2.3.5. Classical kinetic modeling of protein folding/unfolding

In a strict sense, folding/unfolding kinetics should be analyzed by statistical, or stochastic (time-dependent statistical) methods and models. These are *molecular* models. Their methodologies and applications will be presented in § 3. Yet, a common approach has been to analyze and interpret the observed behavior using classical chemical kinetics methodologies. These provide us with conceptually simple kinetic schemes and rate parameters, which could be useful from engineering points of view. A brief summary of simple kinetic models, and corresponding integrated rate laws will be presented here, with reference to proteins whose folding/unfolding kinetics obey one or another scheme.

2.3.5.1. Two-state transition

The simplest type of transition between states U and N is a two-state process, given by the scheme

$$U \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N \qquad (2.V)$$

The differential rate expressions holding in this case are

$$\begin{aligned} d[U]/dt &= - k_f [U] + k_u [N] \\ d[N]/dt &= + k_f [U] - k_u [N] \end{aligned} \qquad (2.3.10)$$

where [U] and [I] are the instantaneous (time-dependent) concentrations of the unfolded and folded conformations, respectively, and k_f and k_u are the folding and unfolding rate constants. Let the initial concentrations be $[U]_0$ and $[N]_0$. In folding experiments, we take $[N]_0 = 0$, and the instantaneous concentration [N] is given by

$$[N] = [U]_0 - [U] \qquad (2.3.11)$$

such that the first expression in equation (2.3.10) reduces to a non homogeneous, first order differential equation

$$d[U]/dt = - (k_f + k_u) [U] + k_u [U]_0 \qquad (2.3.12)$$

the solution of which is

$$[U] / [U]_0 = \frac{k_u}{k_f + k_u} + \frac{k_f}{k_f + k_u} \exp \{-(k_f + k_u)t\} \qquad (2.3.13)$$

A few points deserve attention in this result. First, [U] decreases to an equilibrium value of $[U]_\infty = [U]_0 k_u / (k_f + k_u)$ at long times. The fractional concentration of the two states in equilibrium are given by

$$\begin{aligned} f_U + [U]/([U]+[N[) &= k_u / (k_f + k_u) \\ f_N + [N]/([U]+[N[) &= k_f / (k_f + k_u) \end{aligned} \qquad (2.3.14)$$

Second, this decrease (or the increase in [N]) is single exponential with an apparent rate constant

$$k_{app} = k_f + k_u \qquad (2.3.15)$$

Finally, the equilibrium ($t \dashrightarrow \infty$) concentration of folded conformations obeys the detailed balance equation

$$[U]_\infty k_f = [N]_\infty k_u \qquad (2.3.16)$$

The equilibrium concentration also define the equilibrium constant for the folding reaction

$$K_{UN} = \frac{[N]_\infty}{[U]_\infty} = \frac{k_f}{k_u} \qquad (2.3.17)$$

The equilibrium constant is related to the free energy of unfolding by the equation

$$\Delta G_{UN} = - RT \ln K_{UN} \qquad (2.3.18)$$

Several single domain proteins obey a two-state transition. The folding/unfolding kinetics of these proteins on a macroscopic scale is therefore governed by the above set of equations.

2.3.5.2. Sequential transition from U to N

The transition from U to N has been shown in numerous examples above to proceed through the formation of one or more intermediates. The corresponding kinetic schemes can be written as 2.II or 2.III. Let us consider here the simpler case of a single intermediate. Let k_{XY} designate the rate constant for the passage from state X to state Y. Using this notation, scheme 2.II can be rewritten as

$$U \underset{k_{IU}}{\overset{k_{UI}}{\rightleftharpoons}} I \underset{k_{NI}}{\overset{k_{IN}}{\rightleftharpoons}} N \qquad (2.II)$$

The set of equations for the differential change in concentration become in these case

$$d[U]/dt = - k_{UI}\,[U] + k_{IU}\,[I] \qquad (2.3.19a)$$

$$d[I]/dt = + k_{UI}\,[U] - k_{IU}\,[I] - k_{IN}\,[I] + k_{NI}\,[N] \qquad (2.3.19b)$$

$$d[N]/dt = + k_{IN}[I] - k_{NI}\,[N] \qquad (2.3.19c)$$

These can be conveniently written in a matrix equation formalism as

$$\begin{bmatrix} d[U]/dt \\ d[I]/dt \\ d[N]/dt \end{bmatrix} = \begin{bmatrix} -k_{UI} & k_{IU} & 0 \\ k_{UI} & -k_{IU}-k_{IN} & k_{NI} \\ 0 & k_{IN} & -k_{NI} \end{bmatrix} \begin{bmatrix} [U] \\ [I] \\ [N] \end{bmatrix} \qquad (2.3.20)$$

In concise notation, equation 2.3.20 reads

$$d\mathbf{X}(t)/dt = \mathbf{A}\,\mathbf{X}(t) \qquad (2.3.21)$$

where **X**(t) is the vector of the instantaneous concentrations, and **A** is the matrix of rate constants, shortly referred to as *rate matrix*, of order three in the present case. In § 3.6, a matrix equation similar in form, referred to as *a master equation*, will be adopted for describing the stochastics of the transitions in the space of multiple conformations. Therein, the concentrations are simply replaced by the instantaneous probabilities of the different conformations.

The set of coupled differential equations admit a solution that can be found using analytical methods. However, in this case, and more complex cases where the number of intermediates, and consequently the size of the rate matrix **A** is larger, the solution of the set of differential equation (2.3.19) is conveniently found by matrix algebra methods, using the similarity transformation

$$\mathbf{A} = \mathbf{B}\,\mathbf{\Lambda}\,\mathbf{B}^{-1} \qquad (2.3.22)$$

Here **B** is the matrix of eigenvectors of **A**, and **Λ** is the diagonal matrix of eigenvalues. In the above particular case the matrix **A** admits two nonzero eigenvalues, λ_1 and λ_2, each of which represent the negative of the rate constants apparent in a biexponential time dependence. In other words, the instantaneous concentrations are controlled by the double exponential functions, following the formal solution

$$\mathbf{X}(t) = \mathbf{B}\,\exp\{\mathbf{\Lambda}t\}\,\mathbf{B}^{-1}\,\mathbf{X}(0) \qquad (2.3.23)$$

of equation 2.3.21. Equation 2.3.23 may be rewritten in explicit notation for each state i (X_i = [U], [I]or [N]) as

$$X_i(t) = \sum_k \sum_j B_{ik} \exp\{\lambda_k t\}\; B^{-1}{}_{kj}\, X_j(0) \qquad (2.3.24)$$

where the subscript denote the particular elements of the matrices, or vectors, and the summations are carried over all (in the present case three) elements.

Equation 2.3.24 is similar in form to the multiexponential form generally postulated for describing experimental data. See eq 2.3.2. The empirical rate constants determined in these experiments represent the eigenvalues (negative) of **A**, and the

preexponential (or amplitude) factors are functions of the eigenvectors and initial concentrations of the form $\sum_j B_{ik} B^{-1}_{kj} X_j(0)$ for the kth exponential term.

2.3.5.3. Steady state approximation for the intermediate

Let us explore the result of adopting a *steady-state approximation* for the concentration of intermediate. In this approximation, the rate of change of the concentration of intermediate is zero (d[I]/dt = 0). Using eq (2.3.19b), we obtain

$$[I] = (k_{UI}[U] + k_{NI}[N])/(k_{IU} + k_{IN}) \tag{2.3.25}$$

for the constant concentration of intermediates. The rate of unfolding may then be expressed as

$$d[U]/dt = -\frac{k_{UI}\,k_{IN}}{k_{IU} + k_{IN}}[U] + \frac{k_{IU}\,k_{NI}}{k_{IU} + k_{IN}}[N] \tag{2.3.26}$$

In other words, the effective rate constants for the unfolding and folding processes conforming with the two-state scheme 2.V, are

$$k_{u,eff} = \frac{k_{UI}\,k_{IN}}{k_{IU} + k_{IN}}$$

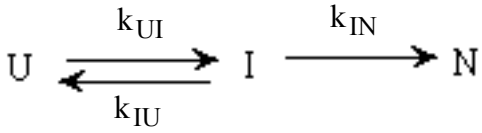
$$k_{f,eff} = \frac{k_{IU}\,k_{NI}}{k_{IU} + k_{IN}} \tag{2.3.27}$$

And the solution of the general scheme 2.II becomes identical in form to eq 2.3.13, with the only exception that $k_{u,eff}$ and $k_{f,eff}$ replace k_u and k_f, respectively.

2.3.5.4. Sequential transition with a pre-equilibrium between initial state and intermediate state

Let us consider a special case of the kinetic scheme 2.II: the transition between states U and I is considerably faster than that between states I and N. Furthermore we

assume that there is practically no reverse transition from state N to state I. Under these circumstances, a *pre-equilibrium* is achieved between states U and I, prior to the rate controlling step I-->N. This type of kinetics, represented by the scheme



(2.II')

has been observed to be effective in a large number of proteing folding experiments. In these studies, the U <--> I step is usually reported to be too fast to be probed by stopped-flow methods and only the corresponding amplitude factor can be measured [Baldwin, 1996 #199]. See § 2.3.3 and 2.3.4.

A similar scheme (2.IV), though in opposite direction, has proved useful in interpreting unfolding experiments probed by HD exchange/2-dimensional NMR. In fact, a pre-equilibrium is quite a general phenomenon. It arises whenever the rate of formation of the intermediate and its decay back into reactants are much faster than its rate of formation of the products. The change may be in the folding or unfolding direction; a passage through a fast forming (on-pathway) intermediate is a common observation. The kinetic equations presented below will refer to the scheme (2.II'). Yet, it should be recalled that the same derivations and/or equalities are valid -with the proper substitution of rate constants- for the case of a pre-equilibrium established between states N and I, preceding complete unfolding.

The set of equations 2.3.19 again applies to scheme 2.II', with the substitution $k_{NI} = 0$. However, the existence of a pre-equilibrium between U and I significantly simplifies the solution, in that the rate of formation of the native state can be simply expressed as

$$d[N]/dt = k_{IN}[I] \qquad (2.3.28)$$

Also, taking advantage of the fact that the concentration of the intermediate reaches a pre-equilibrium $[I] = K_{UI}[U]$ before conversion to [N], the same equation may be rewritten as

$$d[N]/dt = k_{IN} K_{UI} [U] = (k_{IN} k_{UI} / k_{IU}) [U] \quad (2.3.29)$$

Usually, K_{UI} is readily found from the measurements of [U] and [I] at the start of folding experiments, -for example, from the extent of decay in fluorescence or CD at the burst stage of folding. In eq 2.3.29, we assume that the rate of conversion of [I] to [N] is too slow to affect the maintenance of the pre-equilibrium. Alternatively, one may take account of the possibility of I leaking away as it forms U, but being rapidly restored by the fast passages U <--> I. In the case, the net rate of change of I is equal to zero, i.e.

$$d[I]/dt = + k_{UI} [U] - k_{IU} [I] - k_{IN} [I] = 0 \quad (2.3.30)$$

or

$$[I] = + k_{UI} [U] / (k_{IU} + k_{IN}) \quad (2.3.31)$$

Equation 2.3.30 represents the steady state approximation for the intermediate I. Substitution of eq 2.3.31 into eq 2.3.28 yields

$$d[N]/dt = k_{IN} k_{UI} [U] / (k_{IU} + k_{IN}) \quad (2.3.31)$$

which reduces to eq 2.3.29, provided that the rate of formation of N is much smaller than the conversion of the intermediate I back to the unfolded state, $k_{IN} << k_{IU}$. In either case, the rate of conversion from the unfolded to the folded state appears as a first order reaction with rate proportional to k_{IN}. Using the fractional population

$$fI = [I]/ ([U]+[I]) \quad (2.3.32)$$

of I amongst the non native conformations and the conservation equation

$$[U]_0 = [U] + [I] + [N] \quad (2.3.33)$$

the rate law 2.3.28 can be rewritten as a non homogeneous first order differential equation of the form

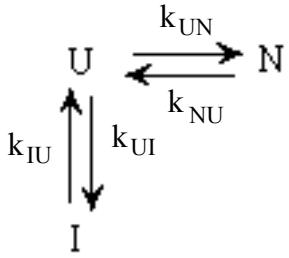
$$d[N]/dt = k_{IN} f_I ([U]_0 - [N]) \quad (2.3.34)$$

