Force-dependent switch in protein unfolding pathways and transition-state movements

Pavel I. Zhuravlev1-4, Michael Hinczewski1-2, Shaon Chakrabarti5, Susan Marqusee2, and D. Thirumalai6

1Biophysics Program, Institute for Physical Science and Technology, Department of Chemistry & Biochemistry, University of Maryland, College Park, MD 20742; 2Department of Physics, Case Western Reserve University, Cleveland, OH 44106; and 3California Institute for Quantitative Biosciences, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

Edited by Ken A. Dill, Stony Brook University, Stony Brook, NY, and approved December 24, 2015 (received for review August 10, 2015)

Although it is known that single-domain proteins fold and unfold by parallel pathways, demonstration of this expectation has been difficult to establish in experiments. Unfolding rate, \( k_u(f) \), as a function of force \( f \), obtained in single-molecule pulling experiments on src SH3 domain, exhibits upward curvature on a log \( k_u(f) \) plot. Similar observations were reported for other proteins for the unfolding rate \( k_u(\Delta) \). These findings imply unfolding in these single-domain proteins involves a switch in the pathway as force \( f \) or \( \Delta \) is increased from a low to a high value. We provide a unified theory demonstrating that if log \( k_u(\Delta) \) as a function of a perturbation \( f \) or \( \Delta \) exhibits upward curvature then the underlying energy landscape must be strongly multidimensional. Using molecular simulations, we provide a structural basis for the switch in the pathways and dramatic shifts in the transition-state ensemble (TSE) in src SH3 domain as force \( f \) is increased. We show that a single-point mutation shifts the upward curvature in log \( k_u(f) \) to a lower force, thus establishing the malleability of the underlying folding landscape. Our theory, applicable to any perturbation that affects the free energy of the protein linearly, readily explains movement in the TSE in a \( \beta \)-sandwich (I27) protein and single-chain monellin as the denaturant concentration is varied. We predict that in the force range accessible in laser optical tweezer experiments there should be a switch in the unfolding pathways in I27 or its mutants.

Protein folding | parallel pathways | single-molecule force spectroscopy

T

hat single-domain proteins should fold by parallel or multiple pathways emerges naturally from theoretical considerations (1–3) and computational studies (4–7). The generality of the conclusions reached in the theoretical studies is sufficiently compelling, which makes it surprising that they are not routinely demonstrated in typical ensemble folding experiments. The reasons for the difficulties in directly observing parallel folding or unfolding pathways in monomeric proteins can be appreciated based on the following arguments. Consider a protein that reaches the folded state by two different pathways. The ratio of flux through these pathways is proportional to \( \exp \left( \frac{\Delta G_f^1 - \Delta G_f^2}{k_BT} \right) \), where \( \Delta G_f^1 \) and \( \Delta G_f^2 \) are, respectively, the free energy barriers separating the folded and unfolded states along the two pathways, \( k_BT \) is the Boltzmann constant, and \( T \) is the temperature. If \( \Delta G^2 = \Delta G_f^2 - \Delta G_f^1 > 0 \) is large compared with \( k_BT \), then the experimental detection of flux through the high free energy barrier pathway \( (H) \) is unlikely. External perturbations (such as mechanical force \( f \) or denaturants \( C \)) could reduce \( \Delta G_f \). The values of \( f \) or \( C \) should fall in an experimentally accessible range for detecting a potential switch between pathways. Despite these inherent limitations, Clarke and coworkers showed in a most insightful study that unfolding of Ig domain (I27) induced by denaturants occurs by parallel routes (8). Subsequently, additional experiments on single-chain monellin (9), using denaturants and spectroscopic probes, have firmly shown the existence of multiple paths. Thus, it seems that multiple folding routes can be detected in standard folding experiments (10, 11) provided the flux through the higher free energy barriers is not so small that it escapes detection. In addition, parallel folding pathways have been observed in repeat proteins, where inherent symmetry in the connectivity of the individual domains (12) results in parallel assembly.

Single-molecule pulling experiments in which \( f \) is applied to specific locations on the protein have demonstrated that unfolding of many proteins follows complex multiple routes. Mechanical force, unlike denaturants, does not alter the effective microscopic interactions between the residues, thus allowing for a cleaner interpretation. More importantly, by following the fate of many individual proteins the underlying heterogeneity in the routes explored by the protein can be revealed. Indeed, pulling experiments and simulations on a variety of single-domain proteins (13–15) show clear signatures of many routes for \( f \)-induced unfolding. It could be argued that in many of these studies the network of states connecting the folded and unfolded states is a consequence of complex topology, although they are all single-domain proteins. However, the src SH3 domain is a small protein with no apparent symmetry in the arrangement of secondary structure elements that folds in an apparent two-state manner. Thus, the discovery that it unfolds using parallel pathways (16, 17) is unexpected and requires a firm theoretical explanation and structural interpretation.

In single-molecule pulling experiments, performed at constant force or constant loading rate, only a 1D coordinate, the molecular extension \( x \), is readily measurable. When performed at constant \( f \) it is possible to generate folding trajectories \( x \) as a function of time, from which equilibrium 1D free energy profiles, \( F(x) \), can be extracted using rigorous theory (18). The utility of \( F(x) \) hinges on the crucial assumption that all other degrees of

Significance

Single-domain proteins with symmetrical arrangement of secondary structural elements in the native state are expected to fold by diverse pathways. However, understanding the origins of pathway diversity, and the experimental signatures for identifying it, are major challenges, especially for small proteins with no obvious symmetry in the folded state. We show rigorously that upward curvature in the logarithm of unfolding rates as a function of force (or denaturants) implies that the folding occurs by diverse pathways. The theoretical concepts are illustrated using simulations of src SH3 domain, which explain the emergence of parallel pathways in single-molecule pulling experiments and provide structural description of the routes to the unfolded state. We make testable predictions illustrating the generality of the theory.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

To whom correspondence may be addressed. Email: pavel.zhuravlev@gmail.com or mhx605@case.edu.

This article contains supporting information online at www.pnas.orglookup/suppl/doi:10.1073/pnas.1515730113–/DCSupplemental.
freedom in the system including the solvent come to equilibrium on time scales much faster than in \( x \), so that \( x \) itself may be considered to be an accurate reaction coordinate.

