Dynamics of hierarchical folding on energy landscapes of hexapeptides

Yaakov Levy, Joshua Jortner, and Oren M. Becker[a)]
Department of Chemical Physics, School of Chemistry, Tel-Aviv University, Ramat Aviv, Tel-Aviv 69978, Israel

(Received 20 August 2001; accepted 13 September 2001)

In this paper we apply the master equation approach to study the effects of the energy landscape topology and topography on the kinetics of folding, and on kinetic transitions of three alanine-hexapeptides analogs which involve polypeptides with neutral and charged groups and a cyclized polypeptide. We rely on the potential-energy landscapes of these molecular systems, which have been constructed using both a topological mapping analysis and a principal component analysis. It was found that the different topology and topography of the energy landscapes result in different ''folding'' time scales and that the systems with geometrical constraints (cyclization and opposite charges at the termini) ''fold'' more slowly than the unconstrained peptide. In addition, for each of the three polypeptide systems, the kinetics is nonexponential at the temperature range 400–600 K. The relaxation kinetics is characterized by logarithmic oscillations, which indicate hierarchical dynamics characterized by multiple time scales of fast (few ps) and slow (few μs) events. At higher temperatures, successive relaxation channels with similar characteristic time scales collapse into a single relaxation channel. While the kinetics of the unconstrained peptide at 600 K can be reasonably well described by a single exponential time scale, the kinetics of the constrained hexapeptides are inherently hierarchical and featured by multiple time scales even at high temperatures. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1415444]

I. INTRODUCTION

The structure and dynamics of peptides and proteins are characterized by their multidimensional complex energy landscapes. Their kinetics, including what is known as the ''protein folding problem,'' is of major interest for both experimental and theoretical studies.[1–6] Theorists promote the idea that the overall kinetics of a polypeptide is a consequence of the multidimensionality and complexity of its underlying energy landscape.[7–13] However, most experiments seem to indicate that a straightforward classical kinetic analysis, based on a small number of states along a one-dimensional reaction coordinate, is sufficient for characterizing the observed protein folding kinetics.[1,14] This kinetics is usually explained in terms of a straightforward two-state kinetic scheme in which a high barrier separates the unfolded state from the native folded one.[1,14] In some cases a limited number of intermediates is introduced along an essentially one-dimensional reaction coordinate.[15,16] Even in cases where the observed kinetics is quite complex, a one-dimensional description is typically employed. For example, a three-state interpretation (Unfolded↔Intermediate ↔Native) was employed to describe the folding kinetics of ubiquitin[17,18] and of cyt c.[19,20] More complicated one-dimensional schemes were employed to describe the kinetics of circularly permuted RNA.[21] Despite the evident success of the simple kinetic scheme, there are however also experimental studies that point to a far more complex kinetic behavior than commonly assumed. Experimental evidence concerning the complexity of the underlying energy landscape for a system with many degrees of freedom can be obtained, for example, by observing a multiplicity of relaxation times.[22–27] The classical studies of Frauenfelder and collaborators[22–25] on the rebinding kinetics of CO to myoglobin described this kinetics in terms of a hierarchy of minima, referred to as ''conformational substates,'' arranged in ''tiers'' corresponding to different conformation energies and barrier heights. The ''substates'' were introduced to explain the different time scales on which certain kinetic phenomena take place in this system. In general, it was assumed that slower phenomena are associated with larger scale motions.

Recently, several experimental studies for peptides and proteins[26] have shown that their kinetics deviates from the simple exponential relaxation, $\exp(-t/\tau)$. Instead, the kinetics were better fitted by a stretched exponential function, $\exp[-(t/\tau)^{\beta}]$ (where $0<\beta<1$), or an asymptotic power-law function $1/(1+t/\tau)^{n}$, (where $n>0$). For example, Gruebele and co-workers[26,27] reported that the refolding kinetics in an aqueous solution of two proteins, the yeast phosphoglycerate kinase and a ubiquitin mutant, is best described by a stretched exponential. Metzler *et al.*[28] found that the S–S recombination in a di-thiotyrosine peptide following bond cleavage is highly nonexponential and can be fitted either to a stretched exponential ($\beta=0.086$) or to an asymptotic power-law ($n=0.331$).

As noted above, during the last decade, the simple kinetic view of protein folding was complemented by a theoretical view, emphasizing the rugged character of the underlying potential energy landscape (PEL) and the resulting heterogeneity of the folding ensemble.[9–11,13,29,30] Wolynes

[a)]Current address: Bio IT (Bio Information Technologies) Ltd., 3 Hayetzira St., Ramat Gan 52521, Israel.

and collaborators[9,10] suggested, through simulations and investigations of model systems, that the overall structure of complex free-energy surfaces, such as those controlling protein folding, is characterized by a multidimensional ''funnel.'' This ''funnel'' represents a built-in bias in the polypeptide's potential energy and free-energy landscapes, steering it efficiently towards its native state. Theoretically, polypeptide energy landscapes have been studied using both simplified models (on-lattice and off-lattice models)[6,7,31,32] and detailed atomistic simulations,[13,33–37] with the resulting landscapes being characterized in a variety of ways using order parameters,[33,34] geometrical measures,[13,37] and topological connectivity mapping.[35–37] In principle, the energy landscape describes all the pathways bridging the unfolded and the folded states, the intermediates along the pathways, and their hierarchy, i.e., which folding channels occur in parallel and which are sequential. Recently, it has been pointed out that the relaxation of a hierarchical system, which is governed by distinct and discrete time scales, can be detected by plotting $d\log(f(t))/d\log(t)$, where $f(t)$ is the relaxation function.[38] Applying this method of analysis for a hierarchical dynamic process results in ''shoulders'' that are frequently called logarithmic oscillations. Each oscillation represents, in the course of time, the contribution of yet another relaxation channel in the decay cascades. Logarithmic oscillations superimposed on relaxation patterns were found for several complex systems, for instance, the kinetics of chloride ions through a channel,[39] and the diffusion process on the Sierpinski gasket.[40]

Protein folding is often considered a hierarchical process.[41,42] The folding of a globular protein involves a sequence of events, each with its characteristic time scale. The kinetic inhomogeneity of the sequential domains is envisioned in terms of a ''downhill folding'' scenario cascading through a hierarchy of protein substates. A sequential dynamics in the reordering of proteins after a ''proteinquake,'' following a hierarchy of protein substates, as well as the existence of equilibrium fluctuations between the protein states, was suggested by Frauenfelder and colleagues.[22–24] The first evidence for logarithmic oscillations has recently been reported for the relaxation of myoglobin after photodissociation.[25]

An approach that is clearly poised to bridge the gap between the averaged macroscopic kinetic experimental observables on the one hand and the detailed microscopic description provided by the energy landscape on the other hand is the master equation approach.[35,43–48] One advantage of this approach is that it can be used to study kinetics on a much longer time scale than the one accessible through direct molecular dynamics simulations. In addition, the master equation describes the relaxation of an ensemble over a range of temperatures without the need for explicit averaging over the separate trajectories. The prerequisite for any master equation study is a knowledge of the underlying energy landscape and of the transition rates between the minima.

The goal of the present study is to characterize the effect of the energy landscape topology and topography on the complex kinetic characteristics of the system. To address this question we rely on the potential energy landscapes of three alanine-hexapeptide analogs, which have recently been constructed using both the topological analysis[36,37] and the principal component analysis.[37] Applying a master equation approach, based on the dual energy landscape analysis, for each of the three hexa-alanine analogs has shown that the kinetics of each peptide is characterized by nonexponential dynamics and unique time scales, which are dictated by the landscape topography and topology.

II. METHODS

A. The molecular systems and their vacuum energy landscapes

Three alanine hexapeptide analogs were studied: (i) alanine hexapeptide with neutral terminal groups (Ala6); (ii) alanine hexapeptide with a positive charge at the N-terminus and a negative charge at the C-terminus (chrg-Ala6); and (iii) a backbone cyclized alanine hexapeptide (cyc-Ala6). Recently, the energy landscapes of these three hexa-alanine analogs were constructed based on conformation samples generated by high-temperature molecular dynamics simulations in a vacuum followed by gradual annealing.[36,37] In order to allow for an efficient landscape analysis, the conformation samples were pruned by removing conformations that are highly similar to other conformations in the sample. The energy landscapes were mapped using two techniques: the principal component analysis, PCA (based only on local minima)[37,49,50] that reflects the landscape's topography, and the topological mapping analysis (based on minima and barrier information)[35–37] that concentrates on the energy landscape topology (i.e., the barrier connectivity).

The PCA method is used to project high dimensional data onto low-dimensional subspaces which contain the ''essential'' information. In the context of macromolecules, the PCA was first used to analyze and visualize molecular dynamics trajectories,[51–56] to obtain and to visualize conformation samples[57–60] and to attempt to perform dynamics in the reduced subspace.[61] This method projects the multidimensional conformation space (represented by a set of conformations) onto a new set of axes that maximizes the variance of the projection along orthogonal directions. In most cases it was found that the optimal 3D projection of the full space held 60%–85% of the information about the distribution, indicating that a 3D visualization of large multidimensional conformation spaces can be obtained while the rest of the information is distributed in small quantities over the other dimensions.[49]

The topological disconnectivity graphs [Figs. 1(a), 1(c), and 1(e)] and the principal component maps of the energy surfaces [Figs. 1(b), 1(d), and 1(f)] for each of the three hexapeptides yielded very similar energy landscape pictures. However, the topographies and topologies of the three landscapes are different. While the energy landscape of Ala6 is characterized by a single broad rough funnel [Figs. 1(a) and 1(b)], the energy landscape topography of chrg-Ala6 is that of a single deep and narrow funnel [Figs. 1(c) and 1(d)], and the energy landscape of cyc-Ala6 is characterized by three competing basins [Figs. 1(e) and 1(f)].