The apparent folding rate constant is therefore proportional to k_{IN} and f_I

$$k_f = k_{IN} f_I = k_{IN} k_{UI} / (k_{IU} + k_{IU}) \qquad (2.3.35)$$

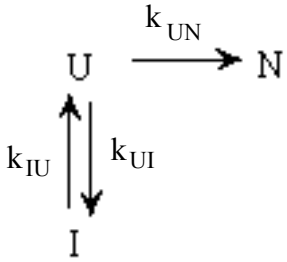
2.3.5.5. Off-pathway intermediate formation

The intermediate formed during folding has been pointed out in a number of experiments to be off-pathway, and slowing down to actual rate of folding. The kinetic scheme



(2.I)

holds in this case. The above scheme can again be solved by the matrix formalism described in eqs 2.3.20-24. For simplicity let us concentrate here on the simplified case of a pre-equilibrium between states U and I, established much faster than the conversion U –-> N. Also, we neglect the back transition from state N to U. These simplifications, expressed as $k_{NU} = 0$, and $k_{UN} << k_{IU}, k_{UI}$, yield the scheme



(2.I')

In this case, the rate of formation of the folded state N is

$$d[N]/dt = k_{UN} [U] \qquad (2.3.36)$$

Using eqs 2.3.32 and 2.3.33 which hold irrespective of the kinetic scheme, the above equation is rewritten as

$$d[N]/dt = k_{UN} (1 - f_I) ([U]_0 - [N]) \qquad (2.3.37)$$

This is a first order nonhomogeneous differential rate equation, in terms of [N], the solution of which is a single exponential rate of formation of [N], with apparent rate constant of

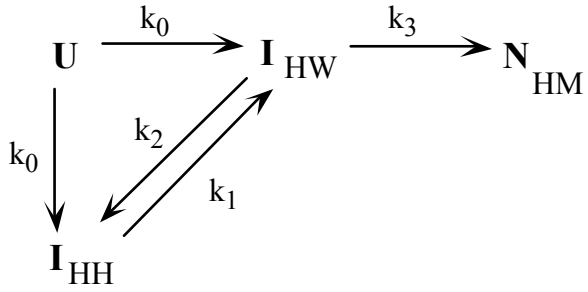
$$k_f = k_{UN} (1-f_I) \qquad (2.3.38)$$

This result deserves special attention. Here, we see that the rate of folding decreases with the extent of formation of the intermediate; whereas eq 2.3.35 indicates the opposite trend, i.e. an increase in folding rate with the extent of intermediate formation. Therefore, accumulation of I has opposite effects in the two schemes (2.I') and 2.II'). It has a productive effect in scheme (2.II'), but a retarding effect in (2.I'). Typical examples in which a slowing down in folding rate has been observed relative to that expected in a two-state process, are the folding of barnase[Matouschek, 1990 #146] and chymotrypsin inhibitor 2. [Jackson, 1991 #260; Otzen, 1994 #227]. The observed slower rates have been attributed to the accumulation of a partially folded intermediate. However, the present analysis demonstrates that the accumulation of intermediate slows down the effective folding rate only if the intermediate is off-pathway. In the case of an on-pathway intermediate, the effect is, on the contrary, an increase in folding rate with the population of intermediate. A typical example illustrating the latter phenomenon is the folding of ubiquitin analyzed in detail by Roder and coworkers [Khorasanizadeh, 1996 #261].

2.3.5.6. More complex kinetic schemes

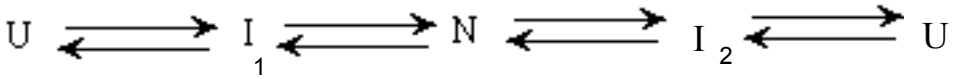
The above examples illustrate how *an apparent single-exponential transition does not necessarily mean the absence of intermediates,* but may result from a pre-equilibrium achieved under certain conditions. Also, intermediates do not always act as kinetic traps, but can have opposite effects depending on their being off-pathway or on-pathway. Finally, the apparent rates in empirical analyses of experimental data, obtained from curve fitting to multiexponential functions, do not represent the microscopic rate constants corresponding to individual steps, but a combination of these rate constants depending on the operating kinetic scheme. All these invite attention to the importance of postulating suitable kinetic schemes for interpreting experimental data. This is what *kinetic modeling* means, in the classical approach.

With the advances in experimental techniques, it becomes apparent that the simplified kinetic schemes described above -though informative and useful from engineering point of view- are too simple. There is a need for more complexes schemes. For example, a kinetic scheme consistent with the folding mechanism of cytochrome *c* displayed in *Figure 2.3.32* is



(2.VI)

Here the slowest and fastest steps are found to have respective rate constants of $k_0^{-1} \leq 100$ μs, and $k_3 \approx 0.03$ s at room temperature. The on-pathway passage U –––> I_{HW} ––> N_{HM} is therefore controlled by the slower step I_{HW} –––> N. The misligated bis-histidine intermediate I_{HH}, on the other hand, cannot be converted to native state, unless a ligand exchange with rate constant k_1 occurs. k_1 and k_2 are reported to be about one order of magnitude smaller than k_3, resulting in the appearance of a much slower folding rate for those species trapped in the misligated form [Yeh, 1997 #259]. The appearance of discontinuous kinetic phases, and the strong sensitivity to ambient conditions is pointed out to cause the overall folding kinetics of cyt c to be extremely complex [Yeh, 1997 #259]. Another example is the urea induced unfolding of barstar [Zaidi, 1997 #262], which exhibits at least two intermediates, and two distinct pathways, each involving a different transition state. The corresponding kinetic scheme can be written as



These complex kinetic schemes suggest that probabilistic or microscopic methods are needed, in conformity with the energy landscape view of folding/unfolding dynamics.

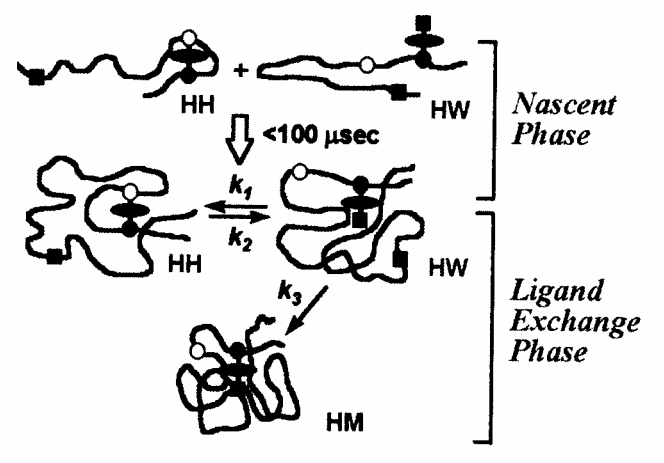


Figure 2.3.32. *Schematic diagram of the folding pathway of cyt c. HH and HW represent the conformations in which the haem group is ligated by two histidines (His 18 and His33/His26), or by one histidine (His18) and a water (W) molecule, respectively. In the nascent phase there is an overall collapse in molecular size. The HW conformation is easily changed into the native conformation (HM for native ligands His18 and Met80) during the ligand exchange phase; whereas the HH conformers (amounting to approximately 40% of the partially compact structures formed at the end of the nascent phase) act as kinetic traps slowing down the completion of the transition to native state. (Figure 9 in[Takahashi, 1997 #258])*

2.3.6. Microscopic modeling of folding kinetics

Deeper insights about protein folding kinetics come from *microscopic* models: statistical mechanical computer simulations{Miller & Dill 1995 6 /id}{Sali, Shakhnovich, et al. 1994 7 /id}{Dinner, Sali, et al. 2000 8 /id}{Pande & Rokhsar 1999 9 /id}{Li, Mirny, et al. 2000 10 /id}. Langevin dynamics of continuum models with different friction coefficients{Veitshans, Klimov, et al. 1997 11 /id}{Klimov & Thirumalai 2000 12 /id}{Thirumalai & Klimov 1999 48 /id} and molecular dynamics (MD) simulations of unfolding{Daggett, Li, et al. 1996 13 /id}{Lazaridis & Karplus

1997 14 /id}{Pande, Grosberg, et al. 1998 15 /id}{Alonso & Daggett 2000 16 /id} or refolding starting from transition states{Pande & Rokhsar 1999 17 /id}.

An exact analysis of folding kinetics cannot be carried due to the astronomically large number of accessible conformations. However, in order to gain an understanding of the microscopic events underlying the observed kinetic behavior, it is possible to perform computations with simple models. An example is the folding of 2-dimensional lattice chains using a Go model for internal interactions. In the Go model, an attractive potential ε is assigned to each native contact. All other contacts have zero interaction energy. To mimic the weakening of hydrophobic interactions by denaturants or temperature, one can vary the ratio ε/kT. High temperatures denature the model proteins, while low temperatures stabilize the folded state, following Boltzmann's law. Folding is initiated by starting with an ensemble in which all conformations are equally probable (infinite temperature limit) and assuming the ambient temperature to be lowered to room temperature at t = 0. The stochastic process of passage from denatured to native state (or an ensemble of configurations in which the native is by far the most stable state) is observed provided that the energy of the native state is significantly lower than that of all other conformations. See Ozkan et al. (2001) for more details.

A complete set (N =802,075 conformations) of all self-avoiding 16-mers on a square lattice were thus generated (Ozkan et al., 2001) to observe the microscopic evolution of all conformations using *a master equation formalism.*. Equations 2.3.21-24 are valid with the redefinition of variables: **P**(t) is the N-dimensional vector of the instantaneous probabilities of the conformations, and **A** is the N x N *transition* (or *rate*) *matrix* describing the rate constants of the transitions between these conformations. By definition, the ijth off-diagonal element of **A** is the rate constant k_{ij} for the passage from conformation j to conformation i. From the principle of detailed balance, $k_{ij}\, P_j^{\circ} = k_{ji}\, P_i^{\circ}$, where P_i° is the equilibrium probability of the ith conformation. The ith diagonal element (k_{ii}) of **A**, on the other hand, represents the overall rate of escape from conformation i. It is found from the negative sum of the off-diagonal elements in the same column, i.e. $k_{ii} = -\,\Sigma_j\, k_{ji}$ (j≠i). The time-dependent probability of occurrence of the ith conformation (i.e. the ith element of **P**(t)) is expressed in terms of the elements of **B**, **Λ**, $\mathbf{B}^{-1}$ and **P**(0) as

$$P_i(t) = \sum_{j=1}^{N}\sum_{k=1}^{N} B_{ik}\exp(\lambda_k t)[B^{-1}]_{kj} P_j(0) = \sum_{j=1}^{N} C(i,t \mid j,0)P_j(0)$$

where $C(i,t \mid j,0)$, denotes the *conditional probability* of (or *transition probability* to) conformation i at time t, given conformation j at $t = 0$. The matrix **C**(t) fully describes the time dependence of NxN *microscopic* transitions. Results can be better understood if analyzed in terms of macroscopic states, each characterized by a given distribution of native contacts. The *time-delayed joint probability* of macroconformations A and B comprising N_A and N_B respective microscopic conformations is

$$P(\mathbf{A}, t_2; \mathbf{B}, t_1) = \sum_{i=1}^{N_A}\sum_{j=1}^{N_B} C(i, t_2 - t_1 \mid \mathrm{j}, 0).P_j(t_1)$$

There are several other studies in this area. Master equation formalisms have been adopted by Scheraga and coworkers{Ye, Ripoll, et al. 1999 28 /id} for analyzing the folding of a subset of 50 conformations (local energy minima) generated for staphylococcal protein A, and by Eaton and coworkers, for modeling the formation of a β-hairpin {Munoz, Henry, et al. 1998 29 /id}{Munoz, Henry, et al. 1998 29 /id}. A well-defined folding pathway was reported{Pande & Rokhsar 1999 9 /id} for a 48-mer on a 3d-cubic lattice, as well as well-defined TS conformations having a common core structure. Likewise, a preferred unfolding pathway was observed by Lazaridis and Karplus{Lazaridis & Karplus 1997 14 /id} in the multiple MD trajectories of chymotrypsin inhibitor 2 (CI2) – a classical example protein that obeys 2-state kinetics - suggesting that a preferred pathway can be compatible with a funnel-like average energy surface, as had been previously noted from lattice model simulations{Miller, Danko, et al. 1992 30 /id}.

2.3.6.1. Rate limiting or rate controlling steps. Apparent activated states

Transition states are ephemeral. These are free energy maxima between two equilibrium states, and the time period spent over such a maximum is too small to be captured by experimental means. In the energy landscape representation, transition states would generally be in the form of a saddle, i.e. a maximum with respect to the transition pathway, and a minimum with respect to transverse directions. Using the energy landscape view, again, one might argue that no single transition state is observable because there are multiple, infinitely many pathways each of which involves the crossing of one or multiple energy barriers. Yet, although at initial stages of folding an infinitely large number of conformations (in the denatured state) are accessible and each may evolve through a distinct route, it is likely that some common dynamic features, preferred from probabilistic points of view, will occur in many pathways, as pointed out above. For example, given the appropriate sequence of residues, it is likely that some segments will undergo an autonomous folding. One or more α-helix formation is a common observation at the burst stage of folding; β-sheet formation, although less common, has also been observed at the earliest folding stage of some proteins (eg. RNase A [Udgaonkar, 1990 #150]). Alternately, at the latest stages, inasmuch as all conformations converge to a unique final state, the native structure, it is highly probably that the pathways will merge, -probably to a MG state close to the native state. The transition from MG to N may or may not be readily established. In fact, many proteins show such a tendency: a rapid conversion to an intermediate structure with MG characteristics, followed by a slow transition to the N state. The last step is slower. This is therefore the rate controlling step.

Therefore in a sequential transition involving one or more probabilistically favorable intermediates -which can be represented by the kinetic schemes 2.II or 2.III given in § 2.3.1.1-, one step, the slowest, will be rate controlling. An activated state can therefore be defined, such as the conformation assumed at the energy maximum to be surmounted during this step. Evidently, a similar -conceptually simpler- definition evidently applies to the single step of an all-or-none transition.

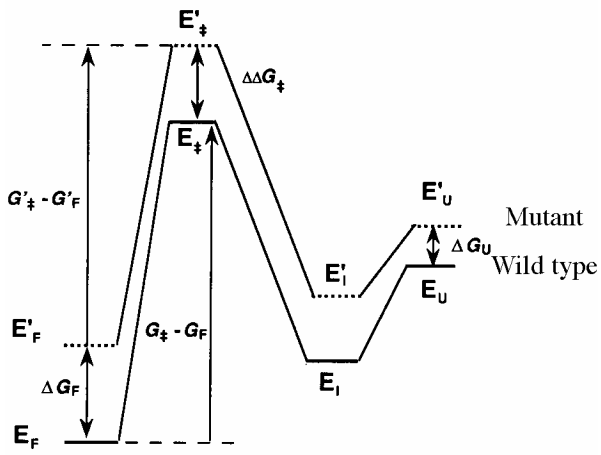
In the case of multiple parallel pathways, on the other hand, the fastest dominates the observed behavior. And if the fastest involves a series of elementary steps, the slowest amongst these control the apparent rate. Thus, for a hypothetical kinetic scheme from the unfolded to the folded state consisting of many parallel pathways each of which involves several steps, the point would be (i) to determine the pathway that involves the lowest energy barriers, and (ii) in this pathway to focus on the highest energy barrier step for an examination of the conformation assumed at the transition state. We may invoke here the concept of *apparent* transition state, since this is simply the most critical one amongst many that can be encountered during the folding process.

Theoretical studies attempting to delineate the most probable pathway are presented in § 3. Again, both classical approaches of reaction kinetics, and modern treatments based on statistical and stochastic models are used to this aim. An elegant modern approach is to use *a master equation formalism* to describe the stochastics of transitions in an ensemble of representative conformations obeying a kinetic scheme involving as many steps as computationally feasible. This approach and its applications for investigating macromolecular dynamics will be presented in § 3.x.

In the next subsection we will present the experimental studies attempting to elucidate the transition state(s) in folding kinetics. An approach based on protein engineering concepts, proposed by Fersht and collaborators, will be presented. This approach is referred to as Φ-value analysis. The use of classical methodologies of chemical kinetics for interpreting kinetic data will be illustrated for a few examples.

2.3.6.2. Effect of point mutations on folding kinetics. Φ-value analysis.

Experiments show that it is, in principle, possible to map out the structures of the transition states, and the intermediates, by studying the folding kinetics of a series of mutants. In these studies, folding is initiated either by a T-jump (in the case of cold denatured proteins), or dilution of denaturant using stopped-flow mixers. A typical example for the former case is the folding of barstar [Nolting, 1995 #183], a peptide inhibitor of barnase; examples for the latter are the folding of chymotrypsin inhibitor 2 (CI2) [Otzen, 1994 #227; Itzhaki, 1995 #228; Fersht, 1995 #229] and barnase [Matouschek, 1990 #146; Matouschek, 1992 #156].

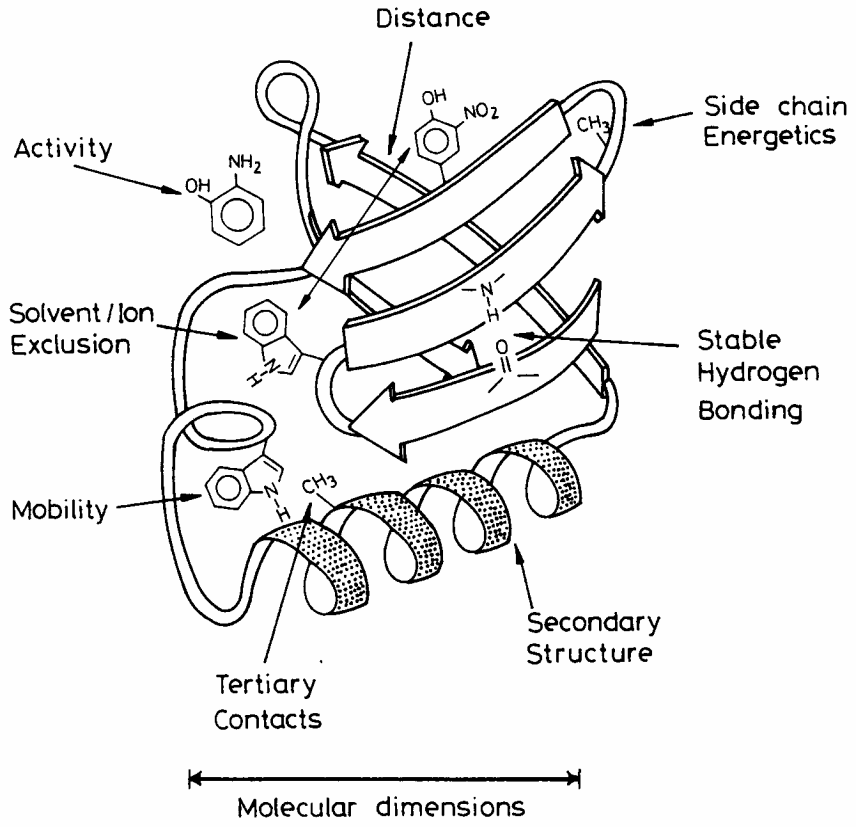


2.3.7. Conclusion: Utility of combining data from different techniques

In the present chapter, we tried to emphasize with several illustrations the need for, and utility of, *jointly* analyzing and interpreting the data from different experimental techniques. Furthermore, we stressed the importance of adopting a statistical viewpoint as a prerequisite for correct interpretation of the data. In principle, a kinetic scheme with infinitely many intermediates communicating via elementary steps is conceivable, and the observed behavior for a particular protein under particular folding or unfolding conditions will necessarily be a simplified

case of this most general statistical behavior. The fact that a two-state transition is observed under given conditions could simply mean that intermediates are not stable enough, or probable enough, to be visible along the transition, in the experimental time window. Yet, these probably do exist as local minima -though relatively shallow under the particular experimental conditions- on the free energy surface. Inasmuch as stability is directly controlled by the depth of free energy minimum, and for a given activated state the latter directly determines the height of the free energy barrier, observed kinetics cannot be abstracted from stability characteristics.

Plaxco and Dobson discussed how complementary methods can be combined to obtain information on the structural and mechanistic aspects of protein folding. There are numerous methods, including NMR, fluorescence, CD, SAXS, mass spectrometry, quasi-elastic light scattering. New developments in many of these techniques, together with improved methods for initiating refolding are now pushing measurements down to nonsecond time regime. These will be leading to a new and more detailed understanding of the protein folding process.


Distance
Activity
Side chain
Energetics
Solvent / Ion
Exclusion
Stable
Hydrogen
Bonding
Mobility
Secondary
Structure
Tertiary
Contacts
Molecular dimensions