A straightforward way to assess whether a 1D picture is adequate is to analyze the \( f \) dependence of the unfolding rate \( k_u(f) \), which can be experimentally obtained at constant \( f \) or computed from unfolding rates measured at different loading rates (18, 19). The observed upward curvature in the \( \log k_u(f) \) plot in src SH3 (16) was shown to be a consequence of unfolding by two pathways, one dominant at low forces and the other at high forces. It was succinctly argued that the measured \( \log k_u(f) \) data cannot be explained by multiple barriers in a 1D \( F(x) \) or a 1D profile with a single barrier in which the unfolding rate is usually fit using the Bell model \( k_u(f) = k_0 \exp(\frac{x}{\alpha}) \), where \( k_0 \) is the unfolding rate at \( f = 0 \), and \( x^2 \) is the location of the barrier in \( F(x) \) at zero force with respect to the folded state. (Throughout this paper, by location of the barrier, or the transition state, we mean the location with respect to the folded state.) The upward curvatures in the monotonous \( \log k_u(f) \) as well as \( \log k_u(C) \) plots, observed experimentally, necessarily imply that parallel routes are involved in the unfolding process. (A nonmonotonic \( \log k_u(f) \) plot suggests catch bond behavior.)

To provide a general framework for a quantitative explanation of a broad class of experiments, we first present a rigorous theoretical proof that upward curvature in \( \log k_u(f) \) (or \( \log k_u(C) \)) implies that the folding landscape is strongly multidimensional (SMD). Hence, such SMD landscapes cannot be reduced to 1D or superposition of meaningful 1D landscapes, which can rationalize the observed convex \( \log k_u(f) \) plots. We note \textit{en passant} that the shape of the measured \( \log k_u(f) \) plot cannot be justified using \( F(x) \) even if \( x^2 \) were allowed to depend on \( f \), moving toward the folded state as \( f \) increases. The theory only hinges on the assumption that the perturbation \( f \) or \( C \) is linearly coupled to the effective energy function of the protein. To illustrate the structural origin of the upward curvature in the \( \log k_u(f) \) plot we also performed simulations of \( f \)-induced unfolding of src SH3 domain, to identify the structural details of the unfolding pathways including the movement of transition states as the force is increased. The results of the simulations semiquantitatively reproduce the experimental \( \log k_u(f) \) plots for both the WT and the V61A mutant. More importantly, we also provide the structural basis of the switch in the unfolding pathways as \( f \) is varied, which cannot be obtained using pulling experiments. We obtain structures of the transition-state ensembles (TSEs), demonstrating the change in the TSE structures as \( f \) is increased from a low to a high value.

### Results

#### Nature of the Energy Landscape from \( \log k_u(f) \) Plots

Let us assume that the unfolding rate \( k_u(\alpha) \), as a function of a controllable external perturbation \( \alpha \), can be measured. We assume that \( \alpha \) decreases the stability of the folded state linearly, as is the case in the pulling experiment with \( \alpha = f \), the force applied at two points of the protein. However, the discussion below is quite general and applies to any external parameter with a linear, additive contribution to the effective protein energy function. For a protein under force, the total free energy has the general form \( U(x, r, f) = U_0(x, r) - f \alpha x \), with a force contribution \( \alpha x \), where \( x \) is the end-to-end extension of the protein. Here, \( U_0 \) is the free energy in the absence of applied tension and the vector \( r \) represents all of the additional conformational degrees of freedom besides \( x \).

In the derivation below, we model the dynamics of the protein as diffusion of a single particle on the multidimensional landscape \( U(x, r, f) \). The unfolding of the protein would correspond to the particle starting in the reference protein conformation \( x_0, r_0 \) in the folded state energy basin \( F \) and diffusing to any other conformation, with a given extension \( x_r > x_0 \), representing the unfolded basin \( U \) (Fig. 1). The unfolding time for a particular trajectory is the time when the particle reaches the target conformation for the first time (known as first passage time). Averaging this time over all trajectories yields the mean first passage time (MFPT) from the unfolded to folded state, which we denote as \( t_u(x_0, r_0, f) \), or the average unfolding time. The unfolding rate is the inverse of the unfolding time, \( k_u(f) = 1/t_u(x_0, r_0, f) \).

We are interested in finding the curvature of \( \log k_u \) as a function of \( f \), and in particular the sign of \( \frac{\partial}{\partial f} \log k_u \). Starting from the diffusion equation, we find expressions for the MFPT from any
conform with extension \( t_0(x, r, f) \), and then for \( \log k_u(f) \) and its first two derivatives. It turns out that if we use the assumptions of a single unfolding pathway, the second derivative is negative and the curvature of \( \log k_u(f) \) has to be downward.

The summary of the subsequent derivation is as follows: (i) We start from the equation for \( t_0(x, r, f) \) which can be obtained from the diffusion equation (20), (ii) integrate it over the \( r \) degrees of freedom, (iii) use two assumptions for evaluating the integral with \( t_0(x, r, f) \) inside, (iv) solve the ordinary differential equation for the unfolding time, and (v) establish that the solution implies certain constraints on the shape of \( \log k_u(f) \) plot. Following this derivation in detail is not necessary for understanding the other parts of the paper.

The equation for the MFPT \( t_0(x, r, f) \) can be obtained from the diffusion equation (in Fokker–Planck form) by integration over \( x, r, \) and \( t \), followed by some rearrangements (20). The result is called the backward Kolmogorov equation:

\[
D(x, r) e^{-\beta U(x, r)}/\partial t_0(x, r, f) \right\} \int dr \left[ e^{-\beta U(x, r)}/\partial t_0(x, r, f) \right] \right\} = -1, \tag{1}
\]

with the boundary condition \( t_0(x, r, f) = 0 \), \( \beta = 1/k_BT \) and \( D(x, r) \) being the diffusion constant, which for simplicity is allowed to depend on the coordinate. By dividing both sides of Eq. 1 by \( D(x, r) \) and integrating over the coordinate \( \mathbf{r} \), we obtain

\[
\frac{\partial}{\partial x} \int dr e^{-\beta U(x, r)}/\partial t_0(x, r, f) \right\} = -\int dr D^{-1}(x, r) e^{-\beta U(x, r)}. \tag{2}
\]

To get the result in Eq. 2, we have assumed that \( U(x, r) \to \infty \) at the integration limits of the coordinate space of \( r \) (i.e., the diffusion process is bounded). We rewrite Eq. 2 as

\[
\frac{\partial}{\partial x} \int dr e^{-\beta U(x, r)}/\partial t_0(x, r, f) \right\} = -\int dr D^{-1}(x, r) e^{-\beta U(x, r)} \tag{3}
\]

where \( G(x) \equiv \int dr D^{-1}(x, r) \exp(-\beta U_0(x, r)) \).

Further simplification of the MFPT expression depends on the nature of the multidimensional free energy \( U_0(x, r) \). In particular, we define a class of free energies that satisfy the following two conditions:

1. \( U_0(x, r) \) has a single minimum with respect to \( r \) at each point \( x \) in the range \( x_0 < x < x_u \). We denote the location of this minimum as \( r_m(x) \).
2. The Boltzmann factor \( \exp(-\beta U_0(x, r)) \) for \( r \) near \( r_m(x) \) is sharply peaked, so the thermodynamic contribution from conformations with coordinates far from \( r_m(x) \) is negligible. In other words, we assume fast equilibration along the \( r \) coordinates at each \( x \), compared with the timescale of first passage between \( N \) and \( U \).

A schematic illustration of a \( U_0(x, r) \) satisfying these requirements is shown in Fig. 1A. Diffusion is essentially confined to a single, narrow reaction pathway in the multidimensional space. We will call any \( U_0(x, r) \) in this category weakly multidimensional (WMD) with respect to \( x \), because the diffusion process is quasi-1D in terms of the reaction coordinate \( x \). In contrast, any \( U_0(x, r) \) that violates either one of the above conditions will be called SMD, because it has characteristics that qualitatively distinguish it from any 1D diffusion process. Note that this categorization makes no other assumptions about the shape of \( U_0(x, r) \) except for those specified above: For example, there could be one or many free energies barriers separating \( N \) and \( U \), or none at all. Fig. 1B and C show two examples of \( U_0(x, r) \) that are SMD. In both cases, condition \( i \) is violated, because in the range \( x_0 < x < x_u \) there is no unique minimum in \( U_0(x, r) \) along \( r \). For Fig. 1B, there are two possible reaction pathways between \( N \) and \( U \), whereas for Fig. 1C there is a single pathway, but it is nonmonotonic in \( x \).

For an energy landscape that is WMD, there are rigorous bounds on the first and second derivatives of \( \log k_u(f) \) with respect to \( f \). To derive these bounds, note that the WMD assumptions allow us to make a saddle-point approximation to the integral over \( r \) on the left-hand side of Eq. 3, setting the value of \( r \) in \( \partial \tau_u(x, r) / \partial r \) to \( r_m(x) \). Because this will be the dominant contribution, we approximate Eq. 3:

\[
\frac{\partial}{\partial x} \left[ \int \frac{\partial \tau_u(x, r_m, f)}{\partial r} \right] \int dr e^{-\beta U_0(x, r)} \approx -e^{-\beta f} G(x). \tag{4}
\]

By simplifying the notation by defining \( \tau_u(x, f) \equiv \tau_u(x, r_m, f) \) and \( I(x) \equiv \int dr \exp(-\beta U_0(x, r)) \), Eq. 4 becomes

\[
\frac{\partial}{\partial x} \left[ \int \frac{\partial \tau_u(x, f)}{\partial r} \right] \int dr e^{-\beta U_0(x, r)} I(x) \approx -e^{-\beta f} G(x). \tag{5}
\]

The solution for \( \tau_u(x, f) \) from Eq. 5, with boundary condition \( \tau_u(x_0, f) = 0 \), can be written as a Laplace transform of a function \( H(x, s) \),

\[
\tau_u(x, f) = \int_0^\infty ds e^{-s \beta} H(x, s), \tag{6}
\]

\[
H(x, s) \equiv \int_x^{x_u} dt 1/(G(x') - s). \tag{7}
\]

Both \( I(x) \) and \( G(x) \) are nonnegative functions [because \( D(x, r) > 0 \) and \( \exp(-\beta U_0(x, r)) \geq 0 \) for all \( x \) and \( r \)], so the function \( H(x, s) \) is likewise nonnegative, \( H(x, s) \geq 0 \) for \( x \leq x_u \). From this property, it follows that \( \tau_u(x, f) \geq 0 \), and

\[
\frac{\partial^2}{\partial x^2} \tau_u(x, f) = \int_0^\infty ds (\beta s)^2 e^{-s \beta} H(x, s) \geq 0, \tag{8}
\]

for \( x \leq x_u \). Because the experimental data are typically plotted in terms of \( \log k_u(f) = -\log \tau_u(x, f) \) with respect to \( f \), we are specifically interested in the corresponding derivatives of \( \log k_u(f) \),

\[
\frac{d}{df} \log k_u = -\frac{1}{\tau_u} \frac{\partial \tau_u}{\partial f} \frac{d^2}{df^2} \log k_u = -\frac{1}{\tau_u} \left[ \frac{\partial \tau_u}{\partial f} \right]^2 - \frac{\tau_u \partial^2 \tau_u}{\partial f^2}. \tag{9}
\]

From Eq. 7 we see that \( d \log k_u/df \geq 0 \). The sign of \( d^2 \log k_u/df^2 \) requires establishing the sign of the term in the square brackets in Eq. 8, which can be done by using the Cauchy–Schwarz inequality. Let us define two functions, \( g_1(x, s) = \Theta(s) \sqrt{e^{-\beta f} H(x, s)} \) and \( g_2(x, s) = \sqrt{e^{-\beta f} H(x, s)} \), where \( \Theta(s) = 1 \) for \( s \geq 0 \) and \( = 0 \) for \( s < 0 \). Then, from Eqs. 6 and 7 we obtain \( \tau_u = \int_0^\infty ds [g_1(x, s)]^2 \frac{d^2}{df^2} = \int_0^\infty ds \log [g_2(x, s)]^2. \) Using the Cauchy–Schwarz inequality

\[
\int_{-\infty}^{\infty} ds g_1(x, s)^2 \leq \int_{-\infty}^{\infty} ds \log [g_2(x, s)]^2 \int_{-\infty}^{\infty} ds [g_2(x, s)]^2 \tag{9}
\]

we find that \( \partial \tau_u / \partial f \)^2 \leq \tau_u \partial^2 \tau_u / \partial f^2 \). Hence, from Eq. 8 we see that \( d^2 \log k_u/df^2 \leq 0 \). In summary, we can now state the full criterion for the validity of WMD for describing force-induced unfolding.

Zhuravlev et al.
Criteria for WMD Unfolding Landscape. The unfolding rate $k_u(f)$ on a WMD free energy landscape under applied force $f$ must satisfy

$$\frac{d}{df} \log k_u \geq 0, \quad \frac{d^2}{df^2} \log k_u \leq 0. \quad [10]$$

If $k_u(f)$ fails to satisfy either of the conditions in Eq. 10, the underlying free energy landscape must be SMD, and analyses of the measured data using the end-to-end distance, $x$, as a reaction coordinate are incorrect.

Scenarios for $[\log k_u(f), f]$ Plots in WMD and SMD. A range of behaviors in $[\log k_u(f), f]$ plots can be obtained depending on the nature of the energy landscape. Stochastic simulations in a WMD (Fig. 1A and D) show that $[\log k_u(f), f]$ has a minor downward curvature, which is readily explained by a generalized Bell model in which the transition-state location is allowed to move toward the folded state in accord with the Hammond effect (21). In contrast, in the SMD (Fig. 1B, C, E, and F) the $[\log k_u(f), f]$ plot shows upward curvature. The upward curvature in Fig. 1E indicates loss of flux from the folded state through two channels in Fig. 1B, similar to parallel pathways in protein unfolding experiments. Interestingly, the upward curvature in Fig. 1F from the SMD landscape in Fig. 1C does not come from parallel pathways. Instead, the lifetime of the folded state first increases followed by the usual decrease as $f$ increases. Such a counterintuitive “catch bond” behavior is well documented in a number of protein complexes (22–24). The results in Fig. 1E and F show that violations of Eq. 10 implies that the underlying energy landscape must be SMD.

Naive Analyses of the $f$ Dependence of $k_u(f)$. Using the data generated by molecular dynamics simulations of force unfolding the src SH3 domain, with force applied to residues 9 and 59, for a set of forces $f_i$, we calculated $k_u(f_i)$ for each force $f_i$ as the inverse of the MFPT from the folded state to the unfolded state, by averaging the set of first passage times to unfolding $t_{f_i}$ over the trajectory index $j$ (Materials and Methods). The $[\log k_u(f), f]$ plot for $f$-induced unfolding is nonlinear with upward curvature implying that the free energy landscape is SMD (Fig. 2B). We note par-enthetically that the inadequacy of the Bell model cannot be fixed using movement of the transition state with $f$ or using a 1D free energy profile with two (or more) barriers. Remarkably, the slope change in the simulations qualitatively coincides with measurements on the same protein (16, 17), where constant force was applied to the residues 7 and 59. Thus, both simulations and experiment show that the condition in Eq. 10 is violated, implying that the free energy landscape for SH3 is SMD.

The observed $k_u(f)$ dependence can be fit using a sum of two exponential functions (16),

$$k_u(f) = k_u^0 \exp \left( \frac{x_L}{kT} \right) + k_u^0 \exp \left( \frac{x_H}{kT} \right). \quad [11]$$

The parameters $k_u^0$ and $k_u^0$ (unfolding rates at $f = 0$) and $x_L$ and $x_H$ (putative locations of the transition states) can be precisely obtained from single-molecule pulling experiments (16).

Fig. 2. (A) The native structure of G. gallus src SH3 domain from Tyr kinase, whose sequence is given below. The N-terminal and C-terminal β-strands are denoted as β4 and β5. In the simulations, the force is applied to residues 9 and 59. RT loop is the longest loop of the domain, which is positioned in sequence right after the N-terminal β-strand. (B) The dependence of unfolding rate $k_u(f)$ obtained from the simulations using the SOP-SC model as a function of $f$ given in terms of $[\log k_u(f), f]$ plot. The solid line is a two-exponential fit (Eq. 11) whereas the dashed line is a fit using Bell model. (C) Same as B except the data are obtained from single-molecule pulling experiments (16).
Histories of unfolding trajectories at 15 and 45 pN show the difference between unfolding pathways. The overlap parameter $6.1 \approx 0.1$ Nanometers $0.08 \approx 0.1$ Nanometers $\approx 0$ is given in Materials and Methods. The inability to is due to the unfolding behavior being greater at higher forces. Our previous work (26) showed that the unfolding rate in denaturants is larger by a factor of $150$, which is similar to the difference between experiment and simulations in Fig. 2. However, because the inference about parallel pathways relies solely on the shape of $\log k_u(f)$ the inability to quantitatively reproduce the precise value of $k_u(f)$ is irrelevant.

Despite the good fits to Eq. 11 neither $x_L$ nor $x_H$ can be associated with transition-state location as is traditionally assumed. We show below that such projections onto a 1D coordinate cease to have physical meaning when the underlying folding landscape is SMD. The apparent barriers to unfolding at $f = 0$ along the pathways can be estimated using $k_u^0 \approx k_{u0} \exp(-\Delta G^u_{HF}/k_B T)$ and $k_L^0 \approx k_{u0} \exp(-\Delta G^u_{HF}/k_B T)$. Using the accepted estimates for the pre-factor ($k_{u0} \approx 10^0 \sim 10^3 s^{-1}$) (27–29), and the values of $k_L^0$ and $k_H^0$ from the fits of experimental data (Table 1), we obtain $\Delta G^u_{HF} \approx 18 - 20 k_B T$, depending on the value of $k_{u0}$ and $\Delta G^u_{HF} \approx (26 - 28) k_B T$. If these values are reasonable then the ratio of fluxes through the two pathways at $f = 0$ is $\sim \exp(-\Delta G^u_{HF}/k_B T) \sim e^{-k} \approx 3 \cdot 10^{-4}$, which is much smaller than those obtained in the simulations by

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Error</th>
<th>Units</th>
<th>Value</th>
<th>Error</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_L^0$</td>
<td>1.5</td>
<td>0.4</td>
<td>Seconds$^{-1}$</td>
<td>2.0</td>
<td>1.1</td>
<td>$10^{-2}$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_H^0$</td>
<td>3.9</td>
<td>10.0</td>
<td>$10^{-3}$ s$^{-1}$</td>
<td>6.1</td>
<td>6.4</td>
<td>$10^{-6}$ s$^{-1}$</td>
</tr>
<tr>
<td>$x_L$</td>
<td>0.40</td>
<td>0.1</td>
<td>Nanometers</td>
<td>0.08</td>
<td>0.1</td>
<td>Nanometers</td>
</tr>
<tr>
<td>$x_H$</td>
<td>1.16</td>
<td>0.3</td>
<td>Nanometers</td>
<td>1.42</td>
<td>0.12</td>
<td>Nanometers</td>
</tr>
</tbody>
</table>

The definition of $P_2/P_1$ is given in Materials and Methods. The errors come out of log-likelihood covariance matrix with respect to parameters.

---

**Fig. 3.** Histograms of unfolding trajectories at 15 and 45 pN show the difference between unfolding pathways. The overlap parameter $f$ is the order parameter describing the global unfolding, and $x_{BLT}$ shows ruptured interactions between the RT loop and the core. The lower panels are blowups of the route to the unfolded (U) state, whereas the upper panels show the profiles connecting F and U states (color scheme is also adapted to higher resolution). The forces are explicitly marked. The rupture process of the RT loop is dramatically different at 15 pN and 45 pN, demonstrating a switch in the unfolding pathway.
direct calculation of the flux through the two pathways. In addition, the finding that \( x_f > x_0 \) also makes no physical sense, because we expect the molecule under higher tension to be more brittle (19). These are the first indications that the fits using Eq. 11 do not provide meaningful parameters.

Structural Basis of \( f \)-Dependent Switch in Pathways. To provide a structural interpretation of the SMD nature of \( f \)-induced unfolding of src SH3, we followed the changes in several variables describing the conformations of SH3 as force applied to residues 9 and 59 is varied in the SOP-SC simulations. Most of these are derived from measures assessing the extent to which structures of various parts of the protein overlap with the conformation in the native state. The structural overlap \( \chi_{AB} \) for two parts of the protein \( A \) and \( B \) is the fraction of broken native contacts between \( A \) and \( B \) (30),

\[
\chi_{AB}(r) = \frac{1}{M_{AB}} \sum_{\langle i,j \rangle A,B} \Theta(|r_i - r_j| - |r_i^0 - r_j^0| - \Delta),
\]

where the summation is over the coarse-grained beads belonging to the parts \( A \) and \( B \), \( M_{AB} \) is the number of contacts between \( A \) and \( B \) in the native state, \( \Theta(x) \) is the Heaviside function, \( \Delta \approx 2 \AA \) is the tolerance in the definition of a contact, and \( r_i^0 \) and \( r_j^0 \), respectively, are the coordinates of the beads in a given conformation \( \{r_i\} \) and the native state. Two of the most relevant sets of contacts in the forced rupture of SH3 are the ones between the N-terminal (\( \beta_4 \)) and C-terminal (\( \beta_5 \)) \( \beta \)-strands (Fig. 2A), computed using the structural overlap \( \chi_{\beta_4\beta_5} \) and contacts between the RT loop (residues 15–31) and the protein core (strands \( \beta_1 \) and \( \beta_2 \), residues 42–57) quantified by \( \chi_{\text{RTL}} \). When these structural elements unravel the structural overlap values become close to unity, signaling the global unfolding of the SH3 domain.

Depending on \( f \), in some trajectories the RT loop ruptures from the protein first (\( \chi_{\text{RTL}} \) sharply approaches 1), followed by the break between \( \beta_4 \) and \( \beta_5 \) strands (\( \chi_{\beta_4\beta_5} \approx 1 \)). In other trajectories, the order is opposite, with \( \beta_4\beta_5 \) sheet melting first, without the RT-loop rupture (Figs. 3 and 4 and Fig. S1). We calculated the fraction, \( P_{\text{RTL}} \), of trajectories that unravel through rupture of the RT-loop pathway depends strongly on force, suggesting that these are the two major pathways responsible for the change in the slope of the \( \log(k_f(f),f) \) plot (Fig. S5C). At low forces (15 pN) \( P_{\text{RTL}} \approx 0.8 \), implying that \( \approx 80\% \) of the trajectories unfold through the RT-loop pathway, and this fraction decreases monotonically to \( P_{\text{RTL}} \approx 15\% \) at 45 pN. Movies S1 and S2 illustrate the two unfolding scenarios.

**Effect of Cysteine Cross-Linking.** To further illustrate that the slope change in Fig. 2 is due to the switching of the unfolding routes between the particular pathways discussed above, we created an in silico mutant by adding a disulfide bond between the RT loop and \( \beta_2 \) (mimicking a potential experiment with L24C/G54C mutant). In the cross-link mutant, the enhanced stability of the RT loop to the protein core blocks the \( \Delta \text{RTL} \) unfolding pathway. We generated six 1,500-ms unfolding trajectories at 15 pN and did not observe unfolding in any of them, thus obtaining an estimate for the upper bound of unfolding rate of \( \approx 0.6 \text{ s}^{-1} \) for this mutant. Comparing this unfolding rate to the rate at 15 pN for the WT (without the disulfide bridge) of 5.2 s\(^{-1} \) shows that blocking of the pathway decreases the average unfolding rate at 15 pN. The mutant simulations with the disulfide bridge suggest that the RTL pathway plays a major role at low forces, and the unfolding through the \( \beta_4\beta_5 \) pathways is much slower at low force. Furthermore, these simulations also show that rupture of the protein through the \( \beta_4\beta_5 \) pathway occurs at a very slow rate at low forces even when the unfolding flux along the RTL pathway is muted. Taken together these simulations explain the structural basis of rupture in the two major unfolding pathways.

**Pathway Switch Occurs at a Lower Force in V61A Mutant.** To examine the effect of point mutations, we calculated \( k_f(f) \) as a function of \( f \) for the V61A mutant. In the laser optical trapping (LOT) experiments, V61A mutant does not show upward curvature in the same force range, and the \( \log(k_f(f),f) \) plot in that range is linear. However, the curvature can be seen at lower forces. In simulations, we observe the same qualitative change with respect to the WT upon mutation (Fig. 6 and Table S1). If only data for forces above 15 pN are taken into account, the single-exponential model becomes slightly more likely than the double-exponential model, but inclusion of the lower forces data shows double exponential, with pathway switching occurring at a lower force than for WT. The fraction of trajectories going through the RT-loop pathway

**Fig. 4.** Histograms of unfolding trajectories at 15 and 45 pN using a 2D map with the global order parameter (\( \chi \)) and the distance between Leu24 and Glu52 side chains, which shows whether the RT loop interacts with the protein core (Fig. 2A). The profile offers an alternative view of the one in terms of \( x_f \) and \( x_{\text{RTL}} \) (Fig. 3).

**Fig. 5.** (A) A snapshot of SH3 with the intact contacts between the RT loop and the core but with ruptured \( \beta_4 \) and \( \beta_5 \) strands (state \( \Delta \text{RTL} \)). (B) A snapshot with the RT loop peeled off, but \( \beta_4\beta_5 \) intact (state \( \Delta \text{RTL} \)). (C) The flux of molecules through the \( \Delta \text{RTL} \) state as a function of force. The blue symbols are obtained from simulations. The black line, which is the prediction using the double-exponential fit of the \( \log(k_f(f),f) \) plot with Eq. 11, differs from direct calculations based on simulations.
Before the unfolded state starting from \(x_L\) or \(x_H\). If \(x_L\) and \(x_H\) truly correspond to distances to transition states then \(P_{\text{fold}} \approx 0.5\) (32), that is, the TSE should correspond to structures that have equal probability of reaching folded or unfolded state, starting from \(x_L\) or \(x_H\). In sharp contrast to this expectation, the states with \(x_L\) are visited hundreds of times before unfolding (Fig. S2), which means \(P_{\text{fold}}(x_L) \approx 1\). Thus, the usual interpretation of \(x_L\) or \(x_H\) ceases to have physical meaning, which is a consequence of the strong multidimensionality of the unfolding landscape of SH3.

**Force-Dependent Movement of the TSE.** The results in Fig. S2 show that the extracted values of \(x_L\) and \(x_H\) cannot represent the TSE. Because the underlying reaction coordinates for the inherently SMD nature of folding landscapes are difficult to guess, the TSE can only be ascertained with a method that does not use a pre-determined form of the reaction coordinate. We use the \(P_{\text{fold}}\) based on the theory that the TSE should correspond to structures that have equal probability \(P_{\text{fold}} \approx 0.5\) of reaching the folded or unfolded state. To determine the TSEs in our simulations, we picked the putative transition-state structures from the saddle point of the 2D \((x, E)\) histogram of the unfolding trajectories \([0.66 < x < 0.73; 82 < E < 105\) (kcal/mol) for \(f = 15\) pN and \(0.59 < x < 0.73; 82 < E < 106\) (kcal/mol) for \(f = 45\) pN], where \(E\) is the total energy of the protein. We ran multiple trajectories from each of the candidate TS structures, noting when the trajectory reaches the folded or the unfolded state first, to determine the \(P_{\text{fold}}\). The set of structures with \(0.4 < P_{\text{fold}} < 0.6\) is identified as the TSE. The \(P_{\text{fold}}\) value for the whole ensemble is the total number of trajectories starting from all of the candidate structures that reach the folded state first, divided by the total number of trajectories (or, the average of the individual \(P_{\text{fold}}\) values).

The TSEs for 15 pN and 45 pN are given in Fig. 8. For both sets the \(P_{\text{fold}} \approx 0.5 \pm 0.07\). The low-force TSE shows that the RT loop is disconnected from the core (\(\Delta \text{RTL}\) state) and the 45 pN TSE has structures where the loop interacts with the core, but the contacts between N- and C-terminal \(\beta\)-strands are broken. The explicit TSE calculations confirm that the TSEs are similar to those found in unfolding trajectories with \(x_{\text{RTL}}\) and \(x_{\text{\beta\beta\beta}}\).

The experimental analysis of transition states of SH3 using mechanical \(\Phi\)-values (17) suggests that in the high-force pathway decreases compared with the WT [i.e., \(P_{\text{WT}}^{\Delta \text{RTL}}(f) < P_{\text{WT}}^{\Delta \text{RTL}}(f)\) for all \(f\)] (Fig. 6C). The loss of upward curvature in the force range above 15 pN can be explained by the more prominent role of the \(\beta4\beta5\) pathway at low forces, leading to lesser degree of switching between the pathways. The V61A mutation is in the \(\beta5\) strand, making interactions between \(\beta4\) and \(\beta5\) weaker, thus enabling the sheet to rupture more readily. Parenthetically we note that this is a remarkable result, considering that in the SOP-SC force field is only minimal, which further illustrates that our model also captures the effect of point mutations.

**Free Energy Profiles and Transition States.** Let us assume that the free energy landscape projected onto extension as the reaction coordinate accurately captures the \(f\)-dependent unfolding kinetics. In this case, we expect the Bell model or its variation would hold, and \(x\) (assumed to be \(f\)-independent) obtained from the fitting to that model would be the distance to the transition state with respect to the folded state \(x_1\). If the underlying free energy landscape were SMD it is still possible to formally construct a 1D free energy profile using experimental (18) or simulation data. It is tempting to associate the distances in the projected 1D profiles with transition-state locations with respect to the folded state, as is customarily done in analyzing single-molecule force spectroscopy (SMFS) data. Such an interpretation suggests that \(x_L\) and \(x_H\) should correspond to the distances to the two transition states in the two pathways, with \(x_1\) increasing with force in an apparent anti-Hammond behavior. To assess whether this is realized, we constructed 1D free energy profiles (of the WT protein) at forces 15, 30, and 45 pN to determine \(x_1\). It turns out that \(x_1\) decreases rather than increases with force, demonstrating the normally expected Hammond behavior (Fig. 7), as force destabilizes the native state (21, 31) (Discussion).

We now demonstrate that \(x_L\) and \(x_H\) cannot be identified with transition-state locations by calculating the committor probability, \(P_{\text{fold}}\) (32), the fraction of trajectories that reach the folded state
the important residues are Phe-10 and Val-61 (which are in the β4 and β5), along with a core residue Leu-44. For the bulk (low/zero force) pathway, Phe-10, Ile-56, and Val-61 are also apparently important in TSE, as is the RT-loop residue Leu-24, which interacts with the protein core. Our simulation results, which provide a complete structural description of the TSEs, support the experimental interpretation, namely, loss of interaction between the RT loop and the core at low forces and rupture of the β4/β5 sheet at high forces.

It is interesting to compute the mean extensions of the two major TSEs. The average distance between force application points for these structures is $x_{F}^{a} = 2.25 \pm 0.2$ nm for 15 pN and $x_{F}^{b} = 2.7 \pm 0.3$ nm for 45 pN, which (given the distances in the folded state of $x_{F}^{a} = 1.96$ and $x_{F}^{b} = 2.05$ nm) translates to the transition states of $x_{F}^{a} = 0.29 \pm 0.18$ and $x_{F}^{b} = 0.65 \pm 0.33$ nm, respectively. These values have no relation to $x_{H}$ and $x_{HT}$, further underscoring the inadequacy of using Eq. 11 to interpret $[\log k(f), f]$ plots in SMD.

**Discussion**

**Hammond Behavior.** Protein folding could be viewed using a chemical reaction framework. Just like in a chemical reaction, transitions occur from a minimum on a free energy landscape (corresponding to reactant or unfolded state) to another minimum (corresponding to a product representing the folded state, or an intermediate) by crossing a free energy barrier. The top of the free energy barrier corresponds to a transition state.

Besides determining the structures of the unfolded and folded states, one of the main goals in protein folding is to identify the TSE and characterize the extent of its heterogeneity. When viewed within the chemical reaction framework, the Hammond postulate provides a qualitative description of the structure of the transition state if it is unique. The Hammond postulate states that, if two states, as, for example, a transition state and an unstable intermediate, occur consecutively during a reaction process and have nearly the same energy content, their interconversion will involve only a small reorganization of the molecular structure” (33). A corollary of the Hammond postulate is that the TS structure likely resembles the least stable species in the folding reaction.

To apply the Hammond postulate to a protein free energy landscape, perturbed by $f$, let us assume that at $f = f_{0}$, the states $F$ and $U$, with equal free energy, are separated by a transition state. Increasing $f$ will generally destabilize $F$, and lower the free energy of $U$. According to Hammond’s postulate, the transition state should be more similar to $F$ than $U$ as $f$ increases. If $f < f_{0}$, then the free energy of $F$ will be lower than $U$, and consequently the transition state will be more $U$-like. As a consequence of this argument, in unfolding induced by force, the transition state should move toward the state that is destabilized by $f$ (31), in accord with Hammond behavior. If the opposite were to happen it could be an indication of anti-Hammond behavior.

In a 1D energy landscape, the distance between a minimum and a barrier is reflected in the slope of the $[\log k(f), f]$ plot (an $m$ value for $[\log k(f)], [C]$), which follows from the Arrhenius law and linear coupling in the free energy. Hammond behavior for the unfolding rate would mean movement of the transition state toward the folded state resulting in the decreasing of the slope of the $[\log k_{U}(f), f]$ plot with $f$. Hence, the temptation to refer to the opposite change of slope (i.e., increasing with $f$) as anti-Hammond behavior is natural. However, because the increase of the slope of the $[\log k_{U}(f), f]$ plot necessarily means that the energy landscape is SMD, referring to movement of the transition state along a single reaction coordinate is not meaningful. Hence, the term “anti-Hammond” behavior in this case does not reflect the opposite of the Hammond postulate in either the original formulation or movement of the transition state. Moreover, even if the energy landscape is formally projected onto the reaction coordinate to which the parameter ($f$ or $[C]$) is coupled (which is possible even in the SMD case albeit without much physical sense), the movement of the transition state on this formal 1D landscape will still obey the Hammond postulate. Such a conclusion follows from a Taylor expansion of the first derivative of the perturbed (by $f$ or $[C]$) free-energy profile $F_{f}(x) = F_{0}(x) - f \hat{x}$ around the barrier top ($\hat{x}$).

$$F_{f}(x)|_{\hat{x}} = (F_{0}(x) - fx)|_{\hat{x}} = F_{0}(x_{f}^{0} + (x_{f} - x_{f}^{0})) - f =$$

$$= F_{0}(x_{f}^{0}) + F_{0}^{f}(x_{f}^{0})(x_{f} - x_{f}^{0}) - f = 0,$$

where $F_{0}(x)$ and $x_{f}^{0}$ are the free energy profile and transition-state position at $f = 0$. Because $F_{0}(x_{f}^{0}) = 0$ and $F_{0}^{f}(x_{f}^{0}) = 0$, we find $x_{f} = x_{f}^{0} = f/F_{0}^{f}(x_{f}^{0}) < 0$, or $x_{f} < x_{f}^{0}$, establishing that the transition state moves toward the native state, in accord with the Hammond behavior. Our conclusion holds for any perturbation $f$ that is linearly coupled to the energy function, and that monotonically destabilizes the folded state. Thus, we surmise that upward curvature in $[\log k_{U}(f), f]$ or $[\log k_{U}(f)], [C]$ plots are not equivalent to anti-Hammond behavior. We note here, though, that the linear coupling of $f$ to the protein Hamiltonian is exact and the perturbation by denaturant is approximate, although the leading order in $[C]$ is linear.

A similar conclusion, that is, a connection between upward curvature and multidimensionality, has been drawn analytically before, in the context of mechanochemistry of small molecules, based on the Taylor expansion of the Bell’s model, similar to Eq. 13 (34, 35). In our work, we started from the most general description rather than from the solution of the Kramer’s problem. The WMD conditions are similar to 1D assumptions when obtaining the Bell’s model, but we do not make any assumptions about the barriers. We also solved directly for the quantity we are interested in, that is, sign of $\frac{d^{2}}{df^{2}} \log k_{U}(f)$, rather than movements of the transition state. Connecting the latter to the curvature of the rate requires some additional steps, which might require more assumptions.
Pathway Switch and Propensity to Aggregate. In our previous work (36) we showed that an excited state \( N^* \) in the spectrum of monomeric src SH3 domain has a propensity to aggregate. The structure of the \( N^* \), which is remarkably close to the very lowly populated structure for Fyn SH3 domain determined using NMR (37), has a ruptured interaction between \( \beta \) and \( \beta \). In other words, the value of \( \chi_{\beta\beta} \) is large. Interestingly, in our simulations unfolding of src SH3 domain occurs by weakening of these interactions at high forces (Fig. 3 and Fig. S1). Thus, the \( N^* \) structures are dominant at high forces. Because the probability of populating of \( N^* \) is low at low forces (\( N^* \) has 1–2% probability of forming at \( f = 0 \) (36, 37)) it follows that SH3 aggregation is unlikely at low forces but can be promoted at high forces. Thus, SH3 domains have evolved to be aggregation-resistant, and only under unusual external conditions they can form fibrils.

Prediction for a Switch in the Force Unfolding of I27. Based on our theory and simulations we can make a testable prediction for forced unfolding of I27. Because there is upward curvature in the denaturation-induced unfolding of I27 (8), we predict that a similar behavior should be observed for force-induced unfolding as well. In other words, there should be a switch in the pathway as the force used to unfold I27 is changed from a low to a high value. It is likely that this prediction has not been investigated because mechanical unfolding of I27 has so far been probed using only atom force microscopy (38), where high forces are used. It would be most welcome to study the unfolding behavior of I27 using LOT experiments to test our prediction.

Conclusions

We have proven that upward curvature in the unfolding rates as a function of a perturbation, which is linearly coupled to the energy function describing a protein in a solvent, implies that the underlying energy landscape is strongly multidimensional. The observation of upward curvature in the \( \log k_u(f) \) plots also implies that unfolding occurs by multiple pathways. In the case of denaturant-induced unfolding of SH3 domain this implies that there is a continuous decrease in the flux of molecules that reach the unfolded state through the low-force pathway as \( f \) increases. The numerical results using model 2D free energy profiles allow us to conjecture that if a protein folds by parallel routes then the unfolding rate as a function of the linear perturbation must exhibit upward curvature. Only downward curvature in the \( \log k_u(f) \) plots can be interpreted using a single-barrier 1D free energy profile with a moving transition state or one with two sequential barriers (39, 40).

Our study leads to experimental proposals. For example, Förster resonance energy transfer (FRET) experiments especially when combined with force would be most welcome to measure the flux through the two paths identified for src SH3 domain. Our simulations suggest that the FRET labels between the RT loop and the protein core should capture the pathway switch, provided there is sufficient temporal resolution to observe the state with the RT loop unfolded. A more direct way is to block the RT-loop pathway with a disulfide bridge between the RT loop and the core, as we demonstrated using simulations, and assess if the unfolding rate decreases dramatically. Our work shows that the richness of data obtained in pulling experiments can only be fully explained by integrating theory and computations done under conditions that are used in these experiments.

Materials and Methods

Force-Dependent Rates for SH3 Domain Using Molecular Simulations. The 56 residue Gallus gallus src SH3 domain from Tyr kinase consists of five \( \beta \) strands (PDB ID code 1SRL), which form \( \beta \)-sheets comprising the tertiary structure of the protein (Fig. 2A). Residues are numbered from 9 to 64. The details of the SOP-SC model are described elsewhere (36). A constant force is applied to the N-terminal end (residue 9) and residue 59 (Fig. 2A). We used Langevin dynamics, in the limit of high friction, to compute the rate-dependent unfolding rates. We covered a range of forces from 12.5 pN to 45 pN generating between 50–100 trajectories at each force. From these unfolding trajectories, we calculated the first time the protein unfolded (\( t_{\text{u}} \approx 5 \text{ nm}, \text{SOP-SC} \)) plots also implies that unfolding occurs by multiple pathways. In the case of denaturant-induced unfolding of SH3 domain this implies that there is a continuous decrease in the flux of molecules that reach the unfolded state through the low-force pathway as \( f \) increases. The numerical results using model 2D free energy profiles allow us to conjecture that if a protein folds by parallel routes then the unfolding rate as a function of the linear perturbation must exhibit upward curvature. Only downward curvature in the \( \log k_u(f) \) plots can be interpreted using a single-barrier 1D free energy profile with a moving transition state or one with two sequential barriers (39, 40).

Our study leads to experimental proposals. For example, Förster resonance energy transfer (FRET) experiments especially when combined with force would be most welcome to measure the flux through the two paths identified for src SH3 domain. Our simulations suggest that the FRET labels between the RT loop and the protein core should capture the pathway switch, provided there is sufficient temporal resolution to observe the state with the RT loop unfolded. A more direct way is to block the RT-loop pathway with a disulfide bridge between the RT loop and the core, as we demonstrated using simulations, and assess if the unfolding rate decreases dramatically. Our work shows that the richness of data obtained in pulling experiments can only be fully explained by integrating theory and computations done under conditions that are used in these experiments.

Maximum Likelihood Estimation. For the set of \( M \) constant forces \( f_i \), with \( N_i \) measurements of the unfolding time at each force, assuming exponential distribution of unfolding times \( P(t) = k \exp(-kt) \), where the rate \( k \) depends on the force, the log-likelihood function is

\[
L = \log \prod_{j=1}^{M} \prod_{i=1}^{N_j} k(f_i) e^{-k(f_i)t_i} = \sum_{i=1}^{M} \sum_{j=1}^{N_i} \log k(f_i) - k(f_i)t_i.
\]

In the above equation \( t_i \) is the unfolding time measured in the \( j \)-th trajectory at force \( f_i \). The exponential distribution allows us to take the sum over \( j \) and use the average unfolding time \( t_i = \sum t_i/N_i \) for each force,

\[
L = \sum_{i=1}^{M} \sum_{j=1}^{N_i} \log k(f_i) - k(f_i)t_i.
\]

For each of the models (single- and double-exponential) the log-likelihood function \( L \) was numerically maximized with the set of data \( \{f_i, N_i, t_i\} \) (from simulations or experiment). The two maximal values of \( L \) (for each model) were plugged into the Akaike information criterion (25) to calculate relative likelihood of the models, that is, the ratio of probabilities that the data are
described by each of the models. The parameters that maximize $L$ are used for fitting the $[\log k_i(t, f)]$ plots.

**Akaike Information Criterion.** The lower value of $AIC = 2n - 2L$ (where $L$ is the log-likelihood function and $n$ is the number of parameters in the model) indicates a more probable model, with the relative likelihood of models with $AIC_1$ and $AIC_2$ given by $P_1/P_2 = \exp(AIC_2 - AIC_1)/2$. Thus, for the comparison of Bell's model ($L = L_1$; $n = 2$) and double-exponential model ($L = L_2$; $n = 4$), the double exponential is more probable by a factor of $P_2/P_1 = \exp(L_2 - L_1 - 2)$, where $L_1$ and $L_2$ are found by maximizing $L$ in Eq. 16.

**ACKNOWLEDGMENTS.** This work was supported by National Science Foundation (NSF) Grant CHE 13-61946 and National Institutes of Health (NIH) Grant GM 089685. S.M. thanks the NIH and NSF for support.

---