddagger}$, measured from the reactant state to the transition state, is a major determinant of the rate constant, and the following expression can be written:

$$k = k_0\,\mathrm{e}^{-\beta\Delta G^{\ddagger}} \tag{2}$$

where $\beta = (k_{\mathrm{B}}T)^{-1}$ with k_{B} Boltzmman's constant and T the absolute temperature, and the prefactor k_0 is dictated by the dynamics of the molecule undergoing the conformational transition. For easy reference, a list of frequently used symbols and their meanings are compiled in table 1.

To illustrate the calculation of k, let us consider a molecule modeled by a one-dimensional energy surface $G(q)$, with reactant and product states around energy minima at q_{rct} and q_{prd}, respectively, and an energy barrier at $q^{\ddagger}$ (figure 1(*A*)).

Table 1. List of frequently used symbols.

Symbols	Explanation
C, c	Concentration
$D, \mathcal{D}$	Diffusion constant or tensor
e	Charge on ion
f	Equilibrium fraction
F, A	Dimensionless factor
$\mathcal{F}$	Force
$g(x)$	Gating function
G	Free energy
h	Half height of reactive patch on DNA
$\mathcal{H}$	Step function
J, j	Flux
k	Rate constant
$k_{\mathrm{B}}T, \beta$	Product of Boltzmann factor and temperature, and inverse
K	Equilibrium constant
l	Sequence separation
L	Length of ion channel
$\mathcal{L}$	Smoluchowski operator
m, M	Index and total number of microstates
n_{f}, s	Numbers of folded and unfolded residues
N	Total number of residues in a protein
$\mathcal{N}$	Normalization constant
O	Contact order
P, p	Probability
$P_i(x)$	Legendre polynomials
q	Reaction coordinate
$\mathbf{q}$	Translational and rotational coordinates
Q	Partition function
r, a, d, b	Distance
$\mathbf{r}$	Position vector
$R(t)$	Relaxation function
$\mathcal{R}$	Random number
S	Entropy
t, τ	Time
U, ε	Energy
v	Velocity
V	Voltage across membrane
w	Conversion or reaction rate
x	Coordinate along track of motor protein or along channel pore
y	Yield of folded protein
z	Coordinate along DNA axis
γ, ψ	Capture probability, and complement
λ	Eigenvalue
ν	Weight of unfolded conformation
θ, α	Polar angle
ρ	Distance to DNA axis
σ	Cross section of ion channel
ω	Conversion rate; oscillation frequency
Ω	Degeneracy (discrete version of density of states)
ζ	Solvent friction

Suppose that the motion of the molecule along the reaction coordinate q is diffusive, with a diffusion coefficient D (which may depend on q); then the probability density of the molecule is governed by the Smoluchowki equation [3]

$$\frac{\partial P(q,t)}{\partial t} = \mathcal{L}P(q,t) = \frac{\partial}{\partial q}\left[D\,\mathrm{e}^{-\beta G(q)}\frac{\partial}{\partial q}\,\mathrm{e}^{\beta G(q)}P(q,t)\right]. \tag{3}$$

This generalizes the familiar free diffusion equation by accounting for the energy function $G(q)$, which leads to a biasing force,

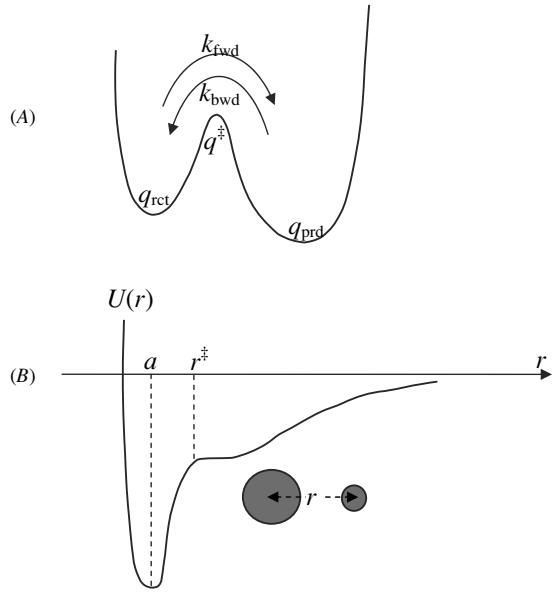

Figure 1. Models for (*A*) unimolecular reaction and (*B*) bimolecular reaction.

$$\mathcal{F} = -\frac{\mathrm{d}G(q)}{\mathrm{d}q}$$

on the molecule. As a result the molecule will have a non-vanishing average velocity, proportional to the bias force,

$$v = \frac{\mathcal{F}}{\zeta}$$

where ζ is the friction coefficient. According to the Einstein relation,

$$D = \frac{k_B T}{\zeta}. \tag{4}$$

The total flux of the molecule now consists of the usual concentration–gradient contribution and a contribution, $vP(q,t)$, from the average velocity,

$$\begin{aligned} j(q,t) &= -D\frac{\partial P(q,t)}{\partial q} + vP(q,t) = -D\frac{\partial P(q,t)}{\partial q} \\ &\quad - \frac{1}{\zeta}\frac{\mathrm{d}G(q)}{\mathrm{d}q}P(q,t) = -D\frac{\partial P(q,t)}{\partial q} \\ &\quad - D\frac{\mathrm{d}[\beta G(q)]}{\mathrm{d}q}P(q,t) = -D\,\mathrm{e}^{-\beta G(q)}\frac{\partial}{\partial q}\mathrm{e}^{\beta G(q)}P(q,t). \end{aligned}$$

Substitution of this expression into the continuity equation $\partial P(q,t)/\partial t = -\partial j(q,t)/\partial q$ then gives equation (3).

The bulk experiment leading to equation (1*b*) can be mimicked by setting the probability density to 0 at $q = q_{prd}$ and to some constant $P(q_{rct})$ at $q = q_{rct}$. With these boundary conditions, the uniform steady-state flux can be determined as

$$j = \frac{P(q_{rct})\,\mathrm{e}^{\beta G(q_{rct})}}{\int_{q_{rct}}^{q_{prd}} \mathrm{d}q[D\,\mathrm{e}^{-\beta G(q)}]^{-1}}.$$

With an equilibrium distribution maintained around q_{rct}, the total probability in the reactant well is $c = \int_{-\infty}^{q^{\ddagger}} \mathrm{d}q\,P(q_{rct})\,\mathrm{e}^{-\beta[G(q)-G(q_{rct})]}$. By analogy to equation (1*b*), the forward rate constant is

$$k_{fwd} = \frac{j}{c} = \frac{1}{\int_{-\infty}^{q^{\ddagger}} \mathrm{d}q\,\mathrm{e}^{-\beta G(q)} \int_{q_{rct}}^{q_{prd}} \mathrm{d}q[D\,\mathrm{e}^{-\beta G(q)}]^{-1}}. \tag{5a}$$

This derivation essentially follows Kramers' classical work for barrier crossing [4]. The backward rate constant can be obtained by analogy,

$$k_{bwd} = \frac{1}{\int_{q^{\ddagger}}^{\infty} \mathrm{d}q\,\mathrm{e}^{-\beta G(q)} \int_{q_{rct}}^{q_{prd}} \mathrm{d}q[D\,\mathrm{e}^{-\beta G(q)}]^{-1}}. \tag{5b}$$

The ratio of the two rate constants is just what is expected from equilibrium distribution:

$$K = \frac{k_{fwd}}{k_{bwd}} = \frac{\int_{q^{\ddagger}}^{\infty} \mathrm{d}q\,\mathrm{e}^{-\beta G(q)T}}{\int_{-\infty}^{q^{\ddagger}} \mathrm{d}q\,\mathrm{e}^{-\beta G(q)}}. \tag{5c}$$

To obtain the rate constant in the form of equation (2), we need only to realize that the two integrals in the denominator of equation (5*a*) are predominantly contributed by small regions around q_{rct} and $q^{\ddagger}$, respectively. Within these regions, the energy function may be approximated as

$$G(q) = \omega_{rct}^2(q - q_{rct})^2/2 + G(q_{rct}) \tag{6a}$$

$$G(q) = \Delta G^{\ddagger} - \omega^{\ddagger 2}(q - q^{\ddagger})^2/2 + G(q_{rct}) \tag{6b}$$

respectively. Then

$$\begin{aligned} k_{fwd} &\approx \frac{D^{\ddagger}\,\mathrm{e}^{-\beta\Delta G^{\ddagger}}}{\int_{-\infty}^{\infty} \mathrm{d}q\,\mathrm{e}^{-\beta\omega_{rct}^2(q-q_{rct})^2} \int_{-\infty}^{\infty} \mathrm{d}q\,\mathrm{e}^{-\beta\omega^{\ddagger 2}(q-q^{\ddagger})^2}} \\ &= \frac{D^{\ddagger}\beta\omega_{rct}\omega^{\ddagger}}{2\pi}\,\mathrm{e}^{-\beta\Delta G^{\ddagger}} \end{aligned} \tag{7}$$

where $D^{\ddagger}$ is the value of the diffusion coefficient at $q^{\ddagger}$.

The dynamics can be modeled as diffusive only in the high-friction regime, where the rate constant is proportional to the diffusion coefficient and thus is inversely proportional to friction. At low friction, inertia effects cannot be ignored. In the inertia limit, the total energy of the molecule is conserved. If the function $G(q)$ is the potential energy, the molecule will be confined to the reactant well when its total energy is lower than $G(q^{\ddagger})$. Then the rate constant will be zero. Adding friction by increasing the interactions of the molecule with its bath will give the molecule a chance to cross energy levels (i.e., energy diffusion), thus increasing the rate constant. The foregoing arguments show that, for a barrier that is purely energetic, the rate constant decreases to zero at both high and low friction, and takes a maximal value at intermediate friction. This friction dependence is known as Kramers' turnover [5].

2.1.1. More rigorous formulation. If all molecules, with concentration C_0, are initially present in the reactant state and left alone, then eventually they will reach an equilibrium distribution with a fraction f_{rct} in the reactant state and a fraction f_{prd} in the product state. Let the concentration of reactant molecules be $C(t)$ at time t. A rate-equation description,

$$\text{Reactant} \underset{k_{bwd}}{\overset{k_{fwd}}{\rightleftharpoons}} \text{Product}$$

leads to an exponential decay, $\exp(-t/\tau_{\rm rlx})$, for the normalized relaxation function,

$$R(t) = \frac{C(t)/C_0 - f_{\rm rct}}{1 - f_{\rm rct}} \tag{8}$$

with the relaxation time given by

$$\tau_{\rm rlx}^{-1} = k_{\rm fwd} + k_{\rm bwd}. \tag{9}$$

$C(t)/C_0$ can be calculated from the conditional probability density, $p(q, t|q_0, 0)$ [6]:

$$C(t)/C_0 = \int_{-\infty}^{q^\ddagger} {\rm d}q \int_{-\infty}^{q^\ddagger} {\rm d}q_0\, p(q, t|q_0, 0)\, {\rm e}^{-\beta G(q_0)} \Big/ \int_{-\infty}^{q^\ddagger} {\rm d}q_0\, {\rm e}^{-\beta G(q_0)}. \tag{10}$$

The time integral of the relaxation function provides the best estimate for the relaxation time:

$$\tau_{\rm rlx} = \int_0^\infty {\rm d}t\, R(t). \tag{11a}$$

For diffusive dynamics, the following exact result is obtained for the time integral [7, 8]:

$$\begin{aligned}\tau_{\rm rlx} = {} & f_{\rm rct} \int_{-\infty}^{q^\ddagger} {\rm d}q [D\, {\rm e}^{-\beta G(q)}]^{-1} \\ & \times \left[\int_{-\infty}^{q} {\rm d}q_0\, {\rm e}^{-\beta G(q_0)}\right]^2 \Big/ \int_{-\infty}^{q^\ddagger} {\rm d}q_0\, {\rm e}^{-\beta G(q_0)} \\ & + f_{\rm prd} \int_{q^\ddagger}^{\infty} {\rm d}q [D\, {\rm e}^{-\beta G(q)}]^{-1} \\ & \times \left[\int_{q}^{\infty} {\rm d}q_0\, {\rm e}^{-\beta G(q_0)}\right]^2 \Big/ \int_{q^\ddagger}^{\infty} {\rm d}q_0\, {\rm e}^{-\beta G(q_0)}. \end{aligned} \tag{11b}$$

With the approximations of equation (6), the result in equation (7) can again be obtained.

The top of the barrier of a one-dimensional energy surface is naturally identified as the transition state. In higher dimensions, identifying the transition state is no longer straightforward. Sometimes a criterion has been invoked that the escape from the transition state to the reactant and product states has equal probabilities [9]. This criterion can be illustrated on the one-dimensional energy surface. It can be shown that $\gamma_{\rm rct}(q)$, the probability for a molecule, starting at q (between $q_{\rm rct}$ and $q_{\rm prd}$), to first reach the reactant well instead of the product well, satisfies the steady-state backward Smoluchowski equation [10]

$$L^\dagger \gamma_{\rm rct}(q) = {\rm e}^{\beta G(q)} \frac{\partial}{\partial q} \left[D\, {\rm e}^{-\beta G(q)} \frac{\partial}{\partial q} \gamma_{\rm rct}(q) \right] = 0 \tag{12}$$

with the boundary conditions

$$\gamma_{\rm rct}(q_{\rm rct}) = 1 \tag{13a}$$

$$\gamma_{\rm rct}(q_{\rm prd}) = 0. \tag{13b}$$

The solution is

$$\gamma_{\rm rct}(q) = \frac{\int_q^{q_{\rm prd}} {\rm d}q [D\, {\rm e}^{-\beta G(q)}]^{-1}}{\int_{q_{\rm rct}}^{q_{\rm prd}} {\rm d}q [D\, {\rm e}^{-\beta G(q)}]^{-1}}. \tag{14}$$

The probability $\gamma_{\rm prd}(q)$ for first reaching the product well can be similarly found and is given by $1 - \gamma_{\rm rct}(q)$. The ratio of the two capture probabilities starting at $q^\ddagger$ is thus

$$\begin{aligned}\frac{\gamma_{\rm prd}(q)}{\gamma_{\rm rct}(q)} &= \frac{\int_{q_{\rm rct}}^{q^\ddagger} {\rm d}q [D\, {\rm e}^{-\beta G(q)}]^{-1}}{\int_{q^\ddagger}^{q_{\rm prd}} {\rm d}q [D\, {\rm e}^{-\beta G(q)}]^{-1}} \\ &\approx \frac{\int_{-\infty}^{q^\ddagger} {\rm d}q\, {\rm e}^{-\beta\omega^{\ddagger 2}(q-q^\ddagger)^2/2}}{\int_{q^\ddagger}^{\infty} {\rm d}q\, {\rm e}^{-\beta\omega^{\ddagger 2}(q-q^\ddagger)^2/2}} = 1. \end{aligned} \tag{15}$$

In higher dimensions, the hyper-surface consisting of points from which the two capture probabilities are equal is known as the stochastic separatrix. The capture probability $\gamma_{\rm rct}$ (or $\gamma_{\rm prd}$) is close to 1 (or 0) within the reactant state, then has a sharp transition around the transition state, and reaches a value close to 0 (or 1) in the product state.

2.2. Bimolecular reactions

The rate equation for a bimolecular reaction

$$\mathrm{A} + \mathrm{B} \xrightarrow{K_{\rm a}} \mathrm{AB}$$

is

$$\frac{{\rm d}[\mathrm{AB}]}{{\rm d}t} = k_{\rm a}[\mathrm{A}][\mathrm{B}]$$

where [A], [B] and [AB] are the concentrations of the species. The simplest model for a bimolecular reaction consists of two spherical molecules interacting with a centrosymmetric potential $U(r)$ (figure 1(*B*)). At the contact distance $r = a$, the potential has a deep minimum corresponding to the bound state. The dissociation process is a unimolecular reaction; thus the theory developed above can be used to find the dissociation rate constant $k_{\rm d}$. It is obvious to identify the radial distance r as the reaction coordinate, and the Boltzmann factor ${\rm e}^{-\beta U(r)}$ and the geometric factor $4\pi r^2$ together define the free energy function:

$${\rm e}^{-\beta G(r)} = 4\pi r^2\, {\rm e}^{-\beta U(r)}.$$

Further identifying $q_{\rm rct}$ with a, $q^\ddagger$ with the outer boundary of the bound-state energy well, and $q_{\rm prd}$ with ∞, the dissociation rate constant is [11]

$$k_{\rm d} = \frac{1}{\int_0^{r^\ddagger} {\rm d}r\, {\rm e}^{-\beta G(r)} \int_a^\infty {\rm d}r [D\, {\rm e}^{-\beta G(r)}]^{-1}} \tag{16}$$

according to equation (5*a*), with D now identified as the mutual diffusion coefficient.

The equilibrium binding constant for this model is given by [12]

$$K_{\rm bi} = \int_0^{r^\ddagger} {\rm d}r\, 4\pi r^2\, {\rm e}^{-\beta U(r)} = \int_0^{r^\ddagger} {\rm d}r\, {\rm e}^{-\beta G(r)} \tag{17}$$

which must be the ratio of the association rate constant $k_{\rm a}$ and the dissociation rate constant $k_{\rm d}$. We thus find the association rate constant

$$k_{\rm a} = K_{\rm bi} k_{\rm d} = \frac{1}{\int_a^\infty {\rm d}r [D\, {\rm e}^{-\beta G(r)}]^{-1}}. \tag{18}$$

This is precisely the result obtained by Debye [13] and reduces to the Smoluchowski result $4\pi Da$ if $U(r) = 0$ for $r > a$ [14].

2.2.1. More realistic model. For binding processes involving protein molecules, either as the binding target or as the binding molecule, the above simple model has limited usage. An important property is binding specificity, i.e., only a very small portion of the target surface is involved in forming a specific complex. Models with small reactive patches on spherical molecules have been proposed [15–17]. Suppose that the two reactant molecules are rigid and their configuration is specified by three relative translation and six individual rotation degrees of freedom, collectively denoted as $\mathbf{q}$. If the interaction potential is $U(\mathbf{q})$, which has a deep minimum corresponding to the bound state, then the equilibrium binding constant, by extending equation (17), is

$$K_{\rm bi} = \int \mathrm{d}\mathbf{q}\mathcal{H}(\mathbf{q})\,\mathrm{e}^{-\beta U(\mathbf{q})}/(8\pi^2)^2 \tag{19}$$

where $\mathcal{H}(\mathbf{q})$ is a function that is 1 inside the bound state and 0 elsewhere, and the factor $(8\pi^2)^2$ is introduced to ensure proper normalization for the six rotation degrees of freedom.

The dissociation rate constant can be obtained by generalizing equations (8)–(10) to higher dimensions, with the bound and unbound states identified as reactant and product, respectively. For the present situation, it should be noted that $f_{\rm rct} \to 0$ and $k_{\rm bwd} = k_{\rm fwd} f_{\rm rct}/f_{\rm prd} \to 0$. Therefore $k_{\rm d}$, which is now equivalent to $k_{\rm fwd}$, is given by

$$k_{\rm d}^{-1} \approx \int_0^\infty \mathrm{d}t\, R(t) \approx \int_0^\infty \mathrm{d}t\, C(t)/C_0 \tag{20a}$$

$$\approx \int_0^\infty \mathrm{d}t \int \mathrm{d}\mathbf{q} \int \mathrm{d}\mathbf{q}_0 \mathcal{H}(\mathbf{q})\mathcal{H}(\mathbf{q}_0) \times p(\mathbf{q}, t|\mathbf{q}_0, 0)\,\mathrm{e}^{-\beta U(q_0)} \Big/ \int \mathrm{d}\mathbf{q}_0 H(\mathbf{q}_0)\,\mathrm{e}^{-\beta U(\mathbf{q}_0)}. \tag{20b}$$

The association rate constant is then obtained from $K_{\rm bi}k_{\rm d}$ as

$$k_{\rm a}^{-1} \approx (8\pi^2)^2 \int_0^\infty \mathrm{d}t \int \mathrm{d}\mathbf{q} \int \mathrm{d}\mathbf{q}_0 \mathcal{H}(\mathbf{q})\mathcal{H}(\mathbf{q}_0) \times p(\mathbf{q}, t|\mathbf{q}_0, 0)\,\mathrm{e}^{-\beta U(\mathbf{q}_0)} \Big/ \left[\int \mathrm{d}\mathbf{q}_0 \mathcal{H}(\mathbf{q}_0)\,\mathrm{e}^{-\beta U(q_0)}\right]^2. \tag{21}$$

This form of the association rate constant was first derived by Doi [18, 19] using the closure approximation of Wilemski and Fixman [20]. It has also been obtained by making an equilibrium distribution approximation [21], which generalizes the constant-flux approximation introduced by Shoup *et al* [15]. Using equation (21), Temkin and Yakobson [22] have carried out explicit calculations for spherical reactant molecules with reactive patches. In particular, for the binding of a pointlike ligand to a binding site spanning a polar angle α on an immobile spherical target with radius a, both of the last two groups of authors obtained the result

$$\frac{k_{\rm a}}{4\pi\, Da} = \frac{(1-\cos\alpha)^2}{\sum_{i=0}^\infty [P_{i-1}(\cos\alpha) - P_{i+1}(\cos\alpha)]^2/(i+1)(2i+1)} \tag{22a}$$

where $P_i(x)$ are Legendre polynomials. When $\alpha \to 0$, equation (22*a*) approaches

$$k_{\rm a}/4\pi\, Da = 3\pi\alpha/32. \tag{22b}$$

Note the linear dependence on α. One might have naively guessed that $k_{\rm a}$ scales with the surface fraction covered by the binding site. That would lead to a quadratic dependence on α and thus a much smaller magnitude for $k_{\rm a}$. The reason for a value of $k_{\rm a}$ much higher than expected from surface fraction is that the ligand, due to its diffusive dynamics, can make repeated attempts at the binding site.

Generalization of equation (22*a*) to the binding of two spherical reactants each bearing a reactive patch has also been obtained [16, 22, 23]. For patches spanning angles (α_1 and α_2) that are small, the result is [16, 23]

$$k_{\rm a}/4\pi\, Da = F_1 \tan(\alpha_2/2) + F_2 \tan(\alpha_1/2) \tag{22c}$$

where $F_i = (1-\cos\alpha_i)/2$ are the surface fractions of the patches. The rate has a third-order dependence on α_1 and α_2, instead of the fourth-order dependence as expected from scaling $4\pi\, Da$ with the surface fractions.

3. Strategies for speed

3.1. Enzyme catalysis

Because of the exponential dependence of rate constant on energy barrier (equation (2)), the most obvious strategy for speeding up a chemical reaction is to lower the energy barrier. This is just what an enzyme does.

Let us first consider the unimolecular reaction of a molecule S that goes through a transition state (denoted as ‡) and then forms the product P:

$$\mathrm{S} \to \ddagger \to \mathrm{P}.$$

In the absence of an enzyme, the rate constant can be written as

$$k_{\rm chem} = k_0\,\mathrm{e}^{-\beta(G^\ddagger - G_{\rm S})}. \tag{23a}$$

The enzyme-catalyzed reaction proceeds by first binding to the enzyme E:

$$\mathrm{E} + \mathrm{S} \to \mathrm{ES} \to \mathrm{E}^\ddagger \to \mathrm{EP} \to \mathrm{E} + \mathrm{P}.$$

After the substrate S is bound to the enzyme, the rate constant of forming the product becomes

$$k_{\rm cat} = k_0\,\mathrm{e}^{-\beta(G_{\mathrm{E}^\ddagger} - G_{\rm ES})} \tag{23b}$$

where it has been assumed that the prefactor k_0 is unchanged by the enzyme and that release of P is not rate-limiting. Taking the ratio of the two rate constants, we find

$$k_{\rm cat}/k_{\rm chem} = \mathrm{e}^{-\beta(G_{\mathrm{E}^\ddagger} - G_{\rm ES})}/\mathrm{e}^{-\beta(G^\ddagger - G_{\rm S})}. \tag{24a}$$

Two alternate routes lead to the enzyme-bound transition state $\mathrm{E}^\ddagger$ from unbound E and S:

$$\begin{array}{ccc} \mathrm{E}+\mathrm{S} & \xrightarrow{K_{\rm s}} & \mathrm{ES} \\ \downarrow & & \downarrow \\ \mathrm{E}+\mathrm{S}^\ddagger & \xrightarrow{K_\ddagger} & \mathrm{E}^\ddagger. \end{array}$$

Thermodynamically the routes are equivalent,

$$\mathrm{e}^{-\beta(G^\ddagger - G_{\rm S})} K^\ddagger = K_{\rm S}\,\mathrm{e}^{-\beta(G_{\mathrm{E}^\ddagger} - G_{\rm ES})}$$

where $K^\ddagger$ and $K_{\rm S}$ are the binding constants of the transition state and the substrate, respectively, to the enzyme. Therefore

$$k_{\rm cat}/k_{\rm chem} = K^\ddagger/K_{\rm S} \tag{24b}$$

which indicates that an enzyme speeds up a reaction by having higher affinity for the transition state than for the substrate. This is an idea first proposed by Pauling [24] and now being stressed once again [25, 26]. At first glance it may suggest designing enzymes through lowering affinity for the substrate. However, an enzyme must have a sufficiently high affinity for the substrate to entice its binding—only then can the catalytic reaction proceed. A good enzyme molecule, therefore, should also have the flexibility for it to reach the more favorable binding with the transition state.

Now consider the bimolecular reaction of two molecules S_1 and S_2 to form a product P (e.g., the synthesis of ATP from ADP and phosphate). Suppose that the reaction proceeds by first forming a complex S_1S_2, which then rearranges to form the product:

$$S_1 + S_2 \rightarrow S_1S_2 \xrightarrow{k_r} P.$$

If the rearrangement rate constant k_r is much smaller than the dissociation rate constant of S_1S_2, then the overall product formation rate constant can be approximated as

$$k_{chem} = K_0 k_r \tag{25a}$$

where K_0 is the equilibrium constant for forming the S_1S_2 complex. The enzyme-catalyzed reaction proceeds first by the binding of substrate S_1 to the enzyme (with binding constant K_{S1}):

$$E + S_1 \rightarrow ES_1.$$

At this point substrate S_2 binds to the enzyme, thereby positioning the two substrates in proximity for them to rearrange into the product:

$$ES_1 + S_2 \rightarrow ES_1S_2 \xrightarrow{K_s} EP \rightarrow E + P.$$

Suppose that the rearrangement rate constant is still k_r and is again much smaller than the dissociation rate constant of S_2 (this time from the ES_1 complex). Then the product formation rate constant becomes

$$k_{cat} = K_{S_2} k_r \tag{25b}$$

where K_{S_2} is the binding constant of S_2 to the ES_1 complex. The ratio of the two rate constants is

$$k_{cat}/k_{chem} = K_{S_2}/K_0. \tag{26a}$$

To make a connection between equations (26*a*) and (24*b*), let us introduce the fictitious route of forming the ternary complex ES_1S_2 by the binding of E with the S_1S_2 complex (with binding constant $K_{S_1S_2}$):

$$\begin{array}{ccc} E + S_1 + S_2 & \xrightarrow{K_{s_1}} & ES_1 + S_2 \\ K_0 \downarrow & & \downarrow K_{S_2} \\ E + S_1S_2 & \xrightarrow{K_{s_1s_2}} & ES_1S_2. \end{array}$$

The two alternate routes of forming ES_1S_2 from unbound E, S_1 and S_2 are thermodynamically equivalent; thus

$$K_0 K_{S_1S_2} = K_{S_1} K_{S_2}.$$

Equation (26*a*) now becomes

$$k_{cat}/k_{chem} = K_{S_1S_2}/K_{S_1} \tag{26b}$$

which suggests that here S_1S_2 plays the role of transition state. Again, the speed up is achieved through preferential binding, now S_1S_2 over S_1. The binding of S_1S_2 to E is bidentate in the sense that the enzyme has two specific binding sites, one for S_1 and one for S_1. The affinity enhancement of bidentate binding has been studied in simple models [27].

Good enzymes are endowed with both high efficiency and high selectivity. Selectivity can easily be obtained by burying the active site, with only a narrow channel leading to it. The narrow channel potentially limits the rate of substrate binding (k_S), which in turn limits the rate of the enzymatic reaction. However, it has been shown that rapid conformational gating along the channel can retain selectivity while minimally sacrificing k_S [28]. Once inside the channel, the substrate makes repeated diffusional attempts at entering the gate, and rapid fluctuations of the gate between open and closed allow for a successful entry.

3.1.1. Substrate channeling. In recent years a number of bifunctional enzymes have been discovered to have their distant active sites connected by a tunnel [29]. This molecular architecture immediately suggests channeling of reaction intermediates from one active site to another, thereby increasing catalytic efficiency. Another mechanism proposed for substrate channeling is a route along the enzyme surface, perhaps maintained by electrostatic interactions [30, 31].

Consider the reaction of a molecule S that goes through an intermediate I before forming the product P:

$$S \rightarrow I \rightarrow P.$$

Suppose that a bifunctional enzyme E catalyzes the first step at one active site and the second step at another. Let us now develop the simplest model for the catalyzed reactions that allows for the possibility of intermediate channeling (figure 2(*A*)). First, we assume that the enzyme concentration [E] is in excess. If S binds to E with rate constant k_S, and upon binding an intermediate I is instantaneously released just outside the first active site, then the concentrations of S and such newly generated I (denoted as I*) are

$$[S](t) = [S]_0 \, e^{-k_S[E]t} \tag{27a}$$

$$[I^*](t) = [S]_0 (1 - e^{-k_S[E]t}) \tag{27b}$$

where $[S]_0$ is the initial concentration of S. I* can have three different fates: (1) geminate binding, to the second active site on same enzyme molecule; (2) bimolecular binding, to that site on a different enzyme molecule; and (3) release to the solvent. If the probabilities that I* survives geminate and bimolecular binding are ψ_{gem} and $\psi_{bi}(t)$, respectively, after being generated at $t = 0$, then the probability that it will survive binding altogether (i.e., be released to the solvent) is approximately [32]

$$\psi(t) = \psi_{gem} \psi_{bi}(t). \tag{28}$$

If I binds to E with rate constant k_I, then

$$\psi_{bi}(t) = e^{-k_I[E]t}. \tag{29}$$

We will come back to ψ_{gem} later. At time $t = \tau$, I* is being generated at a rate $d[I^*](\tau)/d\tau$, and this will have a

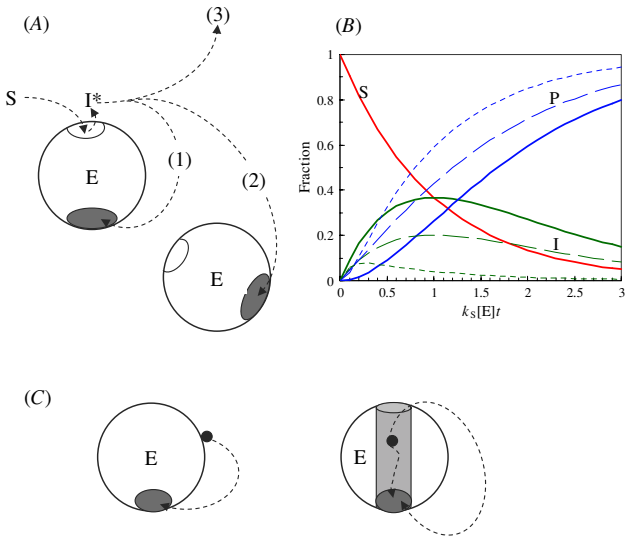

Figure 2. (*A*) Catalytic reaction of a bifunctional enzyme. The newly generated intermediate I* has three possible fates. (*B*) Time dependence of the fractions of S, I and P. Solid, dashed and dotted curves are for (1) $k_I/k_S = 1$ and $\psi_{gem} = 1$, (2) $k_I/k_S = 1$ and $\psi_{gem} = 0.55$, and (3) $k_I/k_S = 10$ and $\psi_{gem} = 1$, respectively. (*C*) Models for calculating ψ_{gem}. For the model on the right, the intermediate can reach the second binding site either directly through the tunnel or by first diffusing out and then coming back to the binding site.

survival probability of $\psi(t-\tau)$ at time t. Accumulating all the surviving intermediates, we find the concentration of I to be

$$[\mathrm{I}](t) = \int_0^t \mathrm{d}\tau\, \psi(t-\tau)\, \mathrm{d}[\mathrm{I}^*](\tau)/\mathrm{d}\tau \tag{30a}$$

$$= [\mathrm{S}]_0 \psi_{gem} k_S (\mathrm{e}^{-k_I[\mathrm{E}]t} - \mathrm{e}^{-k_S[\mathrm{E}]t})/(k_S - k_I). \tag{30b}$$

If the product is formed instantaneously upon binding of I, then the product concentration is

$$[\mathrm{P}](t) = [\mathrm{S}]_0 - [\mathrm{S}](t) - [\mathrm{I}](t) \tag{31a}$$

$$= [\mathrm{S}]_0(1 - \mathrm{e}^{-k_S[\mathrm{E}]t}) - [\mathrm{S}]_0 \psi_{gem} k_S (\mathrm{e}^{-k_I[\mathrm{E}]t} - \mathrm{e}^{-k_S[\mathrm{E}]t})/(k_S - k_I). \tag{31b}$$

From equations (30*b*) and (31*b*), it is clear that the geminate route lowers the intermediate concentration and correspondingly increases the product concentration.

Figure 2(*B*) shows the time dependence of the concentrations of S, I and P. When $k_S = k_I$ and $\psi_{gem} = 1$ (i.e., in the absence of substrate channeling), the intermediate can reach as much as 37% of initial substrate concentration. The presence of substrate channeling, with $\psi_{gem} = 0.55$ (i.e., 45% capture probability by geminate binding), reduces the maximal fraction of I to 20%. Interestingly, increasing k_I over k_S has the same effect of lowering the intermediate concentration. At $k_I/k_S = 10$ (and $\psi_{gem} = 1$), the maximal fraction of I is reduced to just 8%. The last observation may explain recent experimental results of Anderson and co-workers [33, 34]. In the bifunctional enzyme thymidylate synthase-dihydrofolate reductase (TS-DHFR), the TS site catalyzes the transformation of CH_2H_4-folate to the intermediate H_2-folate, which is subsequently transformed to H_4-folate at the DHFR site. The two active sites of *Leishmania major* TS-DHFR have different catalytic rates, with that for DHFR about 10-fold higher. Therefore it is not surprising to see the fraction of H_2-folate peaking at about only 14%. The significantly higher catalytic rate at the DHFR site alleviates the need for (or masks the manifestation of) substrate channeling, which has been hypothesized previously by invoking the electrostatic mechanism [30, 31]. Anderson and co-workers [33] specifically tested the electrostatic channeling hypothesis by reversing or neutralizing charged residues that putatively play the channeling role, and did not find any increase in the intermediate fraction. However, any electrostatic channeling might be tempered by the relatively high ionic strength used in the experiments (and perhaps suppressed by other charged molecules competing against the intermediate [35]). In a separate study on *Cryptosporidium hominis* TS-DHFR, which shows comparable TS and DHFR catalytic rates, Atreya and Anderson [34] found the maximal fraction of the intermediate to rise to 65%. The question of whether there is substrate channeling in TS-DHFR appears still open.

To gain further insight into substrate channeling, let us calculate $\gamma_{gem} = 1 - \psi_{gem}$, the probability of being captured through geminate binding, on a simple model. The model consists of a pointlike intermediate I diffusing to an absorbing binding site located around the south pole, spanning a polar angle α, of a spherical enzyme with radius a (left panel in figure 2(*C*)). The capture probability satisfies the steady-state backward Smoluchowski equation, with a value of 1 on the binding site and 0 at infinity (cf equations (12) and (13)). Of interest to the present situation is the capture probability of I starting at polar angle θ on the enzyme surface, which can be found by the constant-flux approximation as

$$\gamma_{gem}(\theta) = F(\alpha) \sum_{i=0}^{\infty} [P_{i-1}(\cos\alpha) - P_{i+1}(\cos\alpha)] P_i(\cos\theta)/A_i \tag{32a}$$

with

$$F(\alpha) = \frac{(1-\cos\alpha)}{\sum_{i=0}^{\infty} [P_{i-1}(\cos\alpha) - P_{i+1}(\cos\alpha)]^2 / A_i (2i+1)} \tag{32b}$$

$$A_i = i + 1. \tag{32c}$$

For a binding site spanning 5°, $\gamma_{gem}(30^\circ) = 6.5\%$, so the chance for geminate rebinding is small. An attractive potential may keep I around the enzyme surface longer and thus increase the capture probability. In the presence of a square-well potential $U = -\varepsilon_0$ in $a < r < a_1$ and 0 elsewhere, equation (32*c*) becomes

$$A_i = (i+1) \times \frac{1 - (1 - \mathrm{e}^{-\beta\varepsilon_0}) i (a/a_1)^{2i+1} / [i + (i+1)\,\mathrm{e}^{-\beta\varepsilon_0}]}{1 + (1 - \mathrm{e}^{-\beta\varepsilon_0})(i+1)(a/a_1)^{2i+1} / [i + (i+1)\,\mathrm{e}^{-\beta\varepsilon_0}]}. \tag{32d}$$

For a specific starting angle θ in a potential with a given depth ε_0, the capture probability takes a maximal value at a particular width of the potential. For example, for $\beta\varepsilon_0 = 2$, a maximal

value of 24% is obtained for $\gamma_{\rm gem}(30^\circ)$ at $a_1 = 1.13a$. When the width of the potential is too short, the effect of the potential becomes negligible; on the other hand, when the width is too large, the potential will take I away from the enzyme surface.

Another mechanism for increasing the capture probability is suggested by enzyme structures with tunnels connecting the location where I is initially released and the final binding site. The right panel in figure 2(*C*) illustrates such a situation, where the initial location of I is inside a narrow cylindrical tunnel leading to the final binding site at the south pole. The capture probability can be found as

$$\gamma_{\rm gem}(d) = 1 - \frac{F_{\rm e}(\alpha)d/a}{1 + F_{\rm e}(\alpha) + F_{\rm o}(\alpha)} \tag{33}$$

where d is the distance between the initial location and the final binding site, and $F_{\rm e}(\alpha)$ and $F_{\rm o}(\alpha)$ are given in equation (32*b*), except that the summation over i includes only even and odd values, respectively. In the absence of a potential, $\gamma_{\rm gem}(1.5a) = 28\%$ and $\gamma_{\rm gem}(a) = 52\%$. The addition of an attractive potential increases the capture probability. At $\beta\varepsilon_0 = 2$, $\gamma_{\rm gem}(1.5a)$ and $\gamma_{\rm gem}(a)$ become 34% and 56%, respectively, at the same optimal width of $a_1 = 1.41a$.

3.2. Protein folding

A folded protein is stabilized by many specific interactions, as seen in an x-ray or NMR structure. When the protein is unfolded the interactions are lost, but now the protein chain can sample a vast number of configurations; hence the unfolded state is favored by chain entropy. As illustrated in figure 3(*A*), when an unfolded protein molecule starts to fold, it first has to pay chain-entropy cost. As the chain becomes more and more compact, the chain entropy is reduced further and further, but specific interactions also start to accumulate. Eventually the free-energy contributions from the interactions overcome the entropy cost and lead the molecule to the folded state. This description suggests a free-energy barrier for the folding process, with a significant chain-entropy component [36, 37]. This component has been explicitly built into many models for protein folding [38–50]. Below we further analyze the role of chain entropy, with a particular eye toward the concept of transition-state ensemble, address the question of why protein folding is fast, and propose a scenario of dominant folding routes amid vast possibilities.

A simple illustration of the folding problem is provided by the Zwanzig model [38, 39], in which the configuration of a protein with N residues is described by $n_{\rm f}$, the number of folded residues (figure 3(*B*)). Each residue is supposed to have $\nu + 1$ conformations, one of which is folded. Each folded residue is assumed to lower the energy of the protein by ε, and the fully folded configuration (with all N folded residues) is favored by an additional energy term ε_0. The energy function of the protein is thus

$$U(n_{\rm f}) = -\varepsilon n_{\rm f} - \varepsilon_0 \delta_{n_{\rm f}N}. \tag{34}$$

The degeneracy of the microstate with $n_{\rm f}$ folded residues and $s = N - n_{\rm f}$ unfolded residues is

$$\Omega(n_{\rm f}) = \nu^s C_N^s \tag{35}$$

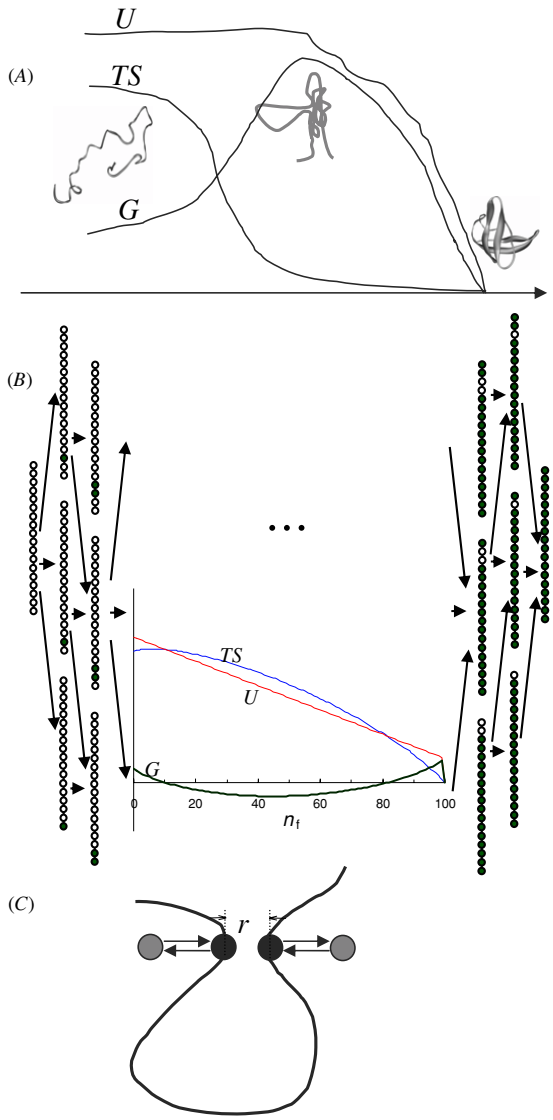

Figure 3. (*A*) Model of protein folding and entropic barrier. (*B*) Zwanzig model. Open and filled circles represent unfolded and folded residues, respectively. Inset shows the energetic and entropic components of the free-energy functional and the functional itself as functions of $n_{\rm f}$. (*C*) Model of contact formation. Residues, represented by small spheres, can fluctuate between native (black) and nonnative (gray) conformations.

where the first factor arises because each unfolded residue has ν choices of conformation and the combinatorial factor accounts for the different choices of selecting s unfolded residues out of the protein chain. The partition function is

$$Q = \sum_{n_{\rm f}=0}^{N} \Omega(n_{\rm f})\, {\rm e}^{-\beta U(n_{\rm f})} = {\rm e}^{\beta(N\varepsilon+\varepsilon_0)} + {\rm e}^{\beta N\varepsilon}[(1+\nu\,{\rm e}^{-\beta\varepsilon})^N - 1]. \tag{36}$$

The equilibrium probability of a microstate is

$$P_{\rm eq}(n_{\rm f}) = \Omega(n_{\rm f})\, {\rm e}^{-\beta U(n_{\rm f})}/Q \tag{37}$$

from which one can form a free-energy functional

$$G(n_f) = -k_B T \ln[Q P_{eq}(n_f)] = U(n_f) - k_B T \ln \Omega(n_f)$$
$$= U(n_f) - s k_B T \ln \nu - k_B T \ln C_N^s. \quad (38)$$

In this functional, $k_B \ln \Omega(n_f)$ appears appropriately as entropy, and $U(n_f)$ may well be a free energy (accounting for solvent-mediated effects such as hydrophobic interactions). For notational convenience, we will continue to refer to contributions arising from specific interactions between residues as energetic, to be distinguished from chain entropy.

In this model, n_f can easily be identified as the (discrete) reaction coordinate, and the functional $G(n_f)$ sets the free energy surface. As figure 3(*B*) shows, the energetic component $U(n_f)$ decreases with increasing n_f, the entropic component $k_B T \ln \Omega(n_f)$ at first increases slightly (due to the combinatorial factor C_N^s), and then steadily decreases. The free-energy surface has a barrier at $n_f = N - 1$, just before reaching the fully folded microstate. It is clear that the free-energy barrier, i.e., transition state, here arises because of the asynchronous decrease of the energetic and entropic components. The equilibrium fraction of the folded state is

$$f_f = P_{eq}(N) = e^{\beta(N\varepsilon + \varepsilon_0)}/Q$$

and the corresponding quantity for the unfolded state is $f_u = 1 - f_f$.

Determination of the folding rate requires the specification of the dynamics among the conformations, which is modeled as rate processes. Specifically, the transition of each residue from unfolded to folded conformation is supposed to occur with rate constant k_0. The conversion from the microstate with n_f folded residues to the microstate with one more folded residue then has a rate of $\omega(n_f \to n_f + 1) = s k_0$, since each of the s unfolded residues has a chance of making a transition. The rate for the reverse process is then determined by the requirement of detailed balance:

$$\omega(n_f \to n_f + 1) P_{eq}(n_f) = \omega(n_f + 1 \to n_f) P_{eq}(n_f + 1)$$

leading to $\omega(n_f + 1 \to n_f) = (n_f + 1) k_0 \nu\, e^{-\varepsilon/k_B T}$ for $n_f < N - 1$ and $N k_0 \nu\, e^{-(\varepsilon+\varepsilon 0)/k_B T}$ for $n_f = N - 1$.

According to TST (equation (1*c*)), the folding rate constant k_f can be estimated as the product of two factors. The first is the equilibrium probability to be at the transition state, and is given by $P_{eq}(N-1)/f_u$. The second is the rate of converting from the transition state to the folded state, which is $\omega(N - 1 \to N) = k_0$. Thus

$$k_{f;TST} = k_0 P_{eq}(N-1)/f_u = k_0 N \nu\, e^{-\beta\varepsilon}/[(1 + \nu\, e^{-\beta\varepsilon})^N - 1].$$

As noted by Zwanzig *et al* [38], the energetic bias $-\varepsilon$ upon folding each residue is essential for bringing the folding rate into the realm of real proteins. In the present context, this bias leads to an appreciable equilibrium probability for reaching the transition state. The unfolding rate constant can be similarly estimated according to TST,

$$k_{u;TST} = k_0 N \nu\, e^{-\beta(\varepsilon+\varepsilon_0)}$$

and the resulting relaxation time is given by

$$\tau_{rlx;TST}^{-1} = k_{f;TST} + k_{u;TST} = k_0 N \nu\, e^{-\beta\varepsilon}/[(1 + \nu\, e^{-\beta\varepsilon})^N - 1]$$
$$+ k_0 N \nu\, e^{-\beta(\varepsilon+\varepsilon_0)}. \quad (39)$$

It can easily be checked that equation (39) is identical to an approximate result obtained by Zwanzig [39] according to the assumption of local thermodynamic equilibrium. The full model can be solved exactly and the result shows that equation (39) is a good approximation. Later, Doyle *et al* [40] and Oliva and Muñoz [51] explored extensions of the Zwanzig model with the Kronecker delta function in equation (34) replaced by more gradual functions. Plaxco *et al* [41] recognized a significant correlation between folding rates and the relative contact order O/N and proposed replacing $U(n_f)$ by $-\varepsilon(n_f/N)^O$ to provide an explanation.

In short, the Zwanzig model touched on some central issues in protein folding. It illustrated that the loss of chain entropy upon folding builds up a free-energy barrier, and that free-energy contributions of native contacts lower this barrier and are a key to the question of why protein folding is fast. However, more realistic models require consideration of multiple reaction coordinates undergoing different dynamic processes, resulting in multiple folding routes and transition-state ensembles with rich features. In some of these models microstates are identified by particular stretches of folded residues [43–46, 48]. To reduce the number of microstates to be enumerated, limits are imposed on the maximal number of stretches of folded residues [44, 45] or number of disordered loops [43]. Free-energy functionals are constructed in the same spirit as equation (38). In forming the energetic term, consideration is given to contributions of native contacts [43, 45, 46], hydrophobic effects as modeled by burial of solvent accessible surface areas [44, 48], and hydrogen bonding [46, 48]. For the entropic term, there is no longer a combinatorial factor since the microstates are now identified by specific residues, and in some models the residue entropy loss $-k_B \ln \nu$ is residue-type specific. In addition, it has been recognized that, when two nonlocal residues form a native contact, besides the entropy loss of these two residues, the intervening unfolded residues also suffer a 'loop entropy' loss because the end-to-end distance r takes a specific value a. Suppose that the probability density $P(r)$ for r is Gaussian,

$$P(r) = 4\pi r^2 p(r) \quad (40a)$$
$$p(r) = (3/2\pi \langle r^2 \rangle)^{3/2}\, e^{-3r^2/2\langle r^2 \rangle} \quad (40b)$$

where $\langle r^2 \rangle$ is the mean square distance and is proportional to the sequence separation l of the end residues,

$$\langle r^2 \rangle = b^2 l. \quad (41)$$

The loop entropy is then [52]

$$S_{loop}(l, a) = k_B \ln[\mathcal{N} P(a)] = k_B \ln[4\pi d^2 \mathcal{N} (3/2\pi b^2)^{3/2}]$$
$$- 3k_B a^2/2b^2 l - (3k_B/2) \ln l \quad (42)$$

where $\mathcal{N}$ is an arbitrary constant with dimensionality of r.

In other models microstates are based on native contacts [42, 49, 50]. Suppose that the native contacts are numbered sequentially as $m = 0, 1, 2, \ldots, M$, each of which implicates two specific residues that in the folded protein are at a distance below a cutoff. The model of Shoemaker and Wolynes [42] centered on the probabilities p_m of forming these contacts. The values of p_m were determined by minimizing a free-energy

functional $G(\{p_m\})$, subject to the constraint that the sum of p_m is a given number of native contacts formed:

$$\sum_{m=0}^{M} p_m = \mu \tag{43}$$

where μ ranges from 0 to M. If the values of p_m thus determined are denoted as $p_m(\mu)$, then the dependence of the free-energy functional on μ (i.e., the number of native contacts formed) is given by $G[\{p_m(\mu)\}]$. A transition state can then be identified as the maximum of $G[\{p_m(\mu)\}]$ between $\mu = 0$ and $\mu = M$.

Shoemaker and Wolynes calculated the loop-entropy loss of each native contact simply by using the sequence separation of that contact. It should be noted that the distribution of the end-to-end distance will be perturbed by contacts formed by any intervening residues. Makarov *et al* [49] explicitly considered the mutual influences of native contacts on chain-entropy loss. On the other hand, Makarov *et al* did not consider any energetic contribution that may arise from cooperative effects among native contacts (which Shoemaker and Wolynes included); there may also be a serious problem with the neglect of excluded-volume effects [53]. In a more coarse-grained model, Weikl *et al* [50] considered microstates that are identified by clusters of native contacts. The loop-entropy loss of each cluster also depends on whether other clusters have formed.

For a sequential model like Zwanzig's, it is easy to identify the transition state as the top of the barrier. For models with more complex connectivities among microstates, various strategies have been proposed to identify the transition-state ensemble. One possibility is to obtain the potential of mean force along a single reaction coordinate. For example, in a model with a free-energy functional $G(m)$ for microstates m, the potential of mean force for a reaction coordinate q is

$$\mathrm{e}^{-\beta G(q)} = \sum_{m} \delta(q)\, \mathrm{e}^{-\beta G(m)} \tag{44}$$

where $\delta(q)$ is 1 if the value of the reaction coordinate is q and 0 otherwise. However, the free-energy barrier along such a reduced reaction coordinate has been found to be broad and rugged [44, 45], making it difficult to identify a transition state with a single value of the reaction coordinate. A more general approach is to extensively sample routes between the fully unfolded and fully folded microstates [43, 44, 50]. Along each route a 'barrier' microstate with the highest free energy is found. The transition-state ensemble then consists of the barrier microstates with the *lowest* free energies. A third approach, requiring the specification of the dynamics among the microstates (see below), is to identify the transition-state ensemble with the stochastic separatrix [9]. From the separatrix, the probabilities of reaching either the fully unfolded or the fully folded microstate first are nearly equal (equation (15)). Berezhkovskii and Szabo [54] have recently developed approximate solutions to the capture probabilities.

The transition-state ensemble provides insight into the folding mechanism. Another important reason for the focus on the transition-state ensemble is that it (perhaps!) can directly be checked by experiments, through a procedure known as Φ-analysis [55]. The Φ-value of a residue is defined by the effects of its mutation on the folding free energy ($= -k_B T \ln(f_f/f_u)$) and on the folding rate (k_f):

$$\Phi = \frac{\ln(k_f'/k_f)}{\ln[(f_f'/f_u')/f_f/f_u)]} \tag{45}$$

where the unprimed and primed quantities refer to those before and after the mutation, respectively. Suppose that the residue contributes $\Delta\Delta G$ to the folded-state free energy and nothing to the unfolded-state energy. The contribution $\Delta\Delta G^{\ddagger}$ of the residue to the transition-state free energy may be as little as zero (if the residue and its would-be contact residues behave like in the unfolded state) or as much as $\Delta\Delta G$ (if the residue and its surrounding are native-like). If $\Delta\Delta G^{\ddagger}$ and $\Delta\Delta G$ are measured by the numerator and denominator respectively, then a Φ-value of 1 or 0 means that the residue is folded or unfolded, respectively, in the transition state. Explanations for Φ-values of individual proteins have motivated much of the theoretical development. However, a note of caution is needed in comparing experimental Φ-values, which are results of both energetics and dynamics, with transition-state ensembles obtained purely on energetic considerations. The proper comparison should be with predicted effects of mutations on the folding kinetics and equilibrium, as has been done in some cases [45, 56]. The very presence of residues with Φ-value close to 1 sends a powerful message. Although in principle there are a vast number of routes that connect the unfolded state to the folded state, there may just be a small number of dominating routes with low free-energy barriers. These dominating routes involve the telltale residues with high Φ-values.

Another set of experimental data motivating theoretical development is the large variation in folding rates among different proteins [41]. In general, proteins forming proportionally more local native contacts (as opposed to nonlocal ones) fold faster. A simple qualitative explanation is that the chain-entropy cost upon forming local contacts is much less relative to nonlocal contacts, leading to a lower free-energy barrier (cf Makarov *et al* [49]). For proteins adopting the same structural fold, folding rates can also vary significantly. Increasing propensities for secondary structures (leading to less residue entropy loss) and strengthening tertiary contacts appear to be good strategies for speeding up folding [57–59].

Given the suggested role of free-energy contributions of native contacts in lowering the free-energy barrier raised by chain entropy, it is puzzling at first sight to find positive activation enthalpies for many proteins at room temperature by analyzing temperature dependence of the folding rate, as done by Akmal and Muñoz [37] (also see Liu and Chan [60] for a somewhat different analysis). A closer look resolves this puzzle and provides some insight into the transition state. Suppose that the free-energy contributions arising from interactions within the protein and with the solvent can be written as

$$G_{\mathrm{int}} = U_{\mathrm{int}} + G_{\mathrm{solv}}(T) \tag{46a}$$

where U_{int} accounts for van der Waals and other interactions within the protein, taken to be temperature independent, and

$G_{\rm solv}(T)$ accounts for solvation effects. The latter can be further decomposed into enthalpy and entropy:

$$G_{\rm solv}(T) = H_{\rm solv} - T S_{\rm solv} \tag{46b}$$

where temperature dependence of $H_{\rm solv}$ and $S_{\rm solv}$ is assumed. From here on we will fix at room temperature and use the unfolded state as reference.

Quite generally it can be assumed that, as the protein goes from the unfolded state through the transition state to the folded state, $U_{\rm int}$ decreases whereas both $H_{\rm solv}$ and $S_{\rm solv}$ increase (but with $G_{\rm solv}$ decreasing). Consider, for example, $U_{\rm int} = -5$ and -70 kcal mol^{-1} and $H_{\rm solv} = 20$ and 40 kcal mol^{-1} at the transition and folded states, respectively. The total enthalpy of the protein, $U_{\rm int} + H_{\rm solv}$, will be 15 and -30 kcal mol^{-1} at the transition and folded states, respectively. The positive activation enthalpy of 15 kcal mol^{-1} thus comes from the asynchronous change of $U_{\rm int}$ and $H_{\rm solv}$, much like the appearance of the transition state from the asynchronous change of chain entropy and $G_{\rm int}$. Akmal and Muñoz also found a positive activation entropy, which can be similarly rationalized by asynchronous change of the chain entropy ($S_{\rm chain}$) and the solvation entropy $S_{\rm solv}$. For example, for CI2, the analysis of Akmal and Muñoz finds $TS_{\rm chain} \sim -40$ and -120 kcal mol^{-1} and $TS_{\rm solv} \sim 50$ and 100 kcal mol^{-1} at the transition and folded states, respectively. The total entropic component of the protein, $T(S_{\rm chain} + S_{\rm solv})$, is thus $\sim$10 and -20 kcal/mol at the transition and folded states, respectively. These illustrative numbers seem to suggest that, as the protein goes first to the transition state and then to the folded state, the changes of $H_{\rm solv}$ and $S_{\rm solv}$ are gradual over the two steps but the changes of $U_{\rm int}$ and $S_{\rm chain}$ are small in the first step and abrupt in the second step. Such a result can be rationalized under the sensible assumption that numerous fine rearrangements of the protein chain and residues occur after the transition state, to which $U_{\rm int}$ and $S_{\rm chain}$ are sensitive, whereas the former quantities, being solvation effects, are less so.

Let us get back to the theoretical models. The question remains: how can the folding rate be calculated? A simple approach is to obtain a free-energy functional for a single reaction coordinate (equation (44)) from a more detailed model, and introduce an effective diffusion constant along this coordinate [45, 47]. The folding rate can then be calculated according to simple rate theory (e.g., equation (5*a*) or (11*b*)). In a more general approach, conversions between microstates are explicitly modeled as rate processes, and the evolution of the microstates is modeled by a master equation [39–41, 48–50, 56, 61–65]. This can generally be written in the form

$$\frac{{\rm d}\mathbf{P}(t)}{{\rm d}t} = \boldsymbol{\omega} \cdot \mathbf{P}(t) \tag{47}$$

where $\boldsymbol{\omega}$ is the matrix of conversion rates satisfying detailed balance

$$\omega_{mm'} P_{\rm eq}(m') = \omega_{m'm} P_{\rm eq}(m) \tag{48a}$$

and with diagonal elements

$$\omega_{mm} = -\sum_{m' \neq m} \omega_{m'm}. \tag{48b}$$

Equation (48) means that $\boldsymbol{\omega}$ has a zero eigenvalue, with the equilibrium probability $\mathbf{P}_{\rm eq}$ as the corresponding eigenvector:

$$\boldsymbol{\omega} \cdot P_{\rm eq} = 0. \tag{48c}$$

All other eigenvalues are negative. Let these be $-\lambda_m$, which are ordered from low to high magnitudes, and the corresponding eigenvectors be $\mathbf{P}_m$. Then the solution of equation (47) can be formally written as

$$\mathbf{P}(t) = A_0 \mathbf{P}_{\rm eq} + A_1 \mathbf{P}_1 \, {\rm e}^{-\lambda_1 t} + A_2 \mathbf{P}_2 \, {\rm e}^{-\lambda_2 t} + \cdots \tag{49}$$

where A_m are to be determined by initial conditions. When $\lambda_1 \ll \lambda_m$ for $m > 1$, only the first exponential is important, with λ_1 giving the inverse of the relaxation time, and the folding kinetics appears as two-state.

When the dimension of the matrix of conversion rates is small (for M up to $\sim$1000), it can be diagonalized numerically. The $\boldsymbol{\omega}$ matrix is symmetrized with the transformation [66]

$$\omega'_{mm'} = [P_{\rm eq}(m)]^{-1/2} \omega_{mm'} [P_{\rm eq}(m')]^{1/2}. \tag{50a}$$

The $\boldsymbol{\omega}$ and $\boldsymbol{\omega}'$ matrices have identical eigenvalues, and the eigenvectors are related by

$$P_m(m') = P'_m(m') [P_{\rm eq}(m')]^{1/2}. \tag{50b}$$

The symmetric matrix $\boldsymbol{\omega}'$ can be diagonalized by, e.g., the Jacobi method [67]. For higher dimensions, the folding rate can be found by kinetic simulations (note that, in this context, Monte Carlo simulations are equivalent to numerical implementation of master equations [68]), or by approximation. Here the TST result is presented. Let us separate the microstates into two states, unfolded and folded. The former will be referred to by indices m_1 (running from 0 to $m^\ddagger$) and the latter by m_2 (running from $m^\ddagger + 1$ to M). The equilibrium fractions are

$$f_{\rm u} = \sum_{m_1=0}^{m^\ddagger} P_{\rm eq}(m_1)$$

$$f_{\rm f} = \sum_{m_2=m^\ddagger+1}^{M} P_{\rm eq}(m_2).$$

Only conversions from unfolded microstates to folded microstates contribute to the folding rate. Summing over all these conversions, we find

$$k_{\rm f;TST} = \frac{1}{f_{\rm u}} \sum_{m_2=m^\ddagger+1}^{M} \sum_{m_1=0}^{m^\ddagger} \omega_{m_2 m_1} P_{\rm eq}(m_1). \tag{51a}$$

Similarly for the unfolding rate we find

$$k_{\rm u;TST} = \frac{1}{f_{\rm f}} \sum_{m_1=0}^{m^\ddagger} \sum_{m_2=m^\ddagger+1}^{M} \omega_{m_1 m_2} P_{\rm eq}(m_2). \tag{51b}$$

These generalize equation (39), which is quite accurate for the Zwanzig model. Both Alm *et al* [48] and Weikl *et al* [50] also showed that TST provides good approximations for their models, suggesting equation (51) may be useful for protein folding modeled by the master equation.

During the folding process, the dynamics of different degrees of freedom may occur at very different time scales. Consider, for example, the elemental step of forming a native

contact. Even in the simplest description, this involves three degrees of freedom (figure 3(*C*)): two discrete variables δ_1 and δ_2 representing the conformations (with values 1 for native and 0 for disordered) of the two residues that eventually form the contact, and the distance r between the residues. Below, the rate of forming a native contact is calculated. This calculation suggests a dynamic reason why protein folding is fast, in addition to the energetic reason discussed earlier.

Suppose that, because of chain statistics, the end distance r of the loop connecting the two residues has a distribution $P(r)$, which for concreteness is set to what is given by equation (40). If a potential $\varepsilon_0(r)$, confined within $r < a$ and with a deep well near $r = a$, represents the energetic contribution from forming the contact, then the free-energy functional for contact formation is

$$G(\delta_1, \delta_2, r) = \delta_1\delta_2\varepsilon_0(r) - k_BT\{\ln[\mathcal{N}P(r)] + (1-\delta_1)\ln\nu_1 + (1-\delta_2)\ln\nu_2\}. \tag{52}$$

Let us first consider the case where the two residues are always in their native conformations (i.e., $\delta_1 = \delta_2 = 1$). Then contact formation can be modeled as a unimolecular reaction, with r as the reaction coordinate. The reactant well has a minimum at $r_m = (2\langle r^2\rangle/3)^{1/2}$, at which $P(r)$ is maximal, and the product well is just inside $r = a$. If the dynamics along the r coordinate is diffusive, with a diffusion constant D, then according to equation (5*b*) the rate constant for contact formation is [69]

$$k_f = \frac{1}{\int_a^\infty \mathrm{d}r\ P(r)\int_a^{r_m}\mathrm{d}q[DP(r)]^{-1}}. \tag{53a}$$

When $a \ll r_m$, equation (53*a*) leads to

$$k_f = 3(6/\pi)^{1/2}Da/\langle r^2\rangle^{3/2} \tag{53b}$$

a result first derived by Szabo *et al* [7].

Let us derive equation (53) by an alternative approach that is more amenable to generalization to more complicated situations. This is done by exploiting connections between the unimolecular reaction between two reactants connected by a flexible loop and the bimolecular reaction without the intervening loop [70, 71]. The equilibrium constants are related by

$$K = p(a)K_{bi}. \tag{54a}$$

The breakup of the contact is dictated by the dynamics within the deep well near $r = a$, and therefore the rates, k_u and k_d, in the unimolecular and bimolecular reactions, respectively, are expected to be identical:

$$k_u = k_d. \tag{54b}$$

These relations can easily be verified for contact formation. For example, the equilibrium constant for the bimolecular reaction is (equation (17))

$$K_{bi} = \int_0^a \mathrm{d}r\ 4\pi r^2\,\mathrm{e}^{-\beta\varepsilon_0(r)}.$$

And for the unimolecular reaction, one has (equation (5*c*))

$$K = \frac{\int_0^a \mathrm{d}r\ P(r)\,\mathrm{e}^{-\beta\varepsilon_0(r)}}{\int_a^\infty \mathrm{d}r\ P(r)} \approx \frac{\int_0^a \mathrm{d}r\ 4\pi r^2 p(r)\,\mathrm{e}^{-\beta\varepsilon_0(r)}}{\int_0^\infty \mathrm{d}r\,P(r)} \approx \int_0^a \mathrm{d}r\ 4\pi r^2 p(a)\,\mathrm{e}^{-\beta\varepsilon_0(r)} = p(a)K_{bi}.$$

From equations (54*a*) and (54*b*), we immediately arrive at

$$k_f = p(a)k_a \tag{54c}$$

for the contact-formation rates. In the present situation, $k_a = 4\pi Da$, and equation (54*c*) leads to the result in equation (53*b*).

A number of applications of equation (54*c*) are of interest. For example, so far we have assumed that the contact-forming residues are uniformly reactive on their surfaces. Results of k_a for non-uniform reactivity (e.g., equation (22*c*)) can be used to account for anisotropic contact formation. Equation (53c) is also valid when the loop is modeled as other types of polymer chains, such as a wormlike chain [72]. Another important extension is the situation where transitions between native and disordered conformations are allowed. If the transitions are modeled as stochastic gating between reactive and non-reactive species, the association rate constant k_{ag} can be worked out [21]. The result is

$$\frac{k_a}{k_{ag}} = 1 + \frac{\omega_{1-}}{\omega_{1+}g(\omega_1)} + \frac{\omega_{2-}}{\omega_{2+}g(\omega_2)} + \frac{\omega_{1-}\omega_{2-}}{\omega_{1+}\omega_{2+}g(\omega_1+\omega_2)} \tag{55a}$$

where ω_{1+} and ω_{1-} are the transition rates from and to the disordered conformations of residue 1, $\omega_1 = \omega_{1+} + \omega_{1-}$; ω_{2+}, ω_{2-}, and ω_2 are the corresponding quantities for residue 2; and $g(x) = 1 + (xa^2/D)^{1/2}$. Equation (53c) then gives the corresponding expression for k_{fg},

$$\frac{k_f}{k_{fg}} = 1 + \frac{\omega_{1-}}{\omega_{1+}g(\omega_1)} + \frac{\omega_{2-}}{\omega_{2+}g(\omega_2)} + \frac{\omega_{1-}\omega_{2-}}{\omega_{1+}\omega_{2+}g(\omega_1+\omega_2)} \tag{55b}$$

which is identical to the result obtained by directly accounting for conformational transitions in unimolecular contact formation [69]. We can now address the question of how much the additional requirement of correct local conformations will slow the rate of contact formation. Equation (55*b*) shows that the answer depends on the time scale of the local conformational fluctuations relative to a^2/D, the time scale for inter-residue diffusional motion. If the fluctuations are slow (i.e.,$g(\omega_1)$, $g(\omega_2)$ and $g(\omega_1+\omega_2) \to 1$), then $k_{fg}/k_f \to \omega_{1+}\omega_{2+}/\omega_1\omega_2 = 1/(1+\nu_1)(1+\nu_2)$, the equilibrium probability for simultaneous occurrence of the native conformations. For example, at $\nu_1 = \nu_2 = 15$ (corresponding to a free energy of 1.6 kcal mol^{-1} due to entropy loss per residue at room temperature), the rate would be reduced by 256-fold. However, when local conformational fluctuations are very fast ($g(\omega_1)$, $g(\omega_2)$ and $g(\omega_1+\omega_2) \to \infty$), k_{fg}/k_f actually approaches 1. In particular, if the local fluctuations occur on a picosecond time scale relative to a nanosecond time scale of inter-residue diffusion, then the rate of contact formation would be reduced by just 7-fold. The compensation of rapid local dynamics for low equilibrium probabilities in achieving a high rate is reminiscent of the situation of enzyme conformational gating in substrate binding to a buried active site [28], and suggests a dynamic reason why protein folding is fast.

In the contact formation model presented here (figure 3(*C*)), the residues can just be taken as secondary structure elements; then the model becomes an extension of the diffusion-collision model [73]. Though in isolation the stability of individual secondary structure elements is

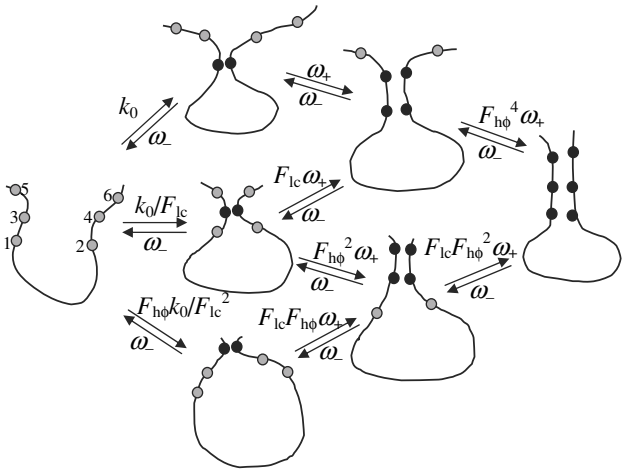

Figure 4. Model for formation of β-hairpin and coiled-coil. For the situation leading to equation (56), $F_{h\phi} = 1$ and $\omega_+/\omega_- > 1$. For the situation leading to equation (57), $F_{h\phi} > 1$ and $\omega_+/\omega_- < 1$. The parameters are chosen so $F_{h\phi}^4 k_0\omega_+^2/\omega_-^3 = 1$, to ensure equal equilibrium probabilities for the fully unfolded and folded microstates.

low, the fast time scales with which the local structures are formed and broken lead to a high probability of forming tertiary contacts upon collision. The contact formation model can provide physically sound transition rates used in master equation models for protein folding. Disparate time scales of local conformational fluctuations and global configurational changes may well play an active role in determining the folding rate.

3.2.1. Formation of β-hairpin and coiled-coil. As illustrated in figure 4, a β-hairpin or coiled-coil is envisioned to form by first forming a native contact and then by zipping up the neighboring contacts. For simplicity we will restrict the number of native contacts to three, designated as 1-2, 3-4 and 5-6, and assume that the unzipping of all contacts has the same rate, ω_-. We also exclude the possibility of skipping a contact, and therefore the 1-2 and 5-6 contacts cannot form simultaneously without forming the 3-4 contact.

First consider the situation where, once the initial contact has been made, the zipping of neighboring contacts is favorable. Let the rate of first forming the 1-2 contact be k_0 and the rates for first forming the 3-4 and 5-6 contacts be k_0/F_{lc} and k_0/F_{lc}^2, respectively, where F_{lc} (>1) accounts for the penalty on the closure of longer loops. Two-state kinetics demands $k_0/\omega_- \ll 1$. The zipping of subsequent contacts on the side of the free terminals is assumed to have the same rate ω_+ ($>\omega_-$), and the zipping of a contact inside the loop is assumed to have the rate $F_{lc}\omega_+$. By requiring $k_0\omega_+{}^2/\omega_-{}^3 = 1$, the microstates without any contact and with all three contacts are made to have the same equilibrium probabilities. By solving the master equation, it can be shown that, when $\omega_+/\omega_- \gg 1$, the folding rate is well approximated by

$$k_f = k_0 + k_0/F_{lc} + k_0/F_{lc}^2 \qquad (56)$$

which is the sum of the rates for forming the first contacts and is the TST result from equation (51*a*) by assuming that the microstate without any contact makes up the unfolded state.

Now consider the situation where formation of the 1-2 and 3-4 contacts are unfavorable (i.e., $\omega_+/\omega_- < 1$), but the formation of the 5-6 contact is favored (e.g., by an accompanying hydrophobic cluster). Suppose that this favorable interaction contributes a factor $F_{h\phi}$ each to the rates for first forming the 5-6 contact and for subsequently forming the 3-4 contact; a factor $F_{h\phi}^2$ each to the rates for forming the 5-6 contact subsequent to the 3-4 contact and for forming the 1-2 contact subsequent to the 3-4 and 5-6 contacts; and a factor $F_{h\phi}^4$ to the rate for forming the 5-6 contact subsequent to the 1-2 and 3-4 contacts. The equal equilibrium probabilities of the fully unfolded and folded microstates are ensured by requiring $F_{h\phi}^4 k_0\omega_+^2/\omega_-^3 = 1$. Again, by solving the master equation, it can be shown that the lesser of following expressions provides a close upper bound for the folding rate:

$$k_f = \frac{(\omega_+/\omega_-)k_0 + k_0/F_{lc} + F_{h\phi}k_0/F_{lp}^2}{1 + k_0/\omega_-} \qquad (57a)$$

$$k_f = \frac{(\omega_+/\omega_-)k_0 + k_0/F_{lc} + F_{h\phi}^2(\omega_+/\omega_-)k_0/F_{lp}}{1 + k_0/\omega_- + F_{h\phi}k_0/\omega_- F_{lc}}. \qquad (57b)$$

These results are the TST approximation (equation (51*a*)) with the border of the unfolded state specified according to the following *general* method. Routes from the fully unfolded microstate to the fully folded microstate are traced along the direction of increasing number of native contacts, and the local-equilibrium reaction flux, $w_{m'm} \equiv \omega_{m'm}P_{eq}(m)$, is evaluated. Along each route, the bottleneck step, with minimum reaction flux, is identified. If a bottleneck step of one route appears in another route with a more severe bottleneck (with a lower minimal reaction flux), the latter bottleneck is eliminated. The collection of bottleneck steps then constitutes the border between the unfolded and folded states. For equation (57*a*), the unfolded state consists of the microstate without any contact and that with the 1-2 contact. For equation (57*b*), the unfolded state has an additional microstate, that with the 5-6 contact.

In both of the cases leading to equations (56) and (57), all the routes are potentially available to form the β-hairpin. However, the dominant routes are different. In the former case, if the loop-closure penalty F_{lc} is large, the dominant folding route is to first form the 1-2 contact. (For smaller F_{lc}, the routes starting with the 3-4 or 5-6 contact also contribute significantly to the folding rate.) In the case of equation (57), if the rate enhancement factor $F_{h\phi}$ due to a hydrophobic cluster accompanying the 5-6 contact is large, then the dominant folding route is to first form the 5-6 contact. The interplay of loop entropy and hydrophobic contact in β-hairpin formation has been studied experimentally by Gai and co-workers [74]. For the β-hairpin formed by the 16 C-terminal residues of streptococcal protein G B1 [75], folding scenarios initiated from the β-turn and started from a hydrophobic cluster near the terminals have both been proposed. The former is by Muñoz *et al* [76] through a model with microstates defined by a single stretch of native residues; the latter is by Dinner *et al* [77] from Monte Carlo simulations of an atomistic model. It is heartening to see that both scenarios can be produced by the same model presented here, through changes of physically

motivated parameters. Further extensions may be able to qualitatively explain why the C-terminal β-hairpin in protein G B1 folds faster than the N-terminal β-hairpin, whereas the order is reversed in the structurally similar peptostreptococcal protein L B1 [78].

We now turn our attention briefly to the formation of coiled-coil. The model is the same, except now the native conformation of each residue is α-helical. According to equation (56), with a large penalty for loop closure, the first formation of a contact closest to the loop is the dominant folding route. If the loop is cleaved, then the folding becomes bimolecular and the penalty for loop closure is eliminated. The formation of all contacts has comparable rates, resulting in multiple competent folding routes. These predictions are just what were observed by Moran *et al* [79], who compared the folding of the dimeric GCN4 coiled-coil and a crosslinked single-chain variant.

3.2.2. Dependence of folding rate on solvent friction. A number of experimental studies have been carried out on the dependence of the folding rate on solvent friction [80–85]. The folding rate is found to be inversely proportional to the solvent friction ζ, as expected from the diffusional motion of the protein chain in forming the folded state (see, e.g., equation (53*b*) with equation (4), $D = k_BT/\zeta$). The extrapolation to zero solvent friction in these studies raises the question about the behavior of k_f in the low-friction regime.

To gain insight into that question, here a theory for the full dependence of the contact formation rate on solvent friction is presented. For simplicity, conformational fluctuations of residues are neglected. The starting point is the result of Bicout *et al* [86] that covers the entire friction range for the reaction rate of two spherical particles with a contact distance a, given by

$$\frac{1}{k_a} = \frac{1}{k_{coll}} + \frac{\zeta}{4\pi k_B T a} \tag{58a}$$

where $k_{coll} = 2^{1/2}\pi a^2\langle v\rangle$ is the rate constant for collisions between the particles [87], with $\langle v\rangle$ the mean velocity of the particles. The first term gives the rate in the zero-friction limit, where the particles undergo ballistic collisions, while the second term gives the rate in the high friction limit, where the particles undergo diffusive motion. Using the connection between unimolecular and bimolecular reaction rates (equation (54*c*)), we find

$$\frac{1}{k_f} = \frac{1}{p(a)k_{coll}} + \frac{\zeta}{4\pi k_B T a p(a)}. \tag{58b}$$

No theoretical result is available for the full friction dependence of the contact formation rate when conformational fluctuations of residues are present. We can speculate that the rate can still be predicted by an interpolation like equation (58*b*). In the zero-friction limit, the contact formation rate will be scaled by the equilibrium native fractions of the residues. When the loop is modeled as a Gaussian chain, the zero-friction limit of the contact formation rate is $k_{fg}(\zeta = 0) = p(a)k_{coll}/(1 + \nu_1)(1 + \nu_2) = 3a^2(3k_BT/\pi\mathcal{M})^{1/2}/2\langle r^2\rangle^{2/3}(1 + \nu_1)(1 + \nu_2)$, where $\mathcal{M}$ is the effective mass of the residues. With $a \sim 4$ Å, $\mathcal{M} \sim 1000$ amu, $\langle r^2\rangle^{1/2} \sim 50$ Å, $\nu_1 \sim \nu_2 \sim 10$, one finds $k_{fg}(\zeta = 0) \sim 10^6\ s^{-1}$, which is within the range of extrapolated values by Hagen and co-workers [83–85] from friction dependence of contact formation and folding rates.

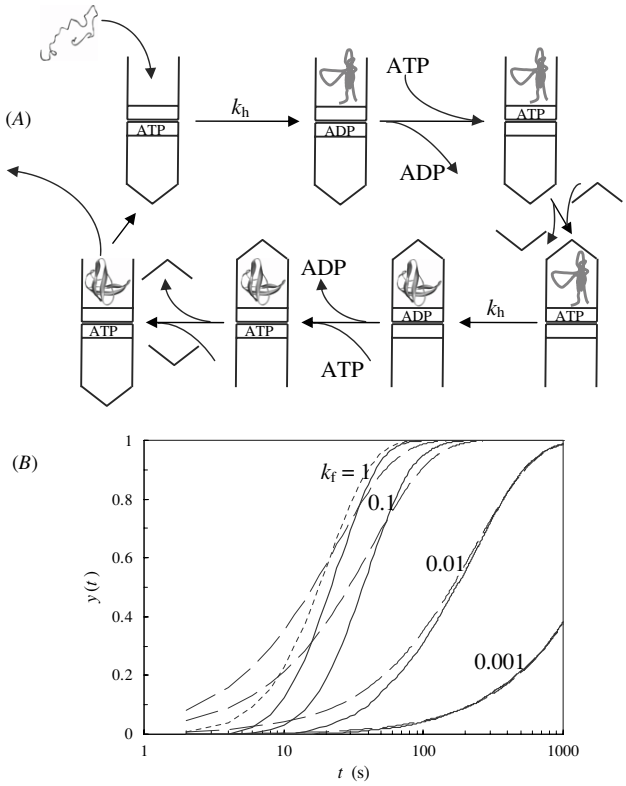

Figure 5. (*A*) Model for chaperonin-assisted protein folding. (*B*) Yield of folded protein, obtained by kinetic simulations (solid curves) and predicted as $1 - \exp(-k_o t)$ (dashed curves). Equation (59) is shown as dotted curve. Values of k_f in s^{-1} are shown. In the simulations, each folding cycle consists of three steps: binding of unfolded protein to *cis* ring of GroEL·ATP·GroES; *trans* ATP hydrolysis; and *cis* ATP hydrolysis. The lifetime of each step was generated from a random number $\mathcal{R}$ uniformly distributed between 0 and 1 as $-\tau\ln(\mathcal{R})$, where $1/\tau = k_1$, k_h and k_h, respectively. After the first cycle, geminate binding can be included. Specifically, a random number $\mathcal{R}$ is compared to the capture probability by geminate binding. If $\mathcal{R}$ is smaller, the lifetime is set to zero; otherwise it is generated randomly with a rate constant of k_1. In each cycle, whether folding has occurred in the *cis* ATP hydrolysis step (with a lifetime t_h) is determined by comparing a random number $\mathcal{R}$ with $1 - \exp(-k_f t_h)$. If $\mathcal{R}$ is smaller, folding has occurred and the total time preceding this event is recorded; otherwise the simulation is continued. The simulation is repeated 5000 times and the yield of folded protein at different times is obtained from the recorded times preceding the folding events. For the parameters selected, the effect of geminate rebinding is negligible.

3.2.3. Chaperonin-assisted folding. In the cell the folding of some proteins is difficult to occur spontaneously due to the tendency of unfolded proteins to aggregate, and is assisted by molecular chaperones. Of these the best studied is the chaperonin system in *Escherichia coli*, consisting of two interacting proteins GroEL and GroES. GroEL has a double-ring structure, each of which has the potential of binding and

hydrolyzing ATP, binding an unfolded polypeptide chain, and binding GroES to form an enclosed folding chamber. In the presence of ATP, the chaperonin-assisted folding of an unfolded protein goes through a cycle illustrated in figure 5(*A*) [88, 89]. ATP-binding to GroEL triggers a conformational change that leads to the tight binding of GroES. However, the two rings have negative cooperativity in binding ATP [90], which means that normally only one ring is ATP- and GroES-bound. The unfolded protein is recognized by the open *cis* ring of such an asymmetric GroEL·ATP·GroES, or LTS, complex. Hydrolysis of the *trans* ATP primes the release of the *trans* GroES [91] and allows the binding of ATP in the *cis* ring. The latter step triggers the release of the *trans* GroES and the binding of the *cis* GroES, which encloses the folding chamber. The encapsulated protein can fold until the hydrolysis of the *cis* ATP, which starts the same sequence of steps listed above with role reversal for the *cis* and *trans* rings, and the concurrent release of the encapsulated protein, either folded or unfolded. In addition, if the chaperonin system is in large excess over the unfolded protein, there is little chance that the *trans* ring can capture another unfolded protein chain before binding GroES. Thus at the end of the cycle, the original asymmetric LTS complex is re-formed. The encapsulated protein has undergone one round of folding, at the expense of ATP hydrolysis in both the *cis* and *trans* rings.

The above cycle of events bears significant similarity to the model of substrate channeling illustrated in figure 2(*A*). Indeed the first two cycles of the chaperonin-assisted folding model (right after the unfolded protein is mixed with excess LTS) reduce to the model of substrate channeling in the limit that ATP hydrolysis is infinitely fast and that the protein is folded at the end of the second folding cycle. With a finite ATP hydrolysis rate constant k_h and a finite folding rate constant k_f inside the folding chamber, the yield, $y(t)$, of folded protein can be obtained through kinetic simulations of the folding cycles. The results for $y(t)$ at $k_h = 0.1\ s^{-1}$, $k_1 = k_a[\mathrm{LTS}] = 1\ s^{-1}$, and four values of k_f are displayed in figure 5(*B*). The association rate constant of unfolded protein to the *cis* ring of LTS is assumed to be the same in different folding cycles. The effect of possible geminate rebinding, i.e., rebinding of a discharged unfolded chain to the same *cis* ring, is very small, since loading unfolded protein onto a *cis* ring is not rate-limiting.

When k_f is very large, only one cycle of folding is required; thus the yield of folded protein can be found by solving the following kinetic scheme:

$$\mathrm{LTS} \xrightarrow{k_1} \mathrm{pLTS} \xrightarrow{k_h} \mathrm{STpL} \xrightarrow{k_h} \mathrm{LTS} + \mathrm{p}.$$

The result is

$$y_\infty(t) = 1 - \left(\frac{(k_1 - k_h)^2 - k_h^2}{(k_1 - k_h)^2} + \frac{k_1 k_h t}{k_1 - k_h}\right) e^{-k_h t} - \frac{k_h^2}{(k_1 - k_h)^2} e^{-k_1 t}. \tag{59}$$

The mean lifetime of the first cycle can be calculated as the time integral of $y_\infty(t)$. The result, $1/k_1 + 2/k_h = (2 + k_h/k_1)/k_h$, gives the best-fitting value for the inverse of the overall folding rate constant k_o at $k_f \to \infty$ if equation (59) is fitted to a single exponential function $1 - \exp(-k_o t)$. Neglecting geminate rebinding, this is also the lifetime for each subsequent cycle. When k_f is very small, the average number of cycles required for yielding a folded protein is k_h/k_f. Then $1/k_o \approx (k_h/k_f)[(2 + k_h/k_1)/k_h] = (2 + k_h/k_1)/k_f$. From the limits at large and small k_f, the following interpolating formula may be proposed:

$$\frac{2 + k_h/k_1}{k_o} = \frac{1}{k_h} + \frac{1}{k_f}. \tag{60}$$

(Note the similarity in structure between equations (58) and (60)). Figure 5(*B*) shows that the time dependence of the yield of folded protein fits reasonably well to the single exponential function $1 - \exp(-k_o t)$.

With nucleotides (ATP and ADP), the protein substrate, and GroES potentially binding on its two rings, GroEL in principle can exist in 144 different species (e.g., L, LTS and pLTS), with a considerable number of possible transition pathways [92]. Yet, as illustrated in figure 5(*A*), the chaperonin-assisted folding appears to occur in an orderly fashion, traversing a small number of GroEL species. This situation is reminiscent of the dominant routes of protein folding discussed earlier. Two major contributing factors to the order here are binding-induced conformational changes (such as induced by ATP binding) and preferential binding (e.g., GroES for ATP-bound GroEL over the nucleotide-free species).

Besides the obvious role of isolating unfolded protein chains to allow for their individual folding (thus preventing aggregation), does the chaperonin system play more active roles in promoting the folding of a protein? Multiple hydrophobic sites around the opening of the GroEL ring provide high affinity for unfolded protein chains [93]. Upon ATP binding and subsequent GroES binding, GroEL undergoes large conformational changes, and the bound protein chain is released to the folding chamber with a surface now lined with hydrophilic residues [94, 95]. The new hydrophilic surface may guide the protein substrate toward the folded state. In addition, it has been suggested that the confinement within the folding chamber eliminates some of the open conformations that are available to the protein if it is in bulk solution, thus reducing the conformational space of the unfolded state [96].

When $k_f < k_h$, multiple cycles of folding are required to obtain the folded protein. Is there any advantage to releasing a protein from the folding chamber before folding is completed? If the encapsulated protein is kinetically trapped, releasing it into the bulk solution gives it a chance to re-equilibrate and move out of the trapped state. Such a pathway may be blocked in the confined space of the folding chamber.

3.3. Protein binding

When two proteins form a complex, the stereospecificity imposes a significant constraint on the association rate. For example, when the proteins are modeled as spheres with respective small reactive patches, the association rate constant is given by equation (22*c*), which will now be denoted as k_{a0}. The extra '0' in the subscript indicates that the result is for the case where interactions between the proteins are absent. This basal rate amounts to 10^5–$10^6\ M^{-1}s^{-1}$ for

patches sizes ∼5°. Much higher rates have been observed for associations, e.g., between barnase and barstar [97], between lac repressor and its DNA operator site [98], between unfolded proteins and GroEL [99–103], and between GroES and ATP-bound GroEL [103, 104]. Two proposed rate-enhancement mechanisms, electrostatic attraction between a protein and its target and nonspecific binding with the target, appear to have firm experimental support. A third mechanism has been proposed for proteins that are unfolded before binding [105, 106]. Whether an unfolded protein chain affords rate enhancement that is not accessible to a folded protein is still a matter of investigation.

We focus on diffusion-controlled association, for which the association rate is determined by absorption at the outer boundary of the bound state (e.g., as represented by $r^{\ddagger}$ in figure 1(*B*)). It has been shown that the rate enhancement by long-range electrostatic attraction in the unbound state can be estimated from a simple formula [107]:

$$k_{\mathrm{a}} = k_{\mathrm{a0}} \langle \mathrm{e}^{-\beta \Delta G_{\mathrm{el}}} \rangle^{\ddagger} \tag{61a}$$

where ΔG_{el} is the free energy of electrostatic interactions between the associating proteins, and $\langle \cdots \rangle^{\ddagger}$ signifies an average over the outer boundary of the bound state. In particular, this formula has been used to quantitatively rationalize the effects of charge mutations on the association rate of barnase and barstar [108].

Equation (61*a*) also offers a mechanistic interpretation. The outer boundary may be viewed as the transition state for association [11, 108]. Note that the concept of transition state originated on unimolecular reactions, but association is a bimolecular process. However, a strong argument can be made for the adoption of this concept here. As two proteins approach each other, their translational and rotational degrees of freedom become restricted. Therefore they suffer a loss in translation and rotational entropy. On the other hand, the stabilization arising from specific interactions between the residues of the two proteins occurs very late, when the two proteins are aligned and in close contact (i.e., nearly as in the final bound structure). Therefore, like the situation in protein folding (see figure 3(*A*)), the proteins apparently encounter a free-energy barrier (with a significant translation/rotational entropy component). This occurs around the outer boundary of the bound state.

With the transition state for association thus identified, its electrostatic free energy, $\Delta G^{\ddagger}_{\mathrm{el}}$, can be defined as (cf equation (44))

$$\mathrm{e}^{-\beta \Delta G^{\ddagger}_{\mathrm{el}}} = \langle \mathrm{e}^{-\beta \Delta G_{\mathrm{el}}} \rangle^{\ddagger}. \tag{61b}$$

Equation (61*a*) then becomes

$$k_{\mathrm{a}} = k_{\mathrm{a0}}\, \mathrm{e}^{-\beta \Delta G^{\ddagger}_{\mathrm{el}}} \tag{61c}$$

which has a straightforward interpretation: electrostatic attraction between the two proteins enhances the association rate simply by lowering the free-energy barrier. Equation (61*c*) also provides a rationalization for a general phenomenon regarding the effects of ionic strength on the kinetics of protein–protein association [11, 109]. The association and dissociation rate constants have been observed to have disparate dependences on ionic strength, with the former showing strong sensitivity and the latter showing weak dependence. Relative to the unbound state (where the two proteins are infinitely apart), the electrostatic attraction in the transition state is screened increasingly, leading to a significant decrease in k_{a} at higher ionic strength. On the other hand, given the closeness of the transition state to the bound state, the electrostatic attraction in these two states may be assumed to be screened to similar extents. Thus, in protein association, ionic strength plays a role that is similar to a mutation with a Φ-value close to 1 in protein folding.

Two caveats should be added to the preceding transition-state view of electrostatic rate enhancement. First, equation (61*a*) is above all a mathematical approximation, with limitations. It was derived specifically for the situation where the binding is specific and the interaction is long-ranged [110]. These conditions happen to be met for electrostatic enhancement of the protein–protein association rate. For example, equation (61*a*) is a poor approximation for two uniformly reactive spheres, as can be checked by the Debye formula (equation 18)). Second, there are important differences between the transition state for the present bimolecular process and those for unimolecular reactions. In particular, the probabilities of the proteins starting from the present transition state to reach the bound state and the unbound state first are far from equal. In fact, the proteins on the outer boundary of the bound state will almost certainly first fall into the deep well in the bound state as opposed to escaping to infinite separation. For example, in equation (32*a*), $\gamma_{\mathrm{gem}}(\theta) \to 1$ as $\theta \to \alpha$; similarly, in equation (33), $\gamma_{\mathrm{gem}}(d) \to 1$ as $d \to 0$. More importantly, unlike in unimolecular reactions, where γ_{prd} has a sharp transition around the transition state, from 1 in the product state to 0 in the reactant state, here the capture probability γ_{gem} by the bound state only gradually decreases to 0, when the proteins become infinitely apart.

Protein binding to a specific site on DNA presents a different scenario for rate enhancement. It was appreciated early on by Adam and Delbruck [111] and by Richter and Eigen [112] that nonspecific binding to the full surface of the target (e.g., the entire length of the DNA) reduces the dimensionality of the search space for the specific site. This idea has subsequently been developed by many others. In particular, Berg *et al* [113] described different modes of nonspecific-binding-facilitated translocation along the DNA (such as sliding and hopping). Berg and Ehrenberg [114] empirically incorporated surface diffusion into the Smoluchowski theory [14] for diffusion-influenced reactions. Recently a more fundamental and realistic approach, in which nonspecific binding is accounted for by a short-range attractive potential around the target surface, was introduced [115]. This approach can be cast into the nonspecific-binding-facilitated diffusion model (see figure 6(*A*)) [116]. In this model, a DNA-targeting protein can diffuse to the DNA surface and nonspecifically bind to it. Once on the DNA surface it will either move along the DNA or equilibrate with the surrounding region, and the process repeats itself. This model encompasses all the various putative modes (e.g., sliding and hopping) of translocation from one site on the DNA to another site.

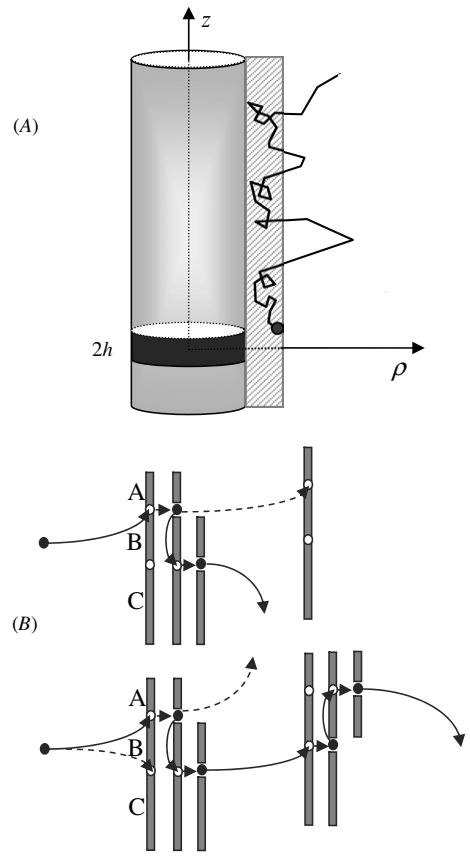

Figure 6. (*A*) The nonspecific-binding-facilitated diffusion model. The black strip with height $2h$ on the DNA surface is the specific binding site; a potential is present in the hatched region just outside the DNA surface. (*B*) Cleavage of a DNA substrate with two specific sites (white circles) by a processive restriction enzyme (black circle). Top: bimolecular binding occurs only at the first site. The second site can be subsequently cleaved by geminate binding. Bottom: bimolecular binding occurs at both sites.

Consider a DNA modeled as a cylinder (with radius a). A potential, $U(\rho)$, that depends on the distance ρ to the DNA axis describes the interaction between a protein and its DNA target. At steady state, the probability density for the DNA-binding protein satisfies

$$\begin{aligned}\mathcal{L}P(\rho,z) &\equiv \nabla\cdot\mathcal{D}\,\mathrm{e}^{-\beta U(\rho)}\nabla\,\mathrm{e}^{\beta U(\rho)}P(\rho,z)\\ &= \frac{D_\perp}{\rho}\frac{\partial}{\partial\rho}\rho\,\mathrm{e}^{-\beta U(\rho)}\frac{\partial}{\partial\rho}\,\mathrm{e}^{\beta U(\rho)}C(\rho,z)\\ &\quad + D_{/\!/}\frac{\partial^2}{\partial z^2}C(\rho,z)=0 \qquad (62)\end{aligned}$$

where $D_\perp$ and $D_{/\!/}$ are the diffusion constants perpendicular and parallel, respectively, to the DNA axis, and z is the coordinate along the axis. The specific binding site is approximated as an absorbing patch located in $|z|<h$ at $\rho=a$. Elsewhere the DNA surface is reflecting. The inner boundary condition is thus

$$P(a,z)=0 \qquad \text{for} \quad |z|<h \qquad (63a)$$

$$j_\perp(z)\equiv D_\perp\mathrm{e}^{-\beta U(\rho)}\frac{\partial}{\partial\rho}\,\mathrm{e}^{\beta U(\rho)}P(\rho,z)\big|_{\rho=a}=0 \quad \text{for } |z|>h. \qquad (63b)$$

For the purpose of calculating the association rate constant k_a, the probability density at infinity is set to 1. In the Smoluchowski theory, k_a is given as the total flux to the absorbing patch:

$$k_\mathrm{a}=2\pi a\int_{-h}^{h}\mathrm{d}z\,j_\perp(z). \qquad (64)$$

Suppose that the potential is restricted to a short range (between $\rho=a$ and a_1) around the DNA surface. The gist of the nonspecific-binding-facilitated diffusion model is to assume equilibration between the surface layer and the bulk phase (where diffusion is taken to be isotropic, i.e., $D_\perp=D_{/\!/}=D$). The probability per unit surface area for finding the protein in the surface layer is

$$u(z)=\int_a^{a_1}\mathrm{d}\rho\,\rho P(\rho,z)/a. \qquad (65a)$$

The equilibrium constant per unit surface area for nonspecific binding is

$$K_\mathrm{ns}=\int_a^{\infty}\mathrm{d}\rho\,\rho[\mathrm{e}^{-\beta U(\rho)}-1]/a. \qquad (65b)$$

Equilibration between surface layer and bulk phase means

$$u(z)=K_\mathrm{ns}P(a_1,z). \qquad (65c)$$

The last equation can be viewed as an implicit boundary condition for the bulk phase, where the protein undergoes isotropic free diffusion. To make the boundary condition more explicit, one finds by integrating equation (62) over ρ from a to a_1:

$$j_\perp(z)=D\frac{\partial}{\partial\rho}\,P(\rho,z)|_{\rho=a_1}+D_{/\!/}\frac{\partial^2}{\partial z^2}u(z) \qquad (66a)$$

$$=0 \qquad \text{for} \quad |z|>h. \qquad (66b)$$

The second step makes use of equation (63*b*). On the reactive patch, one has

$$P(a_1,z)\approx P(a,z)=0 \qquad \text{for} \quad |z|<h. \qquad (66c)$$

It is clear that all the effect of the surface potential now appears through the nonspecific binding constant K_ns. For an infinite DNA, the association rate constant is [115, 116]

$$k_\mathrm{a}^\infty=2\pi^2 DaF^\infty \qquad (67a)$$

with

$$\frac{1}{F^\infty}=\int_0^\infty\mathrm{d}x\,\frac{[\sin(hx/a)/(hx/a)]^2}{xK_1(x)/K_0(x)+x^2D_{/\!/}K_\mathrm{ns}/Da} \qquad (67b)$$

where $K_0(x)$ and $K_1(x)$ are the modified Bessel functions.

A protein that nonspecifically binds to DNA may appear processive, i.e., interacting with two specific sites on the same DNA. Note that processivity here does not imply directionality. It is similar in spirit to substrate channeling in enzyme catalysis, and again can be analyzed through the capture probability, $\gamma_\mathrm{gem}(z)$, the probability that a protein starting on the DNA surface ($\rho=a_1$) at z will be captured by the specific site at $z=0$. The complement, $1-\gamma_\mathrm{gem}(z)$, is the probability $\psi_\mathrm{gem}(z)$ for escaping to infinity, which in turn is identical to $P(a_1,z)$ for the present situation (where the potential is

confined to the surface layer). For an infinitely long DNA, the capture probability is [116]

$$\gamma^{\infty}_{\rm gem}(z) = F^{\infty}\int_0^{\infty} {\rm d}x\,\frac{\cos(xz/a)\sin(hx/a)/(hx/a)}{xK_1(x)/K_0(x) + x^2 D_{//}K_{\rm ns}/Da}. \tag{67c}$$

Corresponding results for finite DNA have been used to rationalize experimental data [117, 118] for the z-dependence of the capture probability [116]. More details can be found later.

When two specific binding sites are present, the association rate constant will be increased over a single site. If the two sites are far apart, the overall rate will be the sum of the rates of binding to the individual sites (with the other site treated as reflecting); if the sites become closer, interference sets in and the overall rate is much less than the sum. An interesting question is which site is reached first. The 'preferential' capture probability, $\gamma_1(\mathbf{r})$, for first reaching the first site starting from position $\mathbf{r}$ satisfies the steady-state backward Smoluchowski equation:

$$\mathcal{L}^{\dagger}\gamma_1(\mathbf{r}) = 0 \tag{68a}$$

with boundary conditions

$$\gamma_1(\mathbf{r}) = 1 \qquad \text{on site 1} \tag{68b}$$
$$= 0 \qquad \text{on site 2} \tag{68c}$$
$$= 0 \qquad \text{at infinity.} \tag{68d}$$

The flux of $\gamma_1(\mathbf{r})$ over site 1 then gives the preferential binding rate $k_{\rm a1}$ (cf equation (64)). Note that $k_{\rm a1}$ specified in this way is different from the rate constant for binding to site 1 alone, when site 2 is treated as reflecting. With preferential binding rate $k_{\rm a2}$ analogously specified for site 2, the overall binding rate is rigorously given by $k_{\rm a1} + k_{\rm a2}$.

3.3.1. Processive restriction enzymes. Restriction enzymes such as EcoRI and EcoRV cleave DNA at specific sites. When two such sites are present on a DNA substrate, the rates at which different fragments are produced contain information about the processivity of the enzyme (figure 6(*B*)) [117, 118]. To simplify the situation, suppose that the substrate concentration [S] is in excess of the enzyme concentration [E] and the experiment is limited to short times. These conditions ensure that [S] can be taken as time-independent and that cleavage at the second site, after the first site has already been cleaved, occurs only as a geminate or processive event. To connect with the model for substrate channeling shown in figure 2(*A*), let us first assume $k_{\rm a2} \to 0$. Four different fragments, A, BC, B and C, will be produced. Their rates of production are:

$$\frac{{\rm d[A]}}{{\rm d}t} = \frac{{\rm d[BC]}}{{\rm d}t} = w_1 \tag{69a}$$
$$\frac{{\rm d[B]}}{{\rm d}t} = \frac{{\rm d[C]}}{{\rm d}t} = \gamma_{1\to 2}w_1 \tag{69b}$$

where $w_1 = k_{\rm a1}[{\rm E}][{\rm S}]$, and $\lambda_{1\to 2}$ is the capture probability by site 2 starting from site 1. Equations (69*a*) and (69*b*) are the $t \to 0$ limits of equations (27*b*), and (31*b*), respectively, for the production of I* and P in the model for substrate channeling. When bimolecular binding to the second site is allowed, five fragments will be generated: A, AB, BC, B and C. Their rates of production are now:

$$\frac{{\rm d[A]}}{{\rm d}t} = w_1 + \gamma_{2\to 1}w_2 \tag{70a}$$
$$\frac{{\rm d[B]}}{{\rm d}t} = \gamma_{1\to 2}w_1 + \gamma_{2\to 1}w_2 \tag{70b}$$
$$\frac{{\rm d[C]}}{{\rm d}t} = w_2 + \gamma_{1\to 2}w_1 \tag{70c}$$
$$\frac{{\rm d[AB]}}{{\rm d}t} = w_2 \tag{70d}$$
$$\frac{{\rm d[BC]}}{{\rm d}t} = w_1 \tag{70e}$$

where $w_2 = k_{\rm a2}[{\rm E}][{\rm S}]$, and $\gamma_{2\to 1}$ is the capture probability by site 1 starting from site 1.

Terry *et al* [117] and Stanford *et al* [118] measured the initial rates of the various fragments cleaved by EcoRI and EcoRV, respectively. Two quantities derived from these rates are particularly interesting. The first is the processivity ratio:

$$F_{\rm p} = \frac{{\rm d[A]}/{\rm d}t + {\rm d[C]}/{\rm d}t - {\rm d[AB]}/{\rm d}t - {\rm d[BC]}/{\rm d}t}{{\rm d[A]}/{\rm d}t + {\rm d[C]}/{\rm d}t + {\rm d[AB]}/{\rm d}t + {\rm d[BC]}/{\rm d}t}. \tag{71a}$$

According to equation (70), we have

$$F_{\rm p} = \frac{\gamma_{1\to 2}w_1 + \gamma_{2\to 1}w_2}{2(w_1 + w_2) + \gamma_{1\to 2}w_1 + \gamma_{2\to 1}w_2} \tag{71b}$$

which reduces to

$$F_{\rm p} = \frac{\gamma_{1\to 2}}{2 + \gamma_{1\to 2}} \tag{71c}$$

when $w_1 = w_2$ and $\gamma_{1\to 2} = \gamma_{2\to 1}$. The experimental result for $F_{\rm p}$ was previously compared to $\eta\gamma_{1\to 2}$, with η an adjustable parameter with value close to 0.5. Equation (71*c*) shows that such an adjustable parameter is really unnecessary.

The second quantity of interest is the preference for site 2 over site 1, defined as

$$\frac{E_2}{E_1} = \frac{{\rm d[AB]}/{\rm d}t}{{\rm d[BC]}/{\rm d}t} \tag{72a}$$

which according to equation (70) is

$$\frac{E_2}{E_1} = \frac{w_2}{w_1} = \frac{k_{\rm a2}}{k_{\rm a1}}. \tag{72b}$$

The preference quantity E_2/E_1 is thus a direct measure of the relative magnitudes of the preferential binding rates $k_{\rm a1}$ and $k_{\rm a2}$.

3.3.2. Folding upon binding. For a protein that is unfolded when dissociated from its binding target, the binding process is accompanied by folding. That process may possibly be separated into an initial folding step followed by the binding step of the folded protein. In this scenario the overall binding rate will be smaller than the binding rate of the folded protein. A different scenario is that the binding site on the target has multiple subsites, each of which may serve as an initial contact point with a particular region of the protein chain. Once such a contact is formed, the protein chain can condense around it and form the bound and folded state. In this scenario, the overall binding rate can be written as

$$k_{\mathrm{a}} = \sum_{\text{subsites}} k_{\mathrm{a}i} F_i \tag{73}$$

where $k_{\mathrm{a}i}$ are the rates of binding to the subsites and F_i are reduction factors accounting for the subsequent condensation step. If F_i is close to 1, then the association rate can be large because each subsite serves as a competent binding route. Such a situation is already noted in the earlier discussion of coiled-coil formation. The multiple hydrophobic sites around the opening of the GroEL ring may not only provide high affinity for unfolded protein chains [93] but may also contribute to the high association rates observed [99–103].

To gain some insight into the reduction factor F_i, consider a condensation step represented by the formation of an additional contact that is a distance d away from the first subsite [106]. Let $k_{\mathrm{d}1}$ be the dissociation rate constant from the first subsite and $k_{\mathrm{f}2}$ be the rate constant for forming the contact when the first subsite is bound. Then

$$F_1 = \frac{k_{\mathrm{f}2}}{k_{\mathrm{d}1} + k_{\mathrm{f}2}}. \tag{74a}$$

If $k_{\mathrm{a}2}$ is the rate constant for forming the contact when the first subsite is not bound, then $k_{\mathrm{f}2} = p(d)k_{\mathrm{a}2}$ according to equation (54*c*). Therefore

$$F_1 = \frac{p(d)k_{\mathrm{a}2}/k_{\mathrm{d}1}}{p(d)k_{\mathrm{a}2}/k_{\mathrm{d}1} + 1}. \tag{74b}$$

A different view of potential rate enhancement in folding upon binding is proposed by Shoemaker *et al* [105]. In this view, the unfolded chain is more open and able to sample different conformations, and therefore may interact with the binding surface on the target at a greater distance than the folded protein. A free energy functional for the binding process was constructed by adding a binding term to their functional for protein folding [42]. The potential of mean force for a given protein–target distance was obtained, and then used in the Debye formula (equation (18)) to find the association rate.

3.4. Directional movement and rotation

Motor proteins use chemical energy to drive directional movement or rotation. The chemical energy may derive from the catalytic turnover of a nucleotide (e.g., in myosin, kinesin, helicases, and the F_1 motor of ATP synthase) or from the translocation of an ion such as proton across an electrochemical gradient (e.g., in the F_0 motor of ATP synthase and the bacterial flagellar motor) [119]. In the examples listed, the first three motors move along a linear track (actin filament, microtubule and DNA, respectively), whereas the last three are rotary. Many aspects of motor proteins have been elucidated through various models [119–131]. The focus here is the most basic question regarding all motor proteins: how is the chemical energy supply coupled to the mechanical motion?

In one of the earlier models that explicitly illustrated the coupling of ATP hydrolysis and directional movement, the motor protein is envisioned to cycle through three species: (1) nucleotide-free, (2) ATP-bound, and (3) ADP-bound (figure 7(*A*)) [121]. The motor protein itself is bound to a one-dimensional track, and undergoes Brownian motion. The

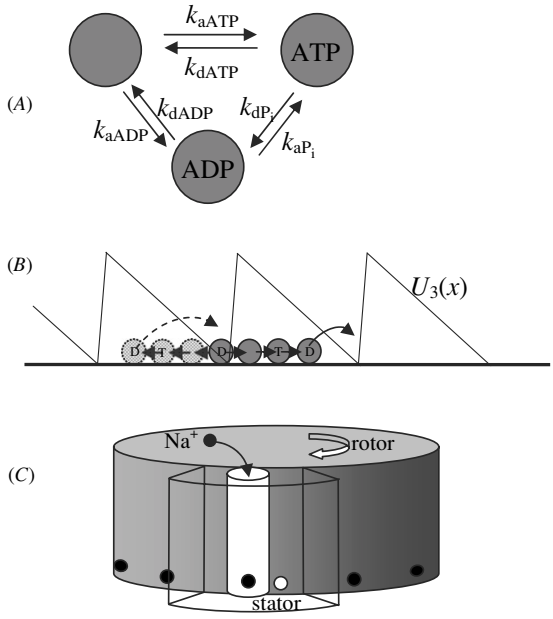

Figure 7. (*A*) Cycle of a motor protein through three species. (*B*) Movement of the motor protein along the track, which also interacts with the ADP-bound species through an asymmetric potential $U_3(x)$. In the depicted situation, $\mathcal{E} > 1$ so the conversions among the three species follow the direction ADP-bound (circle bearing 'D') → nucleotide free (circle without letter) → ATP-bound (circle bearing 'T') → ADP-bound. The ADP-bound species is favored at the bottom of the potential $U_3(x)$. After converting to the nucleotide-free species or subsequently to the ATP-bound species, the motor protein is free to move. When it moves to the left, a longer distance must be traversed before arriving at the potential well to the left, and if conversion to the ADP-bound species occurs before that, it will be driven back to the same well. On the other hand, when the motor protein moves to the right, it only needs to cross a short distance to get to the next well and will be trapped there upon converting to the ADP-bound species. The end result is net movement to the right. (*C*) Model for the F_0 motor, adapted from Xing *et al* [131]. A sodium ion (black circle) in the periplasm is driven down the channel to the binding site (hatched circle) by an electrochemical gradient. Once bound it is repelled to the left by the positive stator charge (white circle). After moving outside the stator, the sodium ion is released to the cytoplasm. The rotor is then free to rotate, bringing another binding site to the bottom of the channel.

Brownian motion is force-free for two of the three species (i.e., $U_1(x) = U_2(x) = 0$), but for one, the ADP-bound species, the motor protein experiences a periodic asymmetric potential $U_3(x)$, arising from interactions with the track (figure 7(*B*)). If the motor protein is restricted to just one species, then at long times it will be distributed according to the probability density,

$$P_{\mathrm{eq}i}(x) = N_i\, \mathrm{e}^{-\beta U_i(x)}, \qquad i = 1, 2 \text{ or } 3 \tag{75}$$

and there is no net motion. Now suppose that the three species are allowed to interconvert with rate constants $\omega_{i\to i'}(x)$. The probability density $P_i(x, t)$ for finding species i at x satisfies the diffusion-reaction equation

$$\frac{\partial P_i(x,t)}{\partial t} = -\frac{\partial j_i(x,t)}{\partial x} + \sum_{i'\neq i} \omega_{i'\to i}(x) P_{i'}(x,t) - \sum_{i'\neq i} \omega_{i\to i'}(x) P_i(x,t) \tag{76a}$$

with

$$j_i(x,t) = -D\,\mathrm{e}^{-\beta U_i(x)}\frac{\partial}{\partial x}[\mathrm{e}^{\beta U_i(x)}P_i(x,t)]. \qquad (76b)$$

At steady state, the total flux,

$$j = \sum_i j_i(x) \qquad (77)$$

is x-independent and can be shown to be the long-time velocity of the motor protein. The directional movement allows the motor protein to carry a load. A drag force W to the left can be accounted for by adding a potential energy Wx to each species.

If the conversion rates are such that detailed balance,

$$\omega_{i\to i'}(x)P_{\mathrm{eq}i}(x) = \omega_{i'\to i}(x)P_{\mathrm{eq}i'}(x), \qquad i \neq i' \qquad (78)$$

is satisfied, then the steady-state solution is just the equilibrium distribution and the velocity is zero. To gain further insight into the driving force for the directional movement in the model, let us calculate the quantity

$$\mathcal{E}(x) \equiv \frac{\omega_{1\to2}(x)\omega_{2\to3}(x)\omega_{3\to1}(x)}{\omega_{2\to1}(x)\omega_{3\to2}(x)\omega_{1\to3}(x)}. \qquad (79a)$$

The value of $\mathcal{E}(x)$ is 1 under the condition of detailed balance. Let us express the conversion rates in terms of the association and dissociation rate constants and the ATP, ADP, and inorganic phosphate (P_i) concentrations. For notational convenience, explicit reference to x is now dropped. The result is

$$\begin{aligned}\mathcal{E} &= \frac{k_{\mathrm{aATP}}[\mathrm{ATP}]k_{\mathrm{dP_i}}k_{\mathrm{dADP}}}{k_{\mathrm{dATP}}k_{\mathrm{aP_i}}[\mathrm{P_i}]k_{\mathrm{aADP}}[\mathrm{ADP}]} \\ &= \frac{k_{\mathrm{aATP}}k_{\mathrm{dP_i}}k_{\mathrm{dADP}}}{k_{\mathrm{dATP}}k_{\mathrm{aP_i}}k_{\mathrm{aADP}}}\frac{[\mathrm{ATP}]}{[\mathrm{P_i}][\mathrm{ADP}]}.\end{aligned} \qquad (79b)$$

The first factor involving the rate constants can be shown to be the inverse of the equilibrium constant K_{eq} for forming ATP from ADP and P_i, thus

$$\mathcal{E} = \frac{1}{K_{\mathrm{eq}}}\frac{[\mathrm{ATP}]}{[\mathrm{P_i}][\mathrm{ADP}]} \qquad (79c)$$

and $\mathcal{E} = 1$ means that the concentrations of ATP, ADP and P_i are maintained at chemical equilibrium. If they are out of chemical equilibrium, then detailed balance is broken and the velocity j of the motor protein will be nonzero. The foregoing analysis makes it clear that the chemical energy supply for the directional movement is the nonequilibrium of ATP, ADP and P_i. It is this nonequilibrium that steers the motor protein through uni-directional cycles of the three species, which together with the asymmetric potential experienced by one of the species consequently lead the motor protein in one direction (figure 7(*B*)). Like actual motor proteins, the direction of motion in the model can be reversed by changing the direction of the catalytic cycle, via changing the relative concentrations of ATP and ADP (e.g., from $\mathcal{E} > 1$ to $\mathcal{E} < 1$).

A very similar approach can be used to understand the basic principle of rotary proteins. For example, Oster and co-workers [131] constructed a mathematical model, based on structural and functional results, for the sodium-driven F_0 motor of *Propionigenium modestum*. In the simplest version (figure 7(*C*)), the rotation angle is the x coordinate, and only two species are present: the rotor's binding site within the channel is either bound or not bound by a sodium ion. In the unbound state, the rotor can move either to the left or right. However, repulsion between the bound ion and the stator charge prevents the bound species from moving to the right. The chemical nonequilibrium between the bound and unbound species arises from the electrochemical gradient, which brings an excess of sodium ions to the binding site. The ions otherwise would not be favored there due to the stator charge.

These models involve rectification (afforded by the asymmetric potential in figure 7(*B*) and the stator charge in figure 7(*C*)), and have been referred to as Brownian ratchets. In another common mechanism, ATP hydrolysis guides the motor protein through an ordered cycle, much like what happens in chaperonin-assisted protein folding (figure 5(*A*)), and directional motion is tied to a large conformational change (i.e., power stroke) at some point along the ATP hydrolysis process. It may also be possible to model this mechanism by a diffusion-reaction equation. The general principle appears to be that directional movement is generated by capturing favorable Brownian motion through intermolecular binding energy [127].

3.5. Conduction of ions through membrane channels

Ion permeation through selective membrane channels is essential for maintaining proper ionic balance across cell membranes, for transmission of neural signals, and for mechanochemical transduction of motor proteins such as the F_0 motor. The hydrophobic environment of the membranes gives rise to a large energy barrier [132]. The fundamental question is how, despite the dielectric barrier, do channels maintain rapid flux *and* high selectivity for specific ions [133]? Much insight has been gained recently, triggered by structure determinations of channel proteins [134–139].

The simplest model for an ion channel is a cylindrical pore (with radius a) through the membrane (figure 8) [140]. The dielectric barrier is described as a potential of mean force, $U(x)$, which goes to zero at the ends of the pore (at $x = 0$ and L). An external voltage V when applied across the membrane contributes a term $eV(1 - x/L)$ to the potential (e: charge on the ion). Let $U'(x) = U(x) + eV(1 - x/L)$. If the motion of ions within the pore is diffusive (with diffusion constant $D_{/\!/}$ along the pore) and the ions are assumed not to interact with each other, then the concentration $c(x, t)$ per unit length along the pore satisfies the Smoluchowski equation

$$\frac{\partial c(x,t)}{\partial t} = -\frac{\partial j(x,t)}{\partial x} \qquad (80a)$$

with flux

$$j(x,t) = -D_{/\!/}\,\mathrm{e}^{-\beta U'(x)}\frac{\partial}{\partial x}[\mathrm{e}^{\beta U'(x)}c(x,t)]. \qquad (80b)$$

Suppose that the bulk concentrations of the ion are C_1 and C_2 on the two sides, and steady state has been reached. At the ends of the pore, the values of $c(x)$ are given by

$$c(0) = \sigma C_1, \qquad c(L) = \sigma C_2 \qquad (81a)$$

where $\sigma = \pi a^2$ is the area of cross section of the pore. The constant flux is then

$$j = D_{/\!/}\sigma H(C_1\,\mathrm{e}^{\beta eV} - C_2) \qquad (81b)$$

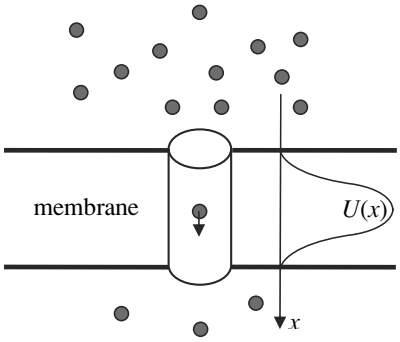

Figure 8. Model of an ion channel.

with

$$1/H = \int_0^L \mathrm{d}x\, \mathrm{e}^{\beta U'(x)}. \tag{81c}$$

The assumption that ions inside the pore do not interact with each other is very crude (see below). A better assumption is perhaps the opposite situation, where at most one ion can occupy the pore. This situation can be accounted for by modifying the boundary conditions of equation (81*a*) [140]:

$$c(0) = \sigma P_0 C_1, \qquad c(L) = \sigma P_0 C_2 \tag{81d}$$

where P_0 is the probability that the pore is empty. Now

$$j = D_{/\!/}\sigma P_0 H(C_1 \mathrm{e}^{\beta eV} - C_2). \tag{81e}$$

By requiring $P_0 = 1 - \int_0^L \mathrm{d}x\, c(x)$ and determining P_0 self-consistently, one finds

$$1/P_0 = 1 + \sigma C_1 \mathrm{e}^{\beta eV} H_1 + \sigma C_2 H_2 \tag{82a}$$

where

$$H_1 = \frac{\int_0^L \mathrm{d}x\, \mathrm{e}^{-\beta U'(x)} \int_x^L \mathrm{d}x'\, \mathrm{e}^{\beta U'(x')}}{\int_0^L \mathrm{d}x\, \mathrm{e}^{\beta U'(x)}} \tag{82b}$$

$$H_2 = \frac{\int_0^L \mathrm{d}x\, \mathrm{e}^{-\beta U'(x)} \int_0^x \mathrm{d}x'\, \mathrm{e}^{\beta U'(x')}}{\int_0^L \mathrm{d}x\, \mathrm{e}^{\beta U'(x)}}. \tag{82c}$$

Berezhkovskii and Bezrukov [141] explicitly considered the rates of entering into and escaping from the channel. In their formulation, the flux is

$$j = P_0'(k_{\mathrm{a}1} C_1 F_1 - k_{\mathrm{a}2} C_2 F_2) \tag{83a}$$

where $k_{\mathrm{a}i}$ are the rate constants for an ion to enter the channel ends and F_i are the probabilities that the ion escapes from the opposite end. The probability $P_0\prime$ for an empty pore is again introduced to enforce the requirement that only one ion can occupy the pore, with the prime indicating a new value, to be given shortly. For the present model of ion channel,

$$k_{\mathrm{a}1} = k_{\mathrm{a}2} = 4Da \tag{83b}$$

$$F_1/F_2 = \mathrm{e}^{\beta eV} \tag{83c}$$

$$1/F_2 = 1 + \mathrm{e}^{\beta eV} + (4D/\pi D_{/\!/} a) \int_0^L \mathrm{d}x\, \mathrm{e}^{\beta U'(x)} \tag{83d}$$

where D is the diffusion constant in the bulk phase. Equation (83*a*) becomes

$$j = D_{/\!/}\sigma P_0' H'(C_1 \mathrm{e}^{\beta eV} - C_2) \tag{84a}$$

with

$$1/H' = 1/H + (\pi D_{/\!/} a/4D)(1 + \mathrm{e}^{\beta eV}). \tag{84b}$$

The new value for the probability of an empty pore is calculated from the average lifetimes of the channel in the empty and occupied states, and given by

$$1/P_0' = 1 + \sigma C_1 \mathrm{e}^{\beta eV} H_1' + \sigma C_2 H_2' \tag{85a}$$

with

$$H_1' = \frac{\int_0^L \mathrm{d}x\, \mathrm{e}^{-\beta U'(x)} \int_x^L \mathrm{d}x'\, \mathrm{e}^{\beta U'(x')} + (\pi D_{/\!/} a/4D) \int_0^L \mathrm{d}x\, \mathrm{e}^{-\beta U'(x)}}{\int_0^L \mathrm{d}x\, \mathrm{e}^{\beta U'(x)} + (\pi D_{/\!/} a/4D)(1 + \mathrm{e}^{\beta eV})} \tag{85b}$$

$$H_2' = \frac{\int_0^L \mathrm{d}x\, \mathrm{e}^{-\beta U'(x)} \int_0^x \mathrm{d}x'\, \mathrm{e}^{\beta U'(x')} + (\pi D_{/\!/} a/4D) \int_0^L \mathrm{d}x\, \mathrm{e}^{-\beta U'(x)}}{\int_0^L \mathrm{d}x\, \mathrm{e}^{\beta U'(x)} + (\pi D_{/\!/} a/4D)(1 + \mathrm{e}^{\beta eV})}. \tag{85c}$$

Comparison of equations (84*a*) and (81*e*) shows that the Berezhkovskii–Bezrukov formulation introduces factors proportional to $\pi D_{/\!/} a/4DL$ to the expressions for H (equation (84*b*)), H_1 (equation (85*b*)), and H_2 (equation (85*c*)). With $a/L \ll 1$ and $D_{/\!/}/D$ expected to be $\ll 1$, the contributions of these additional factors should be small.

The dielectric barrier hinders the flux. The main function of an ion channel is to help overcome the dielectric barrier for the intended ion (such as K^+) while maintaining significant hindrance for competing ions (such as Na^+). For example, in both the gramicidin A channel and the KcsA potassium channel, the large dehydration cost of an ion is compensated by coordination with protein backbone carbonyl oxygens [134, 135, 142]. Multiple specific 'binding' sites for ions can be identified [135, 143]. Occupation of multiple sites gives rise to significant repulsion between the ions, thereby accelerating their conduction [144, 145]. Protein dynamics plays an essential role in facilitating the translocation of ions across the channel by lowering energetic barriers.

Constrictions provide an easy explanation for selectivity against larger ions; it is thus intriguing how potassium channels are selective against the small sodium ion. A contributing factor must be the much more favorable hydration energy of the smaller Na^+ in the bulk phase, leading to a higher dehydration cost. A recent study by Roux and co-workers [146] suggests that the carbonyl dipoles happen to have just the perfect magnitude to compensate the dehydration of K^+ but not Na^+, and repulsion between dynamically fluctuating carbonyls is essential to the robustness of ion selectivity.

4. Conclusion and outlook

A number of simple models have been presented to illustrate the attainment and regulation of biochemical reaction rates and other speeds that measure changes in biological processes. Several unifying themes emerge.

1. Rate limits can be extended by interactions and dynamics of biomolecules. Enzymes provide a prime example in accelerating chemical reactions (from k_{chem} to k_{cat}) through their interactions with substrates, but other examples are abound. Through rapid local

conformational fluctuations, the rate for forming a native contact in a protein increases from $k_f/(1+\nu_1)/(1+\nu_2)$ to k_{fg}. The association rate between two proteins is increased from k_{a0} to k_a by electrostatic attraction. Rate of ion conduction is enhanced by interactions and dynamics of the channel protein.

2. Preferential binding appears as an important vehicle. In enzyme catalysis, preference for the transition state over the substrate is the root cause for rate increase. In chaperonin-assisted protein folding, preference of GroES for ATP-bound GroEL over other nucleotide species contributes to the ordering of the folding cycle. In ion conduction, preferential binding is tantamount to selectivity. Preferential binding is usually accompanied (perhaps afforded) by large-scale conformational rearrangements, e.g., by an enzyme upon substrate binding or by GroEL upon change in nucleotide binding.
3. The potential for multiple pathways can be tamed so that only one or a few routes dominate. This appears to be the case in (spontaneous) protein folding. In chaperonin-assisted folding, though GroEL can be involved in a considerable number of possible binding pathways, it most likely passes through only a few binding species, leading to an ordered folding cycle. In motor proteins, though motion in both directions is experienced, only one direction is favored.
4. Processivity often appears useful. This is true for substrate channeling and for binding to specific sites on DNA. Processivity is also intrinsic to linear motor proteins.
5. The diffusion equation and its discrete version, the master equation, provide a general tool for modeling different processes. Often the models become realistic when the parameters (e.g., equilibrium, diffusion, and rate constants and potential functions) are either adjusted to fit experimental data or extracted from atomistic simulations. The models therefore can bridge between experimental studies and atomistic simulations. However, the lasting value of theoretical models may lie in their intuitive illustration of fundamental ideas. Such fundamental ideas often transcend different phenomenologies and contribute to a unified understanding of diverse biological processes.

Acknowledgments

The author thanks Robert L Baldwin, Hue Sun Chan, Feng Gai, Martin Gruebele and Attila Szabo for stimulating discussion, Adrian Elcock, Stephen Hagen, J Andrew McCammon, Victor Muñoz, George Oster, Benoit Roux and Wei Yang for reading an earlier version of the paper, and Valerie Parsegian for carefully checking for clarity. This work was supported in part by NIH grant GM58187.

Glossary

Speed used here as a general term for measuring the rate of change in biological processes. Examples include the rate of chemical reactions, the velocity of a motor protein and the flux of an ion through a membrane channel.

Diffusion equation describes the average behavior of stochastic motion of a variable. In the discrete case, the diffusion equation becomes the master equation. When both continuous and discrete variables are present, one has the diffusion-reaction equation.

Transition-state ensemble usually defined as the top of a free-energy surface. It may be generalized as consisting of microstates from which local-equilibrium reaction fluxes are minimal.

Preferential binding the tendency of a macromolecule, proteins in particular, to bind a ligand much more strongly over related ligands. One possibility is that, among the conformations assessable to the macromolecule, a better fit can be made with the favored ligand. Another possibility is that the ligand-favoring conformations are reached through the binding of a second ligand.

Processivity the ability of a molecule such as a protein to interact with two or more distinct sites on a target. Nonspecific binding to the target can contribute to processivity. Processivity itself does not imply directional movement, although motor proteins, which do move in a particular direction, typically are processive.

References

[1] Jucovic M and Hartley R W 1996 Protein–protein interaction: a genetic selection for compensating mutations at the barnase-barstar interface *Proc. Natl Acad. Sci. USA* **93** 2343–7
[2] Terlau H, Shon K-J, Grilley M, Stocker M, Stuhmer W and Baldomero O M 1996 Strategy for rapid immobilization of prey by a fish-hunting marine snail *Nature* **381** 148–51
[3] Doi M and Edwards S F 1986 *The Theory of Polymer Dynamics* (Oxford: Clarendon) p 391
[4] Kramers H A 1940 Brownian motion in a field of force and the diffusion model of chemical reactions *Physica* **7** 284–304
[5] Hänggi P, Talkner P and Borkovec M 1990 Reaction-rate theory: fifty years after Kramers *Rev. Mod. Phys.* **62** 251–342
[6] Zhou H-X 1989 The exponential nature of barrier crossings studied by Langevin dynamics *Chem. Phys. Lett.* **164** 285–90
[7] Szabo A, Schulten K and Schulten Z 1980 First passage time approach to diffusion controlled reactions *J. Chem. Phys.* **72** 4350–7
[8] Bicout D J and Szabo A 1997 First passage times, correlation functions, and reaction rates *J. Chem. Phys.* **106** 10292–8
[9] Du R, Pande V S, Grosberg A Y, Tanaka T and Shakhnovich E 1998 On the transition coordinate for protein folding *J. Chem. Phys.* **108** 334–50
[10] Gardiner C W 1985 *Handbook of Stochastic Methods* 2nd edn (Berlin: Springer) p 442
[11] Zhou H-X 2001 Disparate ionic-strength dependencies of on and off rates in protein–protein association *Biopolymers* **59** 427–33
[12] Szabo A and Shoup D 1982 Role of diffusion in ligand binding to macromolecules and cell-bound receptors *Biophys. J.* **40** 33–9

[13] Debye P 1942 Reaction rate in ionic solutions *Trans. Electrochem. Soc.* **82** 265–72

[14] Smoluchowski M 1917 Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen *Z. Phys. Chem.* **92** 129–68

[15] Shoup D, Lipari G and Szabo A 1981 Diffusion-controlled bimolecular reaction rates. The effect of rotational diffusion and orientation constraints *Biophys. J.* **36** 697–714

[16] Zhou H-X 1993 Brownian dynamics study of the influences of electrostatic interaction and diffusion on protein–protein association kinetics *Biophys. J.* **64** 1711–26

[17] Schlosshauer M and Baker D 2002 A general expression for bimolecular association rates with orientational constraints *J. Phys. Chem. B* **106** 12079–83

[18] Doi M 1975 Theory of diffusion-controlled reaction between non-simple molecules: I *Chem. Phys.* **11** 107–13

[19] Doi M 1975 Theory of diffusion-controlled reaction between non-simple molecules: II *Chem. Phys.* **11** 115–21

[20] Wilemski G and Fixman M 1973 General theory of diffusion-controlled reactions *J. Chem. Phys.* **58** 4009–19

[21] Zhou H-X and Szabo A 1996 Theory and simulation of the time-dependent rate coefficients of diffusion-influenced reactions *Biophys. J.* **71** 2440–57

[22] Temkin S I and Yakobson B I 1984 Diffusion-controlled reactions of chemically anisotropic molecules *J. Phys. Chem.* **88** 2679–82

[23] Berg O G 1985 Orientation constraints in diffusion-limited macromolecular association. The role of surface diffusion as a rate-enhancing mechanism *Biophys. J.* **47** 1–14

[24] Pauling L 1948 Nature of forces between large molecules of biological interest *Nature* **161** 707–9

[25] Benkovic S J and Hammes-Schiffer S 2003 A perspective on enzyme catalysis *Science* **301** 1196–202

[26] Garcia-Viloca M, Gao J, Karplus M and Truhlar D G 2004 How enzymes work: analysis by modern rate theory and computer simulations *Science* **303** 186–95

[27] Zhou H-X 2001 The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins *Biochemistry* **40** 15069–73

[28] Zhou H-X, Wlodek S T and McCammon J A 1998 Conformation gating as a mechanism for enzyme specificity *Proc. Natl Acad. Sci. USA* **95** 9280–3

[29] Raushel F M, Thoden J B and Holden H M 2003 Enzymes with molecular tunnels *Acc. Chem. Res.* **36** 539–48

[30] Knighton D R, Kan C-C, Howland E, Janson C A, Hostomska Z, Welsh K M and Matthews D A 1994 Structure of and kinetic channeling in bifunctional dihydrofolate reductase-thymidylate synthase *Nat. Struct. Biol.* **1** 186–94

[31] Elcock A H, Huber G A and McCammon J A 1997 Electrostatic channeling of substrates between enzyme active sites: comparison of simulation and experiment *Biochemistry* **36** 16049–58

[32] Miers J B, Postlewaite J C, Zyung T, Chen S, Roemig G R, Wen X, Dlott D D and Szabo A 1990 Diffusion can explain the nonexponential rebinding of carbon-monoxide to protoheme *J. Chem. Phys.* **93** 8771–6

[33] Atreya C E, Johnson E F, Williamson J, Chang S-Y, Liang P-H and Anderson K S 2003 Probing electrostatic channeling in protozoal bifunctional thymidylate synthase-dihydrofolate reductase using site-directed mutagenesis *J. Biol. Chem.* **278** 28901–11

[34] Atreya C E and Anderson K S 2004 Kinetic characterization of bifunctional thymidylate synthase-dihydrofolate reductase (TS-DHFR) from *Cryptosporidium hominis*. A paradigm shift for TS activity and channeling behavior *J. Biol. Chem.* **279** 18314–22

[35] Elcock A H 2002 Atomistic simulations of competition between substrates binding to an enzyme *Biophys. J.* **82** 2326–32

[36] Gruebele M 2002 Protein folding: the free energy surface *Curr. Opin. Struct. Biol.* **12** 161–8

[37] Akmal A and Muñoz V 2004 The nature of the free-energy barriers to two-state folding *Proteins* **57** 142–52

[38] Zwanzig R, Szabo A and Bagchi B 1992 Levinthal's paradox *Proc. Natl Acad. Sci. USA* **89** 20–2

[39] Zwanzig R 1995 Simple model of protein folding kinetics *Proc. Natl Acad. Sci. USA* **92** 9801–4

[40] Doyle R, Simons K, Qian H and Baker D 1997 Local interactions and the optimization of protein folding *Proteins* **29** 282–91

[41] Plaxco K W, Simons K T and Baker D 1998 Contact order, transition state placement and the refolding rates of single domain proteins *J. Mol. Biol.* **277** 985–94

[42] Shoemaker B A and Wolynes P G 1999 Exploring structures in protein folding funnels with free energy functionals: the denatured ensemble *J. Mol. Biol.* **287** 657–74

[43] Galzitskaya O V and Finkelstein A V 1999 A theoretical search for folding/unfolding nuclei in three-dimensional protein structures *Proc. Natl Acad. Sci. USA* **96** 11299–304

[44] Alm E and Baker D 1999 Prediction of protein-folding mechanisms from free-energy landscapes derived from native structures *Proc. Natl Acad. Sci. USA* **96** 11305–10

[45] Muñoz V and Eaton W A 1999 A simple model for calculating the kinetics of protein folding from three-dimensional structures *Proc. Natl Acad. Sci. USA* **96** 11311–6

[46] Guerois R and Serrano L 2000 The SH3-fold family: experimental evidence and prediction of variations in the folding pathways *J. Mol. Biol.* **304** 967–82

[47] Bicout D J and Szabo A 2000 Entropic barriers, transition states, funnels, and exponential protein folding kinetics: a simple model *Protein Sci.* **9** 452–65

[48] Alm E, Morozov A V, Kortemme T and Baker D 2002 Simple physical models connect theory and experiment in protein folding kinetics *J. Mol. Biol.* **322** 463–76

[49] Makarov D, Keller C A, Plaxco K W and Metiu H 2002 How the folding rate constant of simple, single-domain proteins depends on the number of native contacts *Proc. Natl Acad. Sci. USA* **99** 3535–9

[50] Weikl T R, Palassini M and Dill K A 2004 Cooperativity in two-state protein folding kinetics *Protein Sci.* **13** 822–9

[51] Oliva F Y and Muñoz V 2004 A simple thermodynamic test to discriminate between two-state and downhill folding *J. Am. Chem. Soc.* **126** 8596–7

[52] Jacobsen H and Stockmayer W 1950 Intramolecular reaction in polycondensations: I. The theory of linear systems *J. Chem. Phys.* **18** 1600–6

[53] Wallin S and Chan H S 2005 A critical assessment of the topomer search model of protein folding using a continuum explicit-chain model with extensive conformational sampling *Protein Sci.* **14** 1643–60

[54] Berezhkovskii A and Szabo A 2004 Ensemble of transition states for two-state protein folding from the eigenvectors of rate matrices *J. Chem. Phys.* **121** 9186–7

[55] Matouschek A and Fersht A R 1991 Protein engineering in analysis of protein folding pathways and stability *Methods Enzymol.* **202** 81–112

[56] Merlo C, Dill K A and Weikl T R 2005 Φ values in protein-folding kinetics have energetic and structural components *Proc. Natl Acad. Sci. USA* **102** 10171–5

[57] Burton R E, Myers J K and Oas T G 1998 Protein folding dynamics: quantitative comparison between theory and experiment *Biochemistry* **37** 5337–43

[58] Mayor U, Johnson C M, Daggett V and Fersht A R 2000 Protein folding and unfolding in microseconds to

nanoseconds by experiment and simulation *Proc. Natl Acad. Sci. USA* **97** 13518–22
[59] Qiu L, Pabit S A, Roitberg A E and Hagen S J 2002 Smaller and faster: the 20-residue Trp-cage protein folds in 4 micros *J. Am. Chem. Soc.* **124** 12952–3
[60] Liu Z and Chan H S 2005 Desolvation is a likely origin of robust enthalpic barriers to protein folding *J. Mol. Biol.* **349** 872–89
[61] Kanehisa M I and Tsong T Y 1978 Mechanism of the multiphasic kinetics in the folding and unfolding of globular proteins *J. Mol. Biol.* **124** 177–94
[62] Cieplak M, Henkel M, Karbowski J and Banavar J R 1998 Master equation approach to protein folding and kinetic traps *Phys. Rev. Lett.* **80** 3650–3
[63] Ozkan S B, Dill K A and Bahar I 2002 Fast-folding protein kinetics, hidden intermediates, and the sequential stabilization model *Protein Sci.* **11** 1958–70
[64] Ozkan S B, Dill K A and Bahar I 2003 Computing the transition state populations in simple protein models *Biopolymers* **68** 35–46
[65] Schonbrun J and Dill K A 2003 Fast protein folding kinetics *Proc. Natl Acad. Sci. USA* **100** 12678–82
[66] Van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* (Amsterdam: North-Holland) p 465
[67] Press W H, Flannery B P, Teukolsky S A and Vettering W T 1986 *Numerical Recipes* (Cambridge: Cambridge University Press) p 818
[68] Chan H S and Dill K A 1993 Energy landscapes and the collapse dynamics of homopolymers *J. Chem. Phys.* **99** 2116–27
[69] Zhou H-X 2003 Theory for the rate of contact formation in a polymer chain with local conformational transitions *J. Chem. Phys.* **118** 2010–5
[70] Zhou H-X 2001 Single-chain versus dimeric protein folding: Thermodynamic and kinetic consequences of covalent linkage *J. Am. Chem. Soc.* **123** 6730–1
[71] Zhou H-X 2002 A model for the binding of the inactivation N-terminal to the ion pore of Shaker potassium channel: both electrostatic attraction and covalent linkage are required for rapid inactivation *J. Phys. Chem. B* **106** 2393–7
[72] Zhou H-X 2001 Loops in proteins can be modeled as worm-like chains *J. Phys. Chem. B* **105** 6763–6
[73] Karplus M and Weaver D L 1994 Protein folding dynamics: the diffusion-collision model and experimental data *Protein Sci.* **3** 650–68
[74] Du D, Zhu Y, Huang C Y and Gai F 2004 Understanding the key factors that control the rate of beta-hairpin folding *Proc. Natl Acad. Sci. USA* **101** 15915–20
[75] Muñoz V, Thompson P A, Hofrichter J and Eaton W A 1997 Folding dynamics and mechanism of beta-hairpin formation *Nature* **390** 196–9
[76] Muñoz V, Henry E R, Hofrichter J and Eaton W A 1998 A statistical mechanical model for β-hairpin kinetics *Proc. Natl Acad. Sci. USA* **95** 5872–9
[77] Dinner A R, Lazaridis T and Karplus M 1999 Understanding β-hairpin formation *Proc. Natl Acad. Sci. USA* **96** 9068–73
[78] Karanicolas J and Brooks C L III 2002 The origins of asymmetry in the folding transition states of protein L and protein G *Protein Sci.* **11** 2351–61
[79] Moran L B, Schneider J P, Kentsis A, Reddy G A and Sosnick T R 1999 Transition state heterogeneity in GCN4 coiled coil folding studied by using multisite mutations and crosslinking *Proc. Natl Acad. Sci. USA* **96** 10699–704
[80] Jacob M, Schindler T, Balbach J and Schmid F X 1997 Diffusion control in an elementary protein folding reaction *Proc. Natl Acad. Sci. USA* **94** 5622–7
[81] Plaxco K W and Baker D 1998 Limited internal friction in the rate-limiting step of a two-state protein folding reaction *Proc. Natl Acad. Sci. USA* **95** 13591–6
[82] Bhattacharyya R P and Sosnick T R 1999 Viscosity dependence of the folding kinetics of a dimeric and monomeric coiled coil *Biochemistry* **38** 2601–9
[83] Qiu L, Zachariah C and Hagen S J 2003 Fast chain contraction during protein folding: 'foldability' and collapse dynamics *Phys. Rev. Lett.* **90** 168103
[84] Qiu L and Hagen S J 2004 A limiting speed for protein folding at low solvent viscosity *J. Am. Chem. Soc.* **126** 3398–9
[85] Pabit S A, Roder H and Hagen S J 2004 Internal friction controls the speed of protein folding from a compact configuration *Biochemistry* **43** 12532–8
[86] Bicout D J, Berezhkovskii A M and Szabo A 2001 Irreversible bimolecular reactions of Lagevin particles *J. Chem. Phys.* **114** 2293–303
[87] Zhou H-X and Szabo A 1991 Comparison between molecular dynamics simulations and the Smoluchowski theory of reactions in a hard sphere liquid *J. Chem. Phys.* **95** 5948–52
[88] Hartl F U and Hayer-Hartl M 2002 Molecular chaperones in the cytosol: from nascent chain to folded protein *Science* **295** 1852–8
[89] Saibil H R and Ranson N A 2002 The chaperonin folding machine *Trends Biochem. Sci.* **27** 627–32
[90] Yifrach O and Horovitz A 1995 Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL *Biochemistry* **34** 5303–8
[91] Rye H S, Burston S G, Fenton W A, Beechem J M, Xu Z, Sigler P B and Horwich A L 1997 Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL *Nature* **388** 792–8
[92] Kad N M, Ranson N A, Cliff M J and Clarke A R 1998 Asymmetry, commitment and inhibition in the GroE ATPase cycle impose alternating functions on the two GroEL rings *J. Mol. Biol.* **278** 267–78
[93] Farr G W, Furtak K, Rowland M B, Ranson N A, Saibil H R, Kirchhausen T and Horwich A L 2000 Multivalent binding of nonnative substrate proteins by the chaperonin GroEL *Cell* **100** 561–73
[94] Roseman A M, Chen S, White H, Braig K and Saibil H R 1996 The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL *Cell* **87** 241–51
[95] Xu Z, Horwich A L and Sigler P B 1997 The crystal structure of the asymmetric BroEL-GroES-(ADP) chaperonin complex *Nature* **388** 741–50
[96] Zhou H-X and Dill K A 2001 Stabilization of proteins in confined spaces *Biochemistry* **40** 11289–93
[97] Schreiber G and Fersht A R 1996 Rapid, electrostatically assisted association of proteins *Nature Struct. Biol.* **3** 427–31
[98] Riggs A D, Bourgeois S and Cohn M 1970 The lac repressor-operator interaction. 3. Kinetic studies *J. Mol. Biol.* **53** 401–17
[99] Sparrer H, Lilie H and Buchner J 1996 Dynamics of the GroEL-protein complex: effects of nucleotides and folding mutants *J. Mol. Biol.* **258** 74–87
[100] Clark A C and Frieden C 1997 GroEL-mediated folding of structurally homologous dihydrofolate reductases *J. Mol. Biol.* **268** 512–25
[101] Perrett S, Zahn R, Stenberg G and Fersht A R 1997 Importance of electrostatic interactions in the rapid binding of polypeptides to GroEL *J. Mol. Biol.* **269** 892–901
[102] Tsurupa G P, Ikura T, Makio T and Kuwajima K 1998 Refolding kinetics of staphyloccocal nuclease and its

mutants in the presence of the chaperonin GroEL *J. Mol. Biol.* **277** 733–45
[103] Rye H S, Roseman A M, Chen S, Furtak K, Fenton W A, Saibil H R and Horwich A L 1999 GroEL–GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings *Cell* **97** 325–38
[104] Burston S G, Ranson N A and Clarke A R 1995 The origins and consequences of asymmetry in the chaperonin reaction cycle *J. Mol. Biol.* **249** 138–52
[105] Shoemaker B A, Portman J J and Wolynes P G 2000 Speeding molecular recognition by using the folding funnel: the fly-casting mechanism *Proc. Natl Acad. Sci. USA* **97** 8868–73
[106] Zhou H-X 2004 Polymer models of protein stability, folding, and interactions *Biochemistry* **43** 2141–54
[107] Zhou H-X 1997 Enhancement of protein–protein association rate by interaction potential: accuracy of prediction based on local Boltzmann factor *Biophys. J.* **73** 2441–5
[108] Vijayakumar M, Wong K-Y, Schreiber G, Fersht A R, Szabo A and Zhou H-X 1998 Electrostatic enhancement of diffusion-controlled protein–protein association: comparison of theory and experiment on barnase and barstar *J. Mol. Biol.* **278** 1015–24
[109] Zhou H-X 2003 Association and dissociation kinetics of colicin E3 and immunity protein 3 convergence of theory and experiment *Protein Sci.* **12** 2379–82
[110] Zhou H-X 1996 Effect of interaction potentials in diffusion-influenced reactions with small reactive regions *J. Chem. Phys.* **105** 7235–7
[111] Adam G and Delbruck M 1968 Reduction of dimensionality in biological diffusion processes *Structural Chemistry and Molecular Biology* ed N Davidson (San Francisco: Freeman) pp 198–215
[112] Richter P H and Eigen M 1974 Diffusion controlled reaction rates in spheroidal geometry. Application to repressor-operator association and membrane bound enzymes *Biophys. Chem.* **2** 255–63
[113] Berg O G, Winter R B and von Hippel P H 1981 Diffusion-driven mechanisms of protein translocation on nucleic acids: 1. Models and theory *Biochemistry* **20** 6929–48
[114] Berg O G and Ehrenberg M 1982 Association kinetics with coupled three- and one-dimensional diffusion. Chain-length dependence of the association rate to specific DNA sites *Biophys. Chem.* **15** 41–51
[115] Zhou H-X and Szabo A 2004 Enhancement of association rates by nonspecific binding to DNA and cell membranes *Phys. Rev. Lett.* **93** 178101
[116] Zhou H-X 2005 A model for the mediation of processivity of DNA-targeting proteins by nonspecific binding: Dependence on DNA length and presence of obstacles *Biophys. J.* **88** 1608–15
[117] Terry B J, Jack W E and Modrich P 1985 Facilitated diffusion during catalysis by *Eco* RI endonuclease. Nonspecific interactions in *Eco* RI catalysis *J. Biol. Chem.* **260** 13130–7
[118] Stanford N P, Szczelkun M D, Marko J F and Halford S E 2000 One- and three-dimensional pathways for proteins to reach specific sites *EMBO J.* **19** 6546–57
[119] Bustamante C, Keller D and Oster G 2001 The physics of molecular motors *Acc. Chem. Res.* **34** 412–20
[120] Astumian R D and Bier M 1996 Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis *Biophys. J.* **70** 637–53
[121] Zhou H-X and Chen Y-d 1996 Chemically driven motility of Brownian particles *Phys. Rev. Lett.* **77** 194–7
[122] Astumian R D 1997 Thermodynamics and kinetics of a Brownian motor *Science* **276** 917–22
[123] Elston T, Wang H and Oster G 1998 Energy transduction in ATP synthase *Nature* **391** 510–3
[124] Dimroth P, Wang H, Grabe M and Oster G 1999 Energy transduction in the sodium F-ATPase of *Propionigenium modestum Proc. Natl Acad. Sci. USA* **96** 4924–9
[125] Junge W 1999 ATP synthase and other motor proteins *Proc. Natl Acad. Sci. USA* **96** 4735–7
[126] Okada Y and Hirokawa N 2000 Mechanism of the single-headed processivity: diffusional anchoring between the K-loop of kinesin and the C terminus of tubulin *Proc. Natl Acad. Sci. USA* **97** 640–5
[127] Oster G and Wang H 2003 Rotary protein motors *Trends Biochem. Sci.* **13** 114–21
[128] Bier M 2003 Processive motor protein as an overdamped Brownian stepper *Phys. Rev. Lett.* **91** 148101
[129] Kinosita K Jr, Adachi K and Itoh H 2004 Rotation of F_1-ATPase: how an ATP-driven molecular machine may work *Annu. Rev. Biophys. Biomol. Struct.* **33** 245–68
[130] Yildiz A, Tomishige M, Vale R and Selvin P R 2004 Kinesin walks hand-over-hand *Science* **303** 676–8
[131] Xing J, Wang H, Dimroth P, von Balmoos C and Oster G 2004 Torque generation by the sodium Fo-ATPase *Biophys. J.* **87** 2148–63
[132] Parsegian A 1969 Energy of an ion crossing a low dielectric membrane: solutions to four relevant electrostatic problems *Nature* **221** 844–6
[133] Roux B 2005 Ion conduction and selectivity in K^+ channels *Annu. Rev. Biophys. Biomol. Struct.* **34** 153–71
[134] Ketchem R R, Roux B and Cross T A 1997 High resolution refinement of a solid-state NMR derived structure of gramicidin A in a lipid bilayer environment *Structure* **5** 11655–69
[135] Zhou Y, Morais-Cabral J H, Kaufman A and MacKinnon R 2001 Chemistry of ion coordination and hydration revealed by a K^+ channel-Fab complex at 2.0 Å resolution *Nature* **414** 43–8
[136] Jiang Y, Lee A, Chen J, Cadene M, Chait B T and MacKinnon R 2002 Crystal structure and mechanism of a calcium-gated potassium channel *Nature* **417** 515–22
[137] Nishimura K, Kim S, Zhang L and Cross T A 2002 The closed state of a H^+ channel helical bundle combining precise orientational and distance restraints from solid state NMR *Biochemistry* **41** 13170–7
[138] Jiang Y, Lee A, Chen J, Ruta V, Cadene M, Chalt B T and MacKinnon R 2003 X-ray structure of a voltage-dependent K^+ channel *Nature* **423** 33–41
[139] Kuo A, Gulbis J M, Antcliff J F, Rahman T, Lowe E D, Zimmer J, Cuthbertson J, Ashcroft F M, Ezaki T and Doyle D A 2003 Crystal structure of the potassium channel KirBac1.1 in the closed state *Science* **300** 1922–6
[140] Levitt D G 1986 Interpretation of biological ion channel flux data-reaction-rate versus continuum theory *Annu. Rev. Biophys. Biophys. Chem.* **15** 29–57
[141] Berezhkovskii A M and Bezrukov S M 2004 Optimizing transport of metabolites through large channels: molecular sieves with and without binding *Biophys. J.* **88** L17–9
[142] Roux B, Allen T, Berneche S and Im W 2004 Theoretical and computational models of biological ion channels *Q. Rev. Biophys.* **37** 15–103
[143] Zhou M and MacKinnon R 2004 A mutant KcsA K^+ channel with altered conduction properties and selectivity filter ion distribution *J. Mol. Biol.* **338** 839–46
[144] Morais-Cabral J H, Zhou Y and MacKinnon R 2001 Energetic optimization of ion conduction rate by the K^+ selectivity filter *Nature* **414** 37–42
[145] Berneche S and Roux B 2001 Energetics of ion conduction through the K^+ channel *Nature* **414** 73–7
[146] Noskov S, Berneche S and Roux B 2004 Control of ion selectivity by electrostatic and dynamic properties of carbonyl ligands *Nature* **431** 830–4