To gain insight into the ''folding'' mechanisms of the

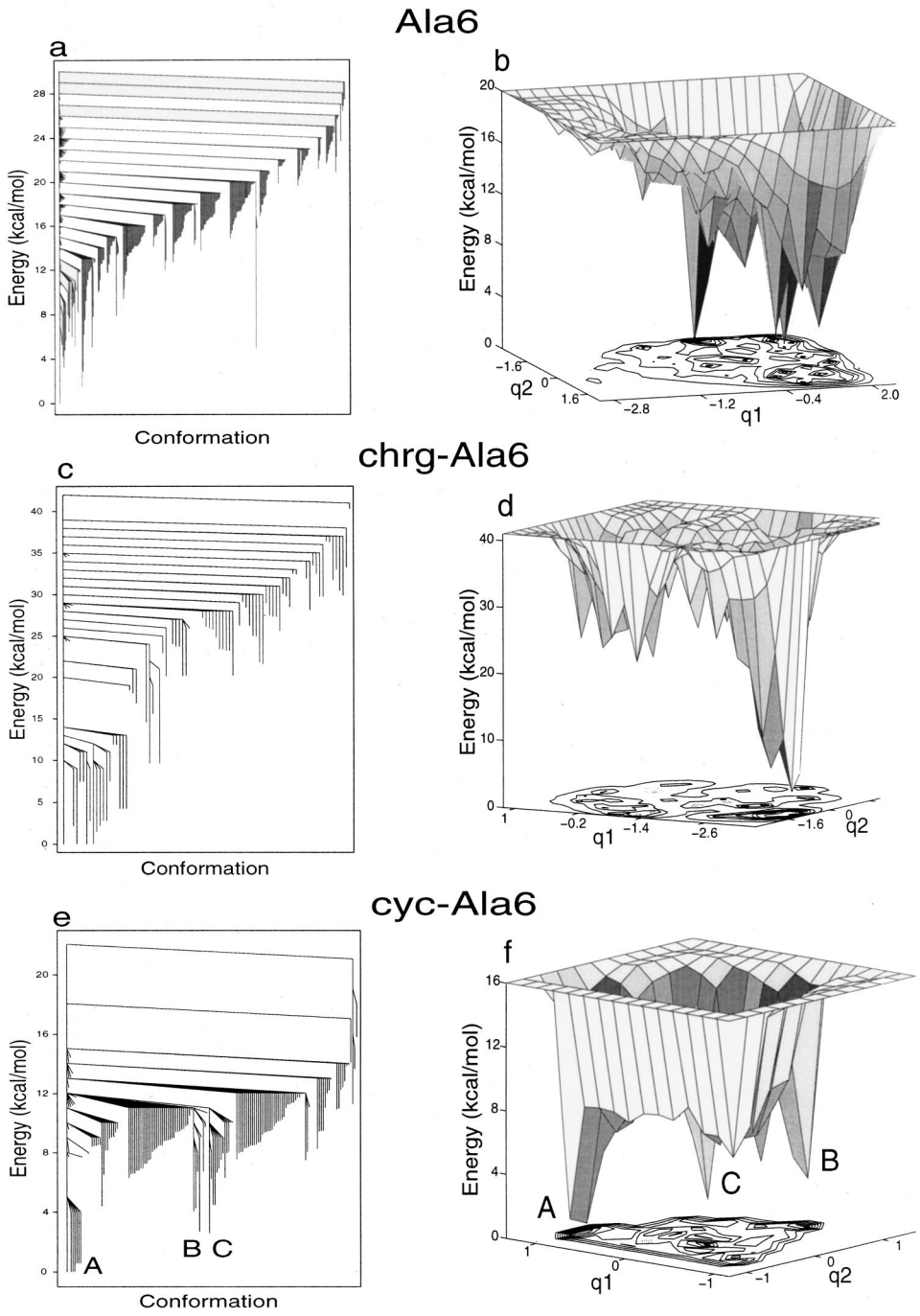

FIG. 1. The energy landscapes obtained by topological mapping and principal component analyses for Ala6 (a and b), chrg-Ala6 (c and d) and cyc-Ala6 (e and f). The graphs indicate a single dominant funnel on the energy landscapes of Ala6 and of chrg-Ala6, and the energy landscape of cyc-Ala6 exhibits three competing basins (marked A, B, and C).

three-polypeptide analogs and into the effects of the conformation constraints, an order parameter ρ was previously defined[37] measuring ''nativeness'' based on the dihedral angles that characterize the native structure. Based on the order parameter as a reaction coordinate it was found that Ala6 exhibits a direct one-step folding, chrg-Ala6 shows a two-step folding, while cyc-Ala6 reveals competing pathways leading from one basin to another, providing information about the effective folding pathways on the complex topography of the landscapes. However, this approach cannot address the central question of the folding kinetics. In what follows, the kinetics of these three molecular systems is studied using the master equation, based on their potential-energy landscapes. In doing so, we assume that the sampled potential surface constitutes a complete representation of each system.

B. The master equation

A powerful approach to the dynamics and kinetics of systems with tens to ten thousands of particles originates from the statistical sampling of linked minima and saddle points on the potential surface of the system. For a system of ten particles, the knowledge of all the minima and saddle points on its potential-energy surface, together with a little additional information about the shape of the potential in the vicinities of these stationary points, is sufficient to construct master equations that reproduce molecular dynamics simula-

tion results at all but the shortest time scales.[45,46,48] For systems of ten or more particles, which are large enough to make a full exploration of the landscapes inconvenient, but not impossible, master equations based only on statistical samples were also able to reproduce the results of molecular dynamics simulations.[62] Recently, the overall kinetics of clusters,[45,46] such as $(KCl)_5$, Ar_9, 13- and 38-atom Lennard-Jones clusters, and of small polypeptides,[35,44,48] e.g., tetra-alanine and octa-alanine, were studied with the aid of master equations, based on a detailed characterization of their energy landscapes.

The master equation represents a fundamental statistical mechanical approach to kinetic transitions among a multitude of states. It is a loss-gain equation that describes the time evolution of the probability $P_i(t)$ for finding the system in a state i.[43] In the language of energy landscapes, local minima represent conformational states of the molecule, i, while saddle points (which represent transition pathways) determine the state-to-state transition rates. The basic form of the master equation, which provides a connection between the topography of the energy landscape and the system's kinetic behavior, is,

$$\frac{dP_i(t)}{dt}=\sum_j\,[W_{ij}P_j(t)-W_{ji}P_i(t)], \tag{1}$$

where W_{ij} is the transition probability from state j to state i. This equation can be rewritten in matrix form $\dot{\mathbf{P}}(t)=\mathbf{W}\mathbf{P}(t)$, where $\mathbf{P}(t)$ is the probability vector at time t, and its formal solution is $\mathbf{P}(t)=\exp(t\mathbf{W})\mathbf{P}(0)$. In this formulation the transition matrix elements are defined as

$$W_{ij}=W_{ij}-\delta_{ij}\left(\sum_k\,W_{ki}\right). \tag{2}$$

The matrix $\mathbf{W}$ has the properties $\mathbf{W}_{ij}\geqslant 0$ for $i\neq j$, and the sum over each column is zero; namely, $\Sigma_i\mathbf{W}_{ij}=0$ for all j. Given the knowledge of the minima and the transition states on the energy landscape it is possible to use the transition state theory (among other methods) to evaluate the transition matrix elements (rate constants k_{ij}), i.e.,

$$k_{ij}\equiv W_{ij}=\frac{kT}{h}\,\frac{Q_{ij}^{\#}}{Q_j}\exp(-E_{ij}^{\#}/kT), \tag{3}$$

where k is the Boltzmann constant, h the Planck constant, Q_j the partition functions of the ''reactant'' state, $Q_{ij}^{\#}$ the partition functions of the transition state, and $E_{ij}^{\#}$ the barrier height measured relative to state j. Histograms of the barrier energies of the three polypeptides (Fig. 2) reveal wide distributions of $E_{ij}^{\#}$ values, with the cyc-Ala6 being characterized by the narrowest distribution ($E_{ij}^{\#}\cong 1-17$ kcal mol^{-1}), while the distributions for Ala6 ($E_{ij}^{\#}\cong 1-32$ kcal mol^{-1}) and for chrg-Ala6 ($E_{ij}^{\#}\cong 1-40$ kcal mol^{-1}) are broader. These wide distributions of barrier heights, which seem to characterize complex systems, such as polypeptides, imply nonexponential kinetics.

To follow the time evolution of the population probability at each minimum, $P_i(t)$, it is necessary to solve the master equation, Eq. (1). This can be done by expanding the probability vector $\mathbf{P}(t)$ in terms of the eigenvectors, $\mathbf{S}_i$, and the eigenvalues, λ_i, of the transition matrix $\mathbf{W}$. The time evolution of the probability vector $P(t)$ can be written as

$$P_i(t)=P_i^{eq}+\sum_k\,C^kS_i^ke^{\lambda_kt} \tag{4}$$

with the sum over k being taken for $\lambda_k<0$.

The coefficients $C^k=[\mathbf{S}^k|\mathbf{P}(0)]$ are determined by the initial distribution, $\mathbf{P}(0)$. Since all $\lambda_k<0$, the equilibrium distribution $\mathbf{P}^{eq}$ is the eigenvector which corresponds to $\lambda_i=0$; i.e., as $t\rightarrow\infty$, $\mathbf{P}(t)\rightarrow\mathbf{P}^{eq}$ and all the other terms decay to zero.

The transition probability W_{ij}, going from a local minimum j to a local minimum i, Eq. (3), is calculated from the barrier $E_{ij}^{\#}$ that separates them,[49,63,64] and from estimates of the partition functions, Q_j and $Q_{ij}^{\#}$. These partition functions were calculated within the harmonic approximation by calculating the normal mode frequencies, υ_l, of each conformation and saddle point using the CHARMM program,[65,66]

$$Q^{\mathrm{Harm}}=\prod_{l=1}^{m}\,\frac{kT}{h\nu_l}, \tag{5}$$

where $m=3N-6$ for each minimum conformation or $m=3N-7$ for each saddle point, while N is the number of atoms.

To trace the folding kinetics for each peptide, the five least stable states with the highest potential energy were initially equally populated. We found that changing the initial

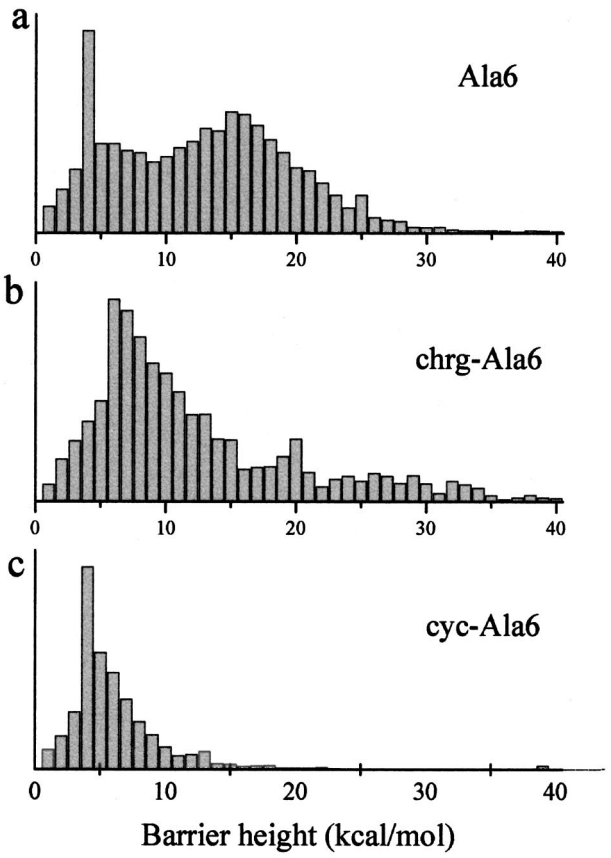

FIG. 2. Histogram plot of the barrier energies of Ala6 (a), chrg-Ala6 (b), and cyc-Ala6 (c).

population to include more than the five least stable local minima, or choosing high-energy local minima, not according to their energy, but according to their connectivity, does not affect the kinetic results. It was found that for chrg-Ala6, at temperatures below 400 K, a number of slightly positive eigenvalues of the transition matrix begins to appear. This numerical artifact can be traced to the small (positive) diagonal elements of **W** at low temperatures. Czerminski and Elber[44] and Wales and co-workers[46] reported similar numerical problems, and therefore restricted their studies to sufficiently high temperatures. Here we concentrate on the kinetics of the three hexa-alanine peptides at a temperature range of 400–600 K.

III. RESULTS

A. Basin population probabilities

Solving the master equation at a given temperature provides the time evolution of the probability for finding the system at any individual state (i.e., conformation) *i*. In principle, it is more informative to focus on the kinetics of transitions between ''basins'' rather than between individual states. These ''basin-to-basin'' transitions[35] are more useful when compared with experimental results, which follow only this type of transitions (e.g., from the ''unfolded'' basin to the ''native'' basin).[1] In a previous study[37] the energy landscapes of the three hexa-alanine analogs were partitioned into disconnected ''basins'' by applying ''topological mapping.''[35] In this context, a basin $R(\alpha)$ is a set of system configurations that map to a single local minimum α by direct minimization[67] or by annealing.[35] ''Topological mapping'' $M^E:\mathbf{R}^{3N-6}\rightarrow\{\alpha'\}^E$ was introduced[35] to map the connectivity between basins as a function of energy (or temperature).

To calculate the accumulated probability of finding the system in a given basin, P_I, the state-to-state results of the master equation are grouped according to the basin partitioning. Namely, P_I is defined as the sum of the probabilities of the individual states *i* that make up basin *I*,

$$P_I(t)=\sum_{i\in I} P_i(t). \tag{6}$$

With this definition of basin probabilities, it is possible to study the effect of the landscape topography on the folding kinetics. In the case of the three hexa-alanine analogs studied here, one would expect that the kinetics on a single basin (single funnel) landscape (Ala6 and chrg-Ala6) will be different from the kinetics on a multiple basin landscape (cyc-Ala6).

Figures 3(a) and 3(b) show the time evolution of the population probability for finding Ala6 and chrg-Ala6, respectively, in the global minimum (GM) basin characterizing their funnel energy landscapes. The population probabilities for each system were calculated at 400 K, starting from an initial population where only the five highest energy conformations were populated. The accumulated population probability of finding the system within the funnel on a ''funnel energy landscape'' is calculated by summing up the prob-

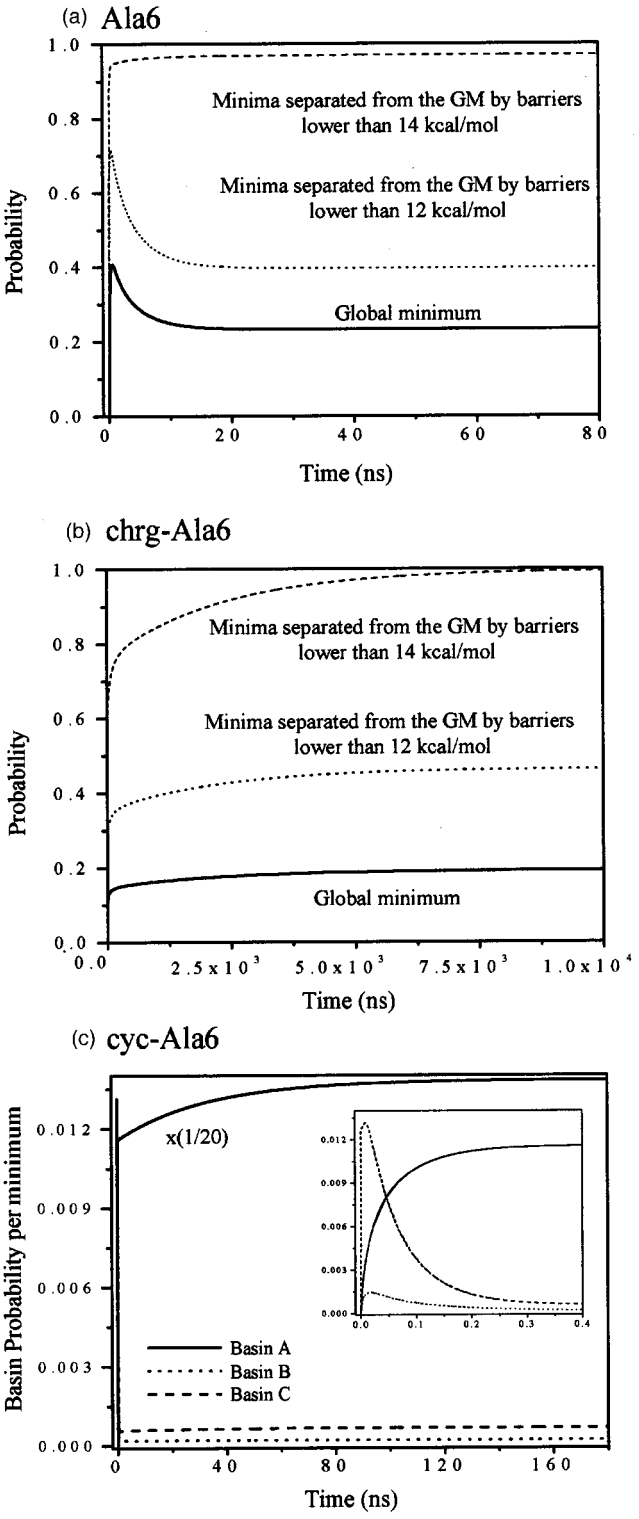

FIG. 3. The time evolution of the population probabilities of finding polyalanines at the global minimum, GM (solid line), and at the funnel bottom (dotted and dashed lines) calculated at 400 K. (a) Data for Ala6. Using the disconnectivity graph [Fig. 1(a)] a conformation is regarded as located at the funnel bottom if the barrier that separates it from the global minimum is smaller than 12 kcal/mol (dotted line, corresponding to 20 conformations) or 14 kcal/mol (dashed line, corresponding to 35 conformations). (b) Data for chrg-Ala6. The dotted and dashed lines correspond to 7 and 18 conformations, respectively. (c) Data for cyc-Ala6. The time evolution of the population probabilities of basin A (solid line), basin B (dashed line), and basin C (dotted line). The insert shows the short time (0–0.4 ns) population probabilities. These probabilities were corrected for size effects by dividing them by the number of minima in each basin.

abilities of the individual conformations located at the bottom of the funnel. The identification of the sets of conformations, which correspond to the bottom of the funnels for Ala6 and chrg-Ala6, rests on the connectivity and topology of their lowest-energy minima on the PELs. Using the disconnectivity graphs of Ala6 [Fig. 1(a)] and of chrg-Ala6 [Fig. 1(c)], we define a conformation as located at the bottom of the funnel if the barrier that separates it from the global minimum is lower than 14 kcal/mol. The time-dependent probability flow in Figs. 3(a) and 3(b) shows the tendency to populate the low-energy regions of the funnel on these PELs. To visualize the gradual population of the funnel, one can follow the population probability of the GM, as well as the accumulated probabilities for all minima separated from the GM by 12 or 14 kcal/mol. We note that the accumulated probability of finding Ala6 and chrg-Ala6 at states separated from the GM by barrier heights lower than 14 kcal/mol (35 conformations of Ala6 and 18 conformations for chrg-Ala6) is higher than the accumulated probability of finding them at states separated by barriers lower than 12 kcal/mol (20 and 7 conformations of Ala6 and chrg-Ala6, respectively). Namely, at 400 K the population probability is not focused on the global minimum of Ala6 and chrg-Ala6, which is located at the bottom of the funnel, but rather it is spread over a broader region including conformations that are located higher up on the PEL.

The energy landscape of Ala6 [Figs. 1(a) and 1(b)] exhibits a funnel with higher energy traps, manifesting population at thermally induced depopulation of the global minimum at short times [Fig. 3(a)]. On the other hand, the energy landscape of chrg-Ala6 [Figs. 1(c) and 1(d)] is characterized by a single funnel, which acts as a sink [Fig. 3(b)]. While the funnel shapes of Ala6 and of chrg-Ala6 result in a qualitatively similar pattern of saturation of the population probabilities for the folding process, marked differences in the ''folding'' time scales, which originate from the specific funnel properties, are observed. The funnel of chrg-Ala6 is deeper and narrower than that of Ala6 and its bottom is isolated from the high-energy states by an energetic gap of about 6 kcal/mol. In a folding process this energetic gap acts as a bottleneck which results in a longer folding time scale for chrg-Ala6 as compared to Ala6. Thus Ala6 ''folds'' within a few ns while chrg-Ala6 ''folds'' within a few μs (Fig. 3). Due to the high barriers, which separate the funnel bottom on the chrg-Ala6 PEL from the rest of the system, only these low-energy minima at the funnel bottom are populated at equilibrium. On the other hand, for Ala6, where no significant barrier separates the low-energy states from the high-energy states, the population is distributed even over higher states, as seen by the fact that the accumulated probability of all low-energy states near Ala6's funnel bottom is less than unity (Fig. 3). Accordingly, one may expect that with raising the temperature, the smearing of the population will become more pronounced for Ala6, while for chrg-Ala6 it will depend on whether the temperature rise will enable the crossing of the gap between the funnel bottom and its higher part. The population of the higher part of the bottom of the funnel can also be seen by the probability flow profiles that first increase and then decrease with time [Fig. 3(a)].

Figures 1(e) and 1(f) reveal that the energy landscape of cyc-Ala6 includes three competing basins. Basin A is the deepest basin on this landscape. It is separated from basins B and C, which are disconnected and include states with a similar energy, by barriers in the range of 10–12 kcal/mol. Despite the similar thermodynamic and kinetic characteristics of basins B and C, the structural properties of the conformations they capture are different. Using an order parameter ρ, which estimates the resemblance of the dihedral angles of a given conformation to the dihedral angles of the ''native'' conformation, one finds that basin C is structurally more similar to basin A than to basin B.[37] The time evolution of the population probabilities, $P_I(t)$, of basins A–C at 400 K, starting from a state in which only the five highest energy conformations are populated, is shown in Fig. 3(c). To correct for size effects, the basin probabilities are divided by the number of minima in each basin (basins A, B, and C include 7, 8, and 13 conformations, respectively). The equilibrium population of basin A, which is the deepest basin on the landscape, is the highest and the equilibrium populations of basins B and C are significantly smaller. One should note that cyc-Ala6 ''folds'' more slowly than Ala6 but faster than chrg-Ala6. The evolution of the probability density for finding the system in basins B and C first increases and then decreases. In protein folding,[1] as well as in general kinetics, this behavior is commonly referred to as evidence for a ''kinetic intermediate.'' However, one should note that the population of basin B is quite negligible before reaching equilibrium and at equilibrium. This may indicate that basin C acts as a kinetic intermediate, which plays a role in the folding process, while basin B is an off pathway intermediate, which does not effectively participate in the folding process.

B. Projection of the probability densities on the principal coordinates

To visualize the kinetics of the polypeptides it may be useful to project the evolution of the probability density onto the first few principal coordinates obtained through the principal component analysis[49] of each of the three systems. To study the spatial and temporal propagations of the probability population distributions obtained by solving the master equation at 400 K for each of the three molecular systems, relative to the underlying energy landscape, they were projected onto the same two principal coordinates, q_1 and q_2, used for charting the energy landscape of each system (Fig. 1). These two coordinates should be interpreted in the context of structural similarity, i.e., the closer two points are on the (q_1,q_2) plane the more similar are the conformations represented by them. For each molecular system five snapshots are presented, starting with the initial population distribution, followed by three snapshots during the system's kinetics, and ending with a population (almost) at equilibrium. As before, initially only the five highest energy states on the PEL are populated. The five snapshots of the projected probability population for Ala6, together with its energy landscape in the q_1-q_2 plane, are presented in Fig. 4. These projected probability populations illustrate that when the highest energy states of the Ala6 energy landscape are popu-

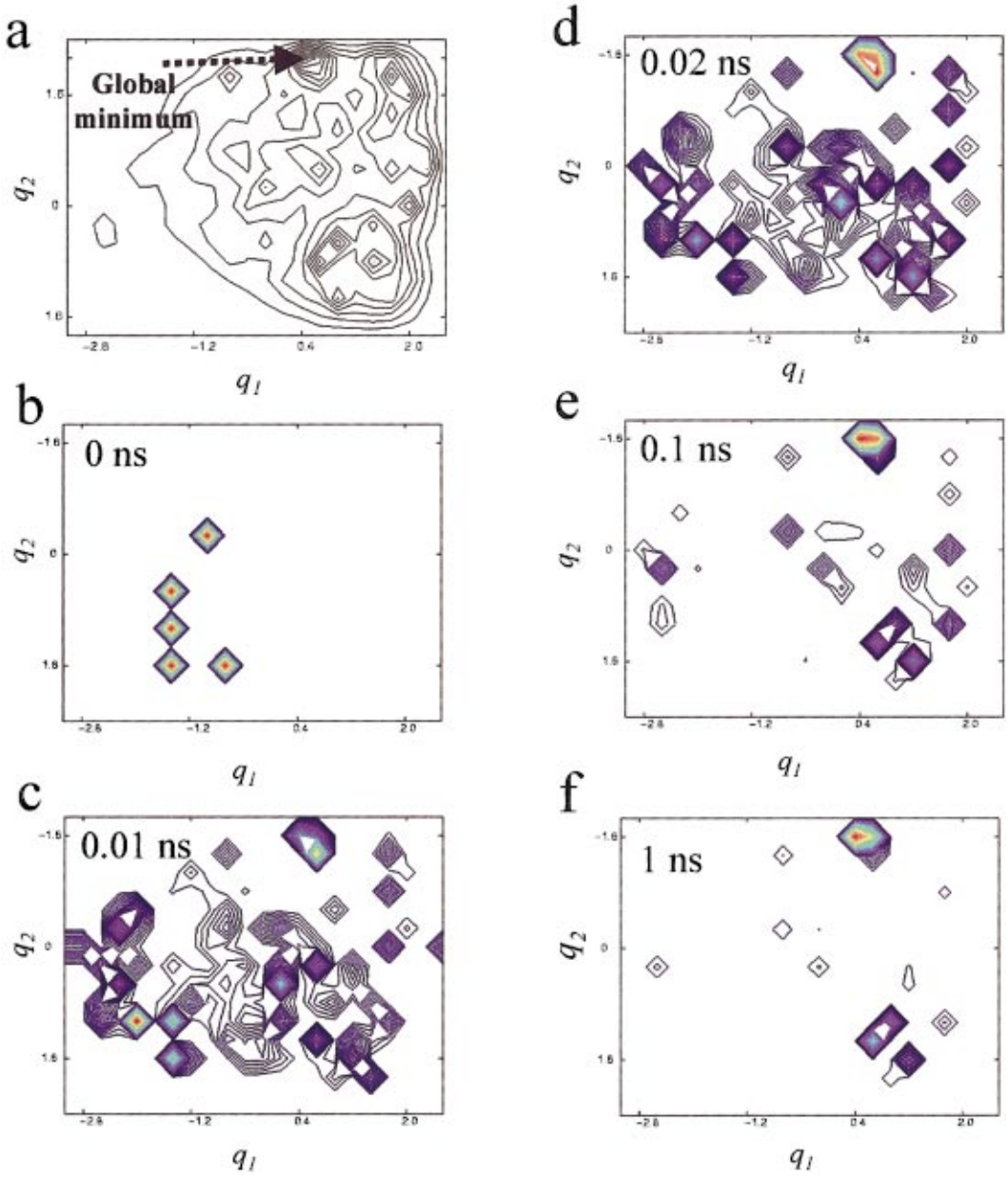

FIG. 4. (Color) (a) A contour plot of the energy landscape of Ala6 [see Fig. 1(b)]. Figures 4(b)–4(f) represent five snapshots of the probability populations of Ala6 at 400 K projected onto the two first principal coordinates that defined its energy landscape. Only probability populations larger than 0.001 are shown. Low and high probability values are shown by blue and red, respectively.

lated [corresponding to small values of q_1, see Fig. 4(a)], the decrease in the probability population along the energy landscape toward low energy states is very fast. The snapshots at 0.01 ns and 0.02 ns [Figs. 4(c) and 4(d), respectively] illustrate that many states with different energies and structures are populated and that the global minimum is characterized by the largest population. The folding process of Ala6 is completed in a few nanoseconds, when the bottom of the funnel is entirely populated. The snapshot at $t=1$ ns [Fig. 4(e)] shows that toward equilibrium not only is the global minimum populated, but so are other conformations that belong to the funnel bottom. The observation that the two populated regions at $t=1$ ns are distinct indicates that the funnel bottom includes conformations with different structural characteristics. This localization into small sub-basins located at the bottom of the funnel energy landscape of Ala6 was previously observed for another peptide, isobutyryl-$(ala)_3$-NH-methyl,[57] and provides an extension of the common funnel concept.

Five temporal snapshots of the probability population of chrg-Ala6 projected onto the two coordinates defining its energy landscape are shown in Fig. 5. Similar to Ala6, the energy landscape of chrg-Ala6 is characterized by a single funnel such that during the ''folding'' kinetics the probability of finding the system in the funnel bottom increases. However, the time scales associated with the kinetic propagation onto these funnel-like peptide energy landscapes are different for the two polypeptides. While Ala6 ''folds'' within a few ns, chrge-Ala6 ''folds'' on a slower time scale of several hundred nanoseconds. This result is somewhat nonintuitive

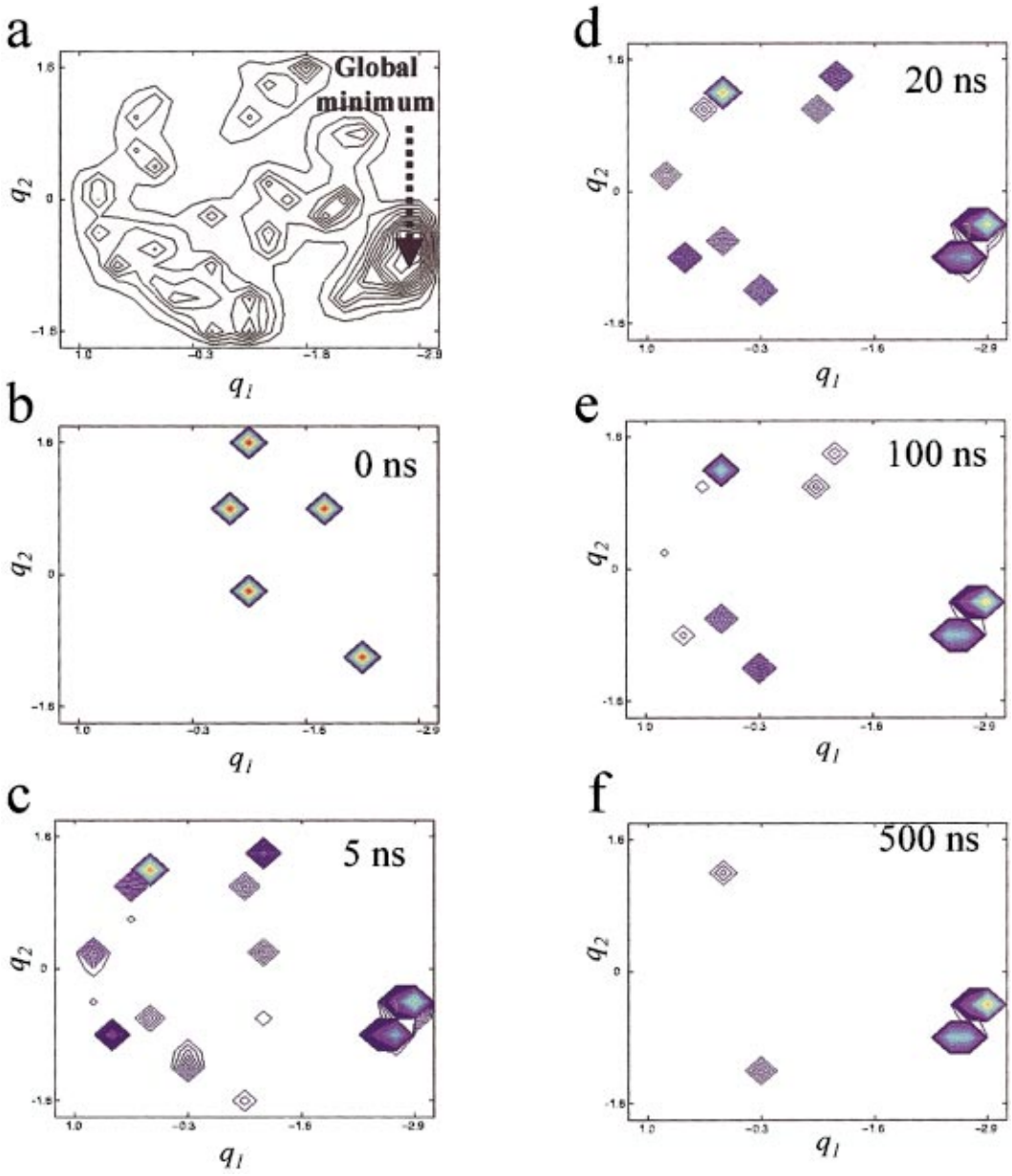

FIG. 5. (Color) A contour plot of the energy landscape (a) and snapshots of probability populations (b)–(f) of chrg-Ala6. Notations identical to Fig. 4.

since it seems that the kinetics on rough, broad, and shallow funnels results in a faster and more efficient folding process than on the smooth, narrow, and deep funnel-like landscape. This observation can be rationalized by noting that chrg-Ala6's funnel is characterized by a barrier of about 6 kcal/mol that separates the ''unfolded'' and the ''folded'' states, which is clearly seen in the topological map of this system [Fig. 1(c)] and is missing from the Ala6 landscape [Fig. 1(a)]. Accordingly, the ''folding'' of chrg-Ala6 is dictated by a bottleneck and thus results in slower ''folding.'' In addition, one can also observe the different topography of the two funnels by the temporal evolution of the probability population toward equilibrium. While the populated region of the Ala6 funnel is large, the corresponding region of chrg-Ala6 is smaller, indicating a narrower funnel.

Figure 6 shows the five snapshots of the probability population projected onto the coordinates defining the energy landscape of cyc-Ala6. In agreement with Fig. 3(c) these snapshots show that basin C is essentially an intermediate basin in the ''folding'' process of cyc-Ala6 which is populated prior to the population of the native basin A. Thus the kinetics of cyc-Ala6 can be qualitatively described by a two-stage process: the populating of basin C and its transition to basin A, which constitutes a slower step. Interestingly, one can observe that the population of basin B is very minor, i.e., while basin C is an on-pathway intermediate, basin B is an off-pathway intermediate. Solving the master equation for this system suggests that not all the pathways connecting the unfolded manifold with the native state participate in the folding process. Namely, the kinetic of cyc-Ala6 is simpler than the ones that might be expected from the three basins energy landscape. The time scale of the ''folding'' process of cyc-Ala6 is about tens ns.

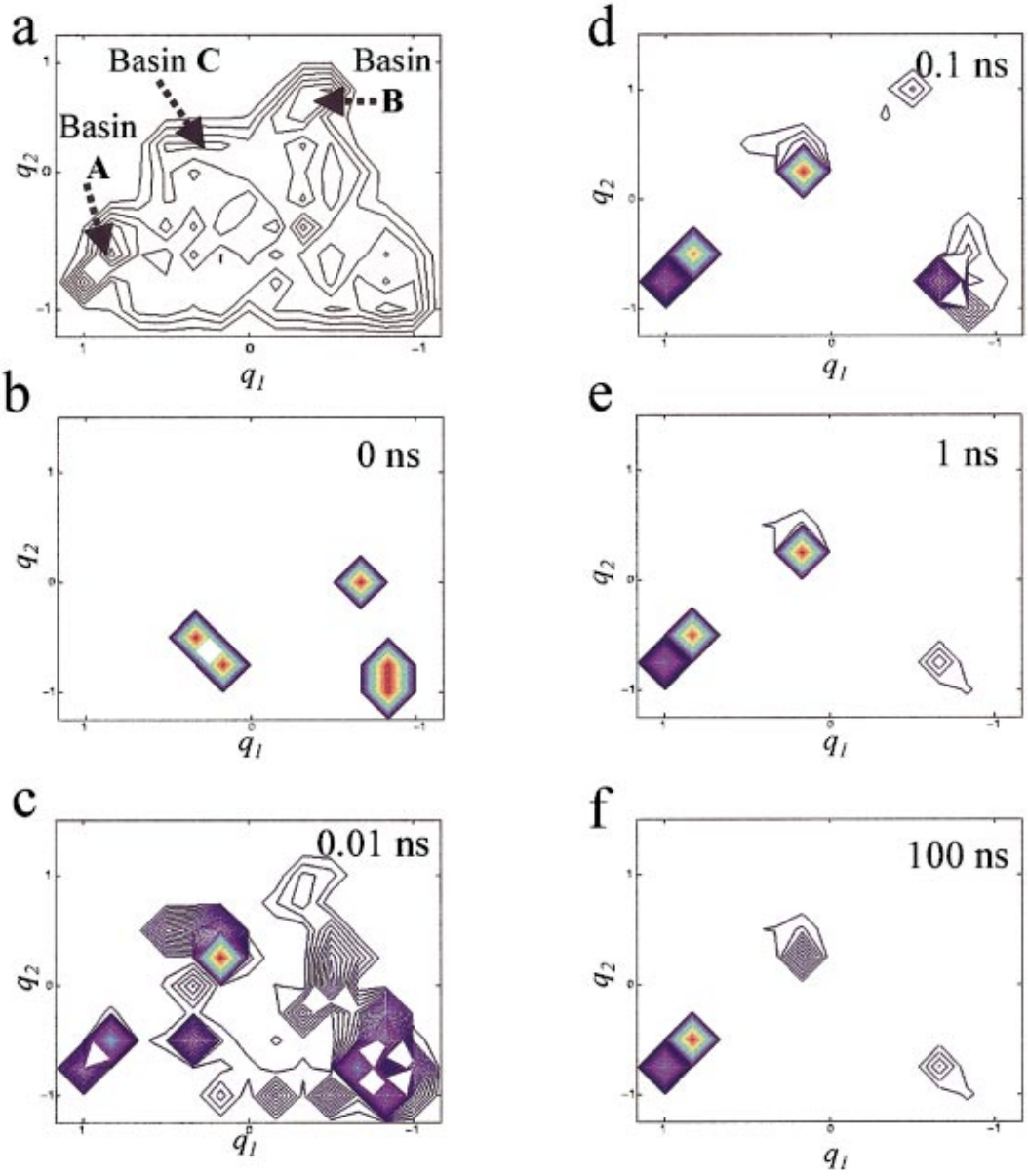

FIG. 6. (Color) A contour plot of the energy landscape (a) and snapshots of probability populations (b)–(f) of cyc-Ala6. Notations identical to Fig. 4.

C. Energy relaxation

The energy landscapes of the three alanine hexapeptides studied here differ in their topologies [Figs. 1(a), 1(c), and 1(e)], topographies [Figs. 1(b), 1(d), and 1(f)], and in the distribution of their barrier heights (Fig. 2). Accordingly, these molecular peptide systems may exhibit not only different ''folding'' time scales, but also different kinetic mechanisms. The nonuniform distributions of the barrier heights across the range of barriers indicate that the ''folding'' kinetics of each system is a sum of multiple microscopic relaxations, each with its own activation energy (Fig. 2). To identify the macroscopic kinetic behavior of a complex system, one has to follow the evolution of the probability vector towards $\mathbf{P}^{eq}$. Here we choose to follow the relaxation of the average energy ($\langle E \rangle$) during a downhill ''folding'' process until it reaches the system's equilibrium energy value, $\langle E \rangle^{eq}$. This average energy is a weighted average over the potential energy E_i of each confirmation i. Following Eq. (4) the time evolution of this average energy can be expressed as,

$$\langle E(t) \rangle = \sum_{i=1}^{n} E_i P_i(t) = \langle E \rangle^{eq} + \sum_{i=1}^{n} \sum_{k} E_i C^k S_i^k e^{\lambda_k t}. \quad (7)$$

Since, initially, only the five conformations with the highest energy values are equally populated, the initial average energy from which the relaxation profile of each system starts is $\langle E(t=0) \rangle = \Sigma_{i=1}^{5} E_i/5$, and as time progresses the difference $[\langle E(t) \rangle - \langle E \rangle^{eq}]$ decays to zero.

The overall energy relaxations, represented by the average energy $\langle E \rangle$ as a function of time at different temperatures for Ala6, chrg-Ala6, and cyc-Ala6 are shown in Figs. 7, 8, and 9. These relaxation profiles do not obey a single-exponent relaxation, nor can they be well-fitted by a stretched exponential or a power-law function. Adequate fit-

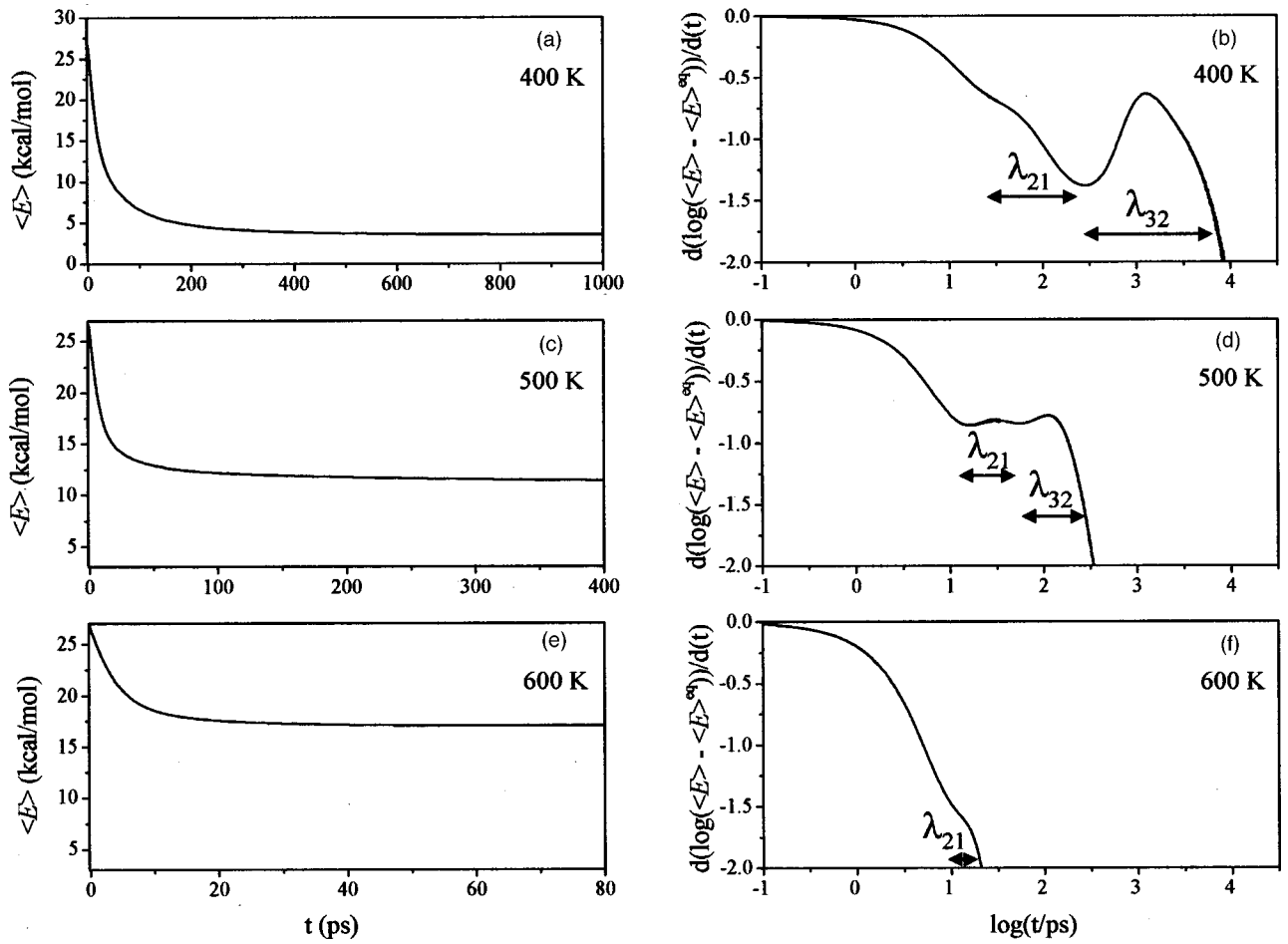

FIG. 7. The energy relaxation of Ala6 at 400 K (a), 500 K (c), and 600 K (e). The logarithmic derivatives of the relaxation data are plotted in (b), (d), and (f), respectively, and reflect the hierarchical behavior of the ''folding'' kinetics. The space between two successive humps, $\lambda_{n+1,n}$, is indicative of the ratio of the characteristic times of the two discrete time scales.

tings of the time evolution are given by using multiexponential functions with different lifetimes τ_i, i.e., $\langle E(t)\rangle - \langle E\rangle^{eq} = \Sigma_i A_i e^{-t/\tau_i}$, with $i=2$, 3, or 4. This analysis indicates that for each system the relaxation is a hierarchical process characterized by fast (short time scales) and by slow (long time scales) events. As such, we applied the analysis proposed by Metzler, Klafter, and Jortner[38] to explore hierarchical relaxation processes. According to this analysis, the multiple down-cascading relaxation pattern of complex systems can be observed by plotting the logarithmic derivative of the data in respect to time versus the logarithm of the time. In doing so, a picture of oscillations (called logarithmic oscillations) emerges. Each hump represents a discrete time scale and the space between two successive humps, $\lambda_{n+1,n}$, is indicative for the ratio of the characteristic times of two successive discrete time scales.[38]

$$\lambda_{n+1,n} = \log(\tau_{n+1}/\tau_n). \tag{8}$$

For the three polypeptide systems studied herein, the dynamic complexity of the energy relaxation, manifested by its nonexponential character, is observed mainly at a low temperature (400 K), and decreases at higher temperatures. The details of these energy relaxations are system dependent. The logarithmic derivatives of the energy relaxation of Ala6 at different temperatures indicates that these relaxations do not obey a single exponential function, rather they are composed

TABLE I. The multiexponential energy relaxation for the folding of hexapeptides. The ratios of the lifetimes were obtained from the $\lambda_{n+1,n}$ parameters, Eq. (8) and Figs. 7–9. The values in brackets are the ratios of lifetimes obtained by the optimal multiexponential fit (see Table II).

	Ala6		chrg-Ala6			cyc-Ala6	
	τ_2/τ_1	τ_3/τ_2	τ_2/τ_1	τ_3/τ_2	τ_4/τ_3	τ_2/τ_1	τ_3/τ_2
400 K	9 (5)	25 (24)	160 (126)	480 (210)	20 (26)	9 (11)	725 (550)
500 K	4 (4)	4 (7)	160 (165)	450 (144)	3 (7)	⋯	260 (296)
600 K	4 (3)	⋯	160 (182)	180 (84)	⋯	⋯	125 (100)

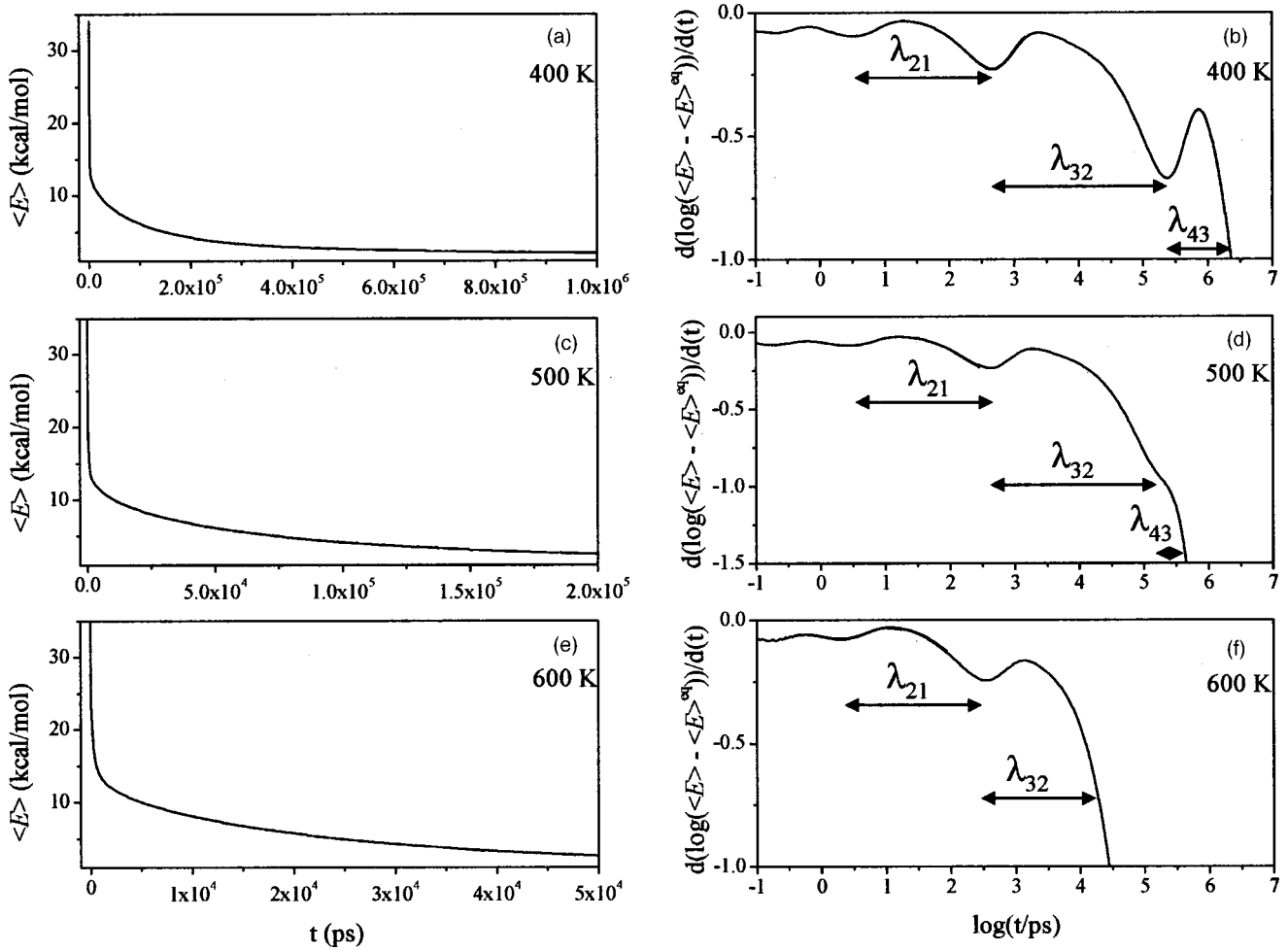

FIG. 8. Energy relaxation of chrg-Ala6. Notations as in Fig. 7.

of several discrete time scales [Figs. 7(b), 7(d), and 7(f)]. In the kinetics of Ala6 at 400 K, one can identify three time scales [Fig. 7(b) and Table I)]. Different amplitudes for each of the three relaxation time scales are manifested by the different hump heights. It should be noted that the slowest time scale does not appear as a shoulder and it can be approximated by the time where the logarithmic derivative decays to a very low value. The spaces between successive humps indicate that the fastest relaxation channel is nine-fold faster than the second relaxation channel and that the latter is 25-fold faster than the slowest relaxation channel (see Table I). The values of the three decay lifetimes, which describe the ''folding'' kinetics of Ala6 at 400 K, were obtained from fitting of the relaxation curve to a three-exponential function (Table II). The ratios of the successive lifetimes τ_{n+1}/τ_n obtained from the multiexponential fit are close (within a numerical factor of ~ 2) to those obtained from the $\lambda_{n+1,n}$ parameters (Table I).

The characteristic times, obtained from the multiexponential fit of the three relaxations, become shorter with increasing temperature (Table II). This observation is also in accord with the shifting of the logarithmic derivative to shorter time values. The temperature dependence of each characteristic time, τ_i, is governed by the effective barrier height, $E_i^{\#}$, as implied by the Arrhenius equation ($\tau_i \sim \exp(E_i^{\#}/kT)$). Accordingly, the larger effect will be manifested for a relaxation channel with a larger barrier height (long time scale) and the ratio between the characteristic times of two successive relaxations, τ_{n+1}/τ_n, becomes smaller with increasing the temperature (Table I). At 500 K the relaxation of Ala6 still consists of contributions from three single exponential decays, although the differences in time scales are considerably smaller than at lower temperatures. A similar behavior is observed at 600 K, where τ_3 is rather close to τ_2 and the energy relaxation is characterized by only two time scales, τ_1 and τ_2 (Table I). At this relatively high temperature, the time scale of the two slowest relaxation channels (defined by τ_2 and τ_3) collapse to a single characteristic time and the joint relaxation channels have a larger amplitude in comparison to the amplitudes of the separated relaxation channel. In addition, the relatively small difference between τ_1 and τ_2 ($\tau_2/\tau_1 = 4$, see Table I) indicates that at higher temperatures the three time scales may collapse into a single time scale and the resulting kinetics would obey the simple relaxation function.

The energy decays of chrg-Ala6 at different temperatures and their logarithmic derivatives are shown in Fig. 8. As expected, at higher temperatures the system reaches equilibrium faster. Again, the logarithmic derivatives of the en-

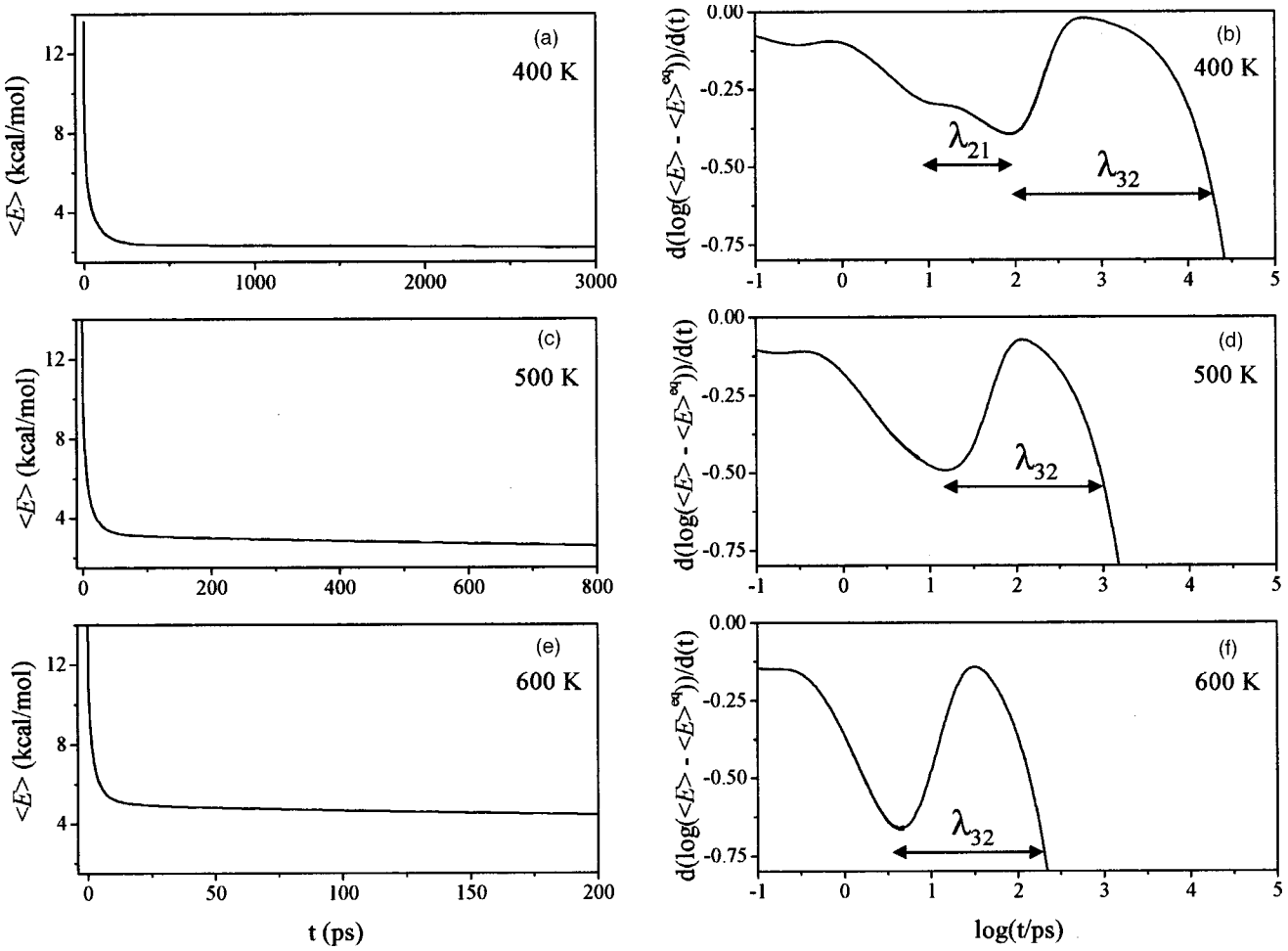

FIG. 9. Energy relaxation of cyc-Ala6. Notations as in Fig. 7.

ergy decays reveal their kinetic complexity. The energy relaxation of chrg-Ala6 at 400 K is characterized by four discrete time scales (Table II). For chrg-Ala6, in comparison with Ala6, the ratio between two successive time scales is significantly larger, indicating a wider range of time scales (Table I). Namely, the ''folding'' of chrg-Ala6 is more hierarchical than that of Ala6. With the increasing temperature, each time scale becomes shorter, as based on the Arrhenius equation, which results in a smaller ratio between two successive time scales. At 600 K τ_3 and τ_4 collapse and thus the relaxation is not characterized by four channels but rather by three discrete channels. However, since for chrg-Ala6 the time scales are significantly different, increasing the temperature is not expected to result in the collapse of the four channels into a single one; i.e., the ''folding'' of chrg-Ala6 is inherently hierarchical.

A similar hierarchical dynamics is observed for cyc-Ala6 (Fig. 9). As for Ala6, the dynamics of cyc-Ala6 is characterized at 400 K by three time scales (Tables I and II). In this case the ratios between the two successive time scales τ_3/τ_2 are closer to those for chrg-Ala6 than for Ala6. The relative difference in the time scales indicates that the ''downhill dynamics'' of cyc-Ala6 should be considered as a hierarchical process even at elevated temperatures.

The logarithmic oscillatory behavior of the three systems inferred from the logarithmic derivative analysis is important in order to extract the minimal number of relaxation channels involved in the decay of the hierarchical process. This infor-

TABLE II. The decay lifetimes, τ_j (in ns) obtained from the optimal multiexponential fit of the relaxation curves of each polypeptide at 400, 500, and 600 K. The number of exponents involved in each fit was determined by the logarithmic oscillations behavior (see Table I).

	Ala6			chrg-Ala6				cyc-Ala6		
	τ_1	τ_2	τ_3	τ_1	τ_2	τ_3	τ_4	τ_1	τ_2	τ_3
400 K	0.024	0.12	2.83	0.0031	0.39	82	2110	0.0053	0.057	31.4
500 K	0.0071	0.026	0.18	0.0026	0.43	62	430	···	0.0051	1.51
600 K	0.0037	0.011	···	0.0017	0.31	26	···	···	0.0018	0.18

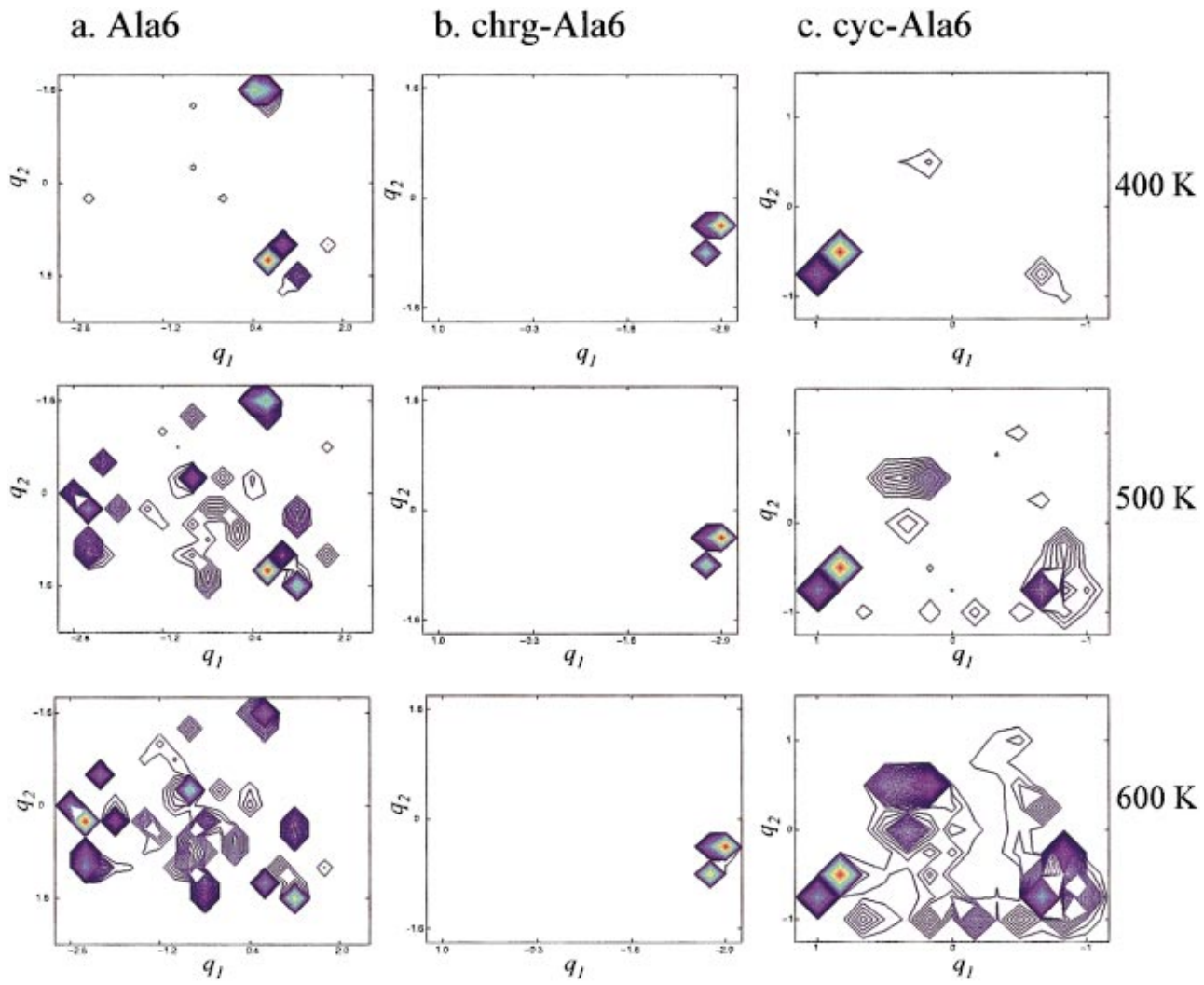

FIG. 10. (Color) Projections of the equilibrium probability populations at 400–600 K for Ala6 (a), chrg-Ala6 (b), and cyc-Ala6 (c). Notations identical to Fig. 4. Blue areas indicate low probability and red area indicate high probability.

mation is crucial prior to any multiexponential fit of the time-dependent data. Moreover, the estimates of the ratio of two successive characteristic times, τ_{n+1}/τ_n, using the $\lambda_{n+1,n}$ parameters may serve to evaluate the initial values of the multiexponential function. The values of τ_{n+1}/τ_n obtained from the $\lambda_{n+1,n}$ parameters and from the values of the decay lifetimes obtained from the fits of all the three hexapeptides are close (Table I). The differences between the τ_{n+1}/τ_n values obtained by the two calculations may originate from the low sensitivity of $\lambda_{n+1,n}$ parameters or from the involvement of possible additional relaxation channels, whose detection was hampered due to their partial overlap with other logarithmic oscillations.

D. Equilibrium probabilities

Figure 10 shows projections of the equilibrium probability distribution calculated for the three molecular systems at 400, 500, and 600 K. The equilibrium probability populations of Ala6, chrg-Ala6, and cyc-Ala6 are projected onto the two principal coordinates (q_1 and q_2) used for charting their energy landscapes. For each peptide these projections are complementary to the five snapshots shown in Figs. 4–6 representing the probability population at $t\rightarrow\infty$. The projected equilibrium probability distributions for Ala6 at different temperatures [Fig. 10(a)] demonstrate that as the temperature increases from 400 K to 600 K, it becomes easier to cross high barriers leading to higher energy states, which are located at high regions of the landscape, being populated at equilibrium. The Ala6 equilibrium energy values, $\langle E\rangle^{eq}$, manifest at different temperatures the system's tendency to populate higher energy states with the increasing temperature (Table III). In contrast to the smearing of the population probability, with increasing the temperature on the energy landscape of Ala6, such an effect is not observed for chrg-

TABLE III. The equilibrium potential energy values (in kcal mol^{-1}) of the three polypeptides.

	Ala6	chrg-Ala6	cyc-Ala6
	$\langle E\rangle^{eq}$	$\langle E\rangle^{eq}$	$\langle E\rangle^{eq}$
400 K	3.28	0.51	0.31
500 K	11.25	0.63	1.49
600 K	17.06	0.86	3.90

Ala6, at least at this temperature range, because the funnel that characterizes its energy landscape is deeper and narrower, effectively acting as a trap. The values of $\langle E \rangle^{eq}$ reflect that there is almost no change in the equilibrium population probabilities of chrg-Ala6 with the increasing temperature (Table III). Very high temperatures are required to escape from this funnel to the higher energy ''unfolded'' states. Finally cyc-Ala6 exhibits a temperature effect similar to that observed for Ala6. As the temperature increases, more high-energy states become populated, although their population remains small in comparison to that of basin A. Interestingly, the most stable ''folded'' basins of chrg-Ala6 and of cyc-Ala6 continue to act as the dominant populated state even at high temperatures, while the native basin of Ala6 is less populated at higher temperatures and the ''unfolded'' states become favorable.

IV. CONCLUSIONS

The interrelationship between ''folding'' kinetics and the corresponding multidimensional potential energy landscapes was explored for three alanine hexapeptides using the master equation approach. To help visualize the kinetics of the polypeptide, the time evolution of the probability density was projected onto the first two principal coordinates obtained through the principal component analysis of each of the three systems. These projections are extremely useful since they offer an insight into the relationship between the temporal evolution of the state-population kinetics calculated by the master equation and the topography of the energy landscape. Our approach highlights the relative role of funnels, traps and ''bottlenecks'' that are likely to exist on complex molecular energy landscapes.

Solving the master equation for these three molecular systems showed that they each exhibit significantly different folding time scales. Linear hexaalanine, Ala6, ''folds'' within a few ns, cyclic hexaalanine, cyc-Ala6, ''folds'' within a few hundred ns, and the hexaalanine with opposite charges at its termini, chrg-Ala6, ''folds'' on a time scale of a few thousand ns. Interestingly, the peptides with geometrical constraints (cyclization and electrostatic interaction) ''fold'' more slowly than the more flexible peptides. This behavior can be traced to the energy landscape characteristics. While the dynamics of Ala6 corresponds to ''folding'' into a rough funnel, the dynamics of cyc-Ala6 involves trapping by an on-pathway intermediate prior to the population of the native basin. The chrg-Ala6 is characterized by a single funnel as well. However, since the single funnel is narrower and is isolated by an energetic gap from the unfolded states, the time scale for the kinetics on this energy landscape is significantly longer.

Furthermore, using a logarithmic derivative analysis, we found that the kinetics of the three hexapeptides studied herein is hierarchical, as indicated by the existence of logarithmic oscillations. Consequently, the kinetics at temperatures between 400–600 K are nonexponential and can be described by several relaxation channels, each with its characteristic time scale. For Ala6, cyc-Ala6, and chrg-Ala6, the ratio between the time scales of the slowest and the fastest relaxation channels at 400 K is about 200, 6000, and $7 \cdot 10^5$, respectively, and reflects the hierarchical characteristics of the kinetics. Although the kinetics for all three systems can be described by several discrete time scales, the six orders of magnitude difference between the fast and the slow time scales of chrg-Ala6 suggests a more pronounced hierarchical character for this system. Indeed, in a previous study of these molecular systems, using an order parameter as the reaction coordinate, it was suggested, only for the ''folding'' of chrg-Ala6, that it occurs via two sequential stages and not as a competition between parallel single-step pathways.[37] Increasing the temperature resulted in the collapse of two relaxation channels into similar time scales and, consequently, the kinetics becomes less hierarchical. While the kinetics of the flexible polypeptide Ala6 at 400 K is characterized by three time scales, its kinetics at 600 K is described only by two similar time scales, which at even higher temperatures seem to collapse into a simple exponential relaxation. In contrast, the kinetics of the constrained peptides, cyc-Ala6 and chrg-Ala6, are inherently hierarchical and their ''folding'' should be described by multiple time scales even at high temperatures.

ACKNOWLEDGMENTS

The authors are grateful to M. Bixon and J. Klafter for stimulating discussions. One of the authors (Y.L.) wishes to acknowledge, with gratitude, the support of the Clore Foundation.

[1] A. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York, 1999).
[2] M. Karplus and E. Shakhnovich, in *Protein Folding* (W. H. Freeman, New York, 1992).
[3] C. M. Dobson, A. Sali, and M. Karplus, Angew. Chem. **37**, 868 (1998).
[4] R. Zwanzig, Proc. Natl. Acad. Sci. U.S.A. **92**, 9801 (1995).
[5] R. Zwanzig, Proc. Natl. Acad. Sci. U.S.A. **94**, 148 (1997).
[6] L. Mirny and E. Shakhnovich, Annu. Rev. Biophys. Biomol. Struct. **30**, 361 (2001).
[7] A. Sali, E. Shakhnovich, and M. Karplus, Nature (London) **369**, 248 (1994).
[8] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins: Struct., Funct., Genet. **21**, 167 (1995).
[9] P. G. Wolynes, J. N. Onuchic, and D. Thirumulai, Science **267**, 1619 (1995).
[10] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, Annu. Rev. Phys. Chem. **48**, 539 (1997).
[11] R. S. Berry, N. Elmaci, J. P. Rose, and B. Vekhter, Proc. Natl. Acad. Sci. U.S.A. **94**, 9520 (1997).
[12] K. A. Dill, Protein Sci. **8**, 1166 (1999).
[13] A. R. Dinner, A. Sali, L. J. Smith, C. M. Dobson, and M. Karplus, Trends Biochem. Sci. **25**, 331 (2000).
[14] S. E. Jackson, Folding Des. **3**, R81 (1998).
[15] R. L. Baldwin, Folding Des. **1**, 1 (1996).
[16] R. L. Baldwin, Nat. Struct. Biol. **8**, 92 (2001).
[17] S. Khorasanizadeh, I. D. Peters, and H. Roder, Nat. Struct. Biol. **3**, 193 (1996).
[18] B. A. Krantz and T. R. Sosnick, Biochemistry **39**, 11696 (2000).
[19] M. C. Shastry and H. Roder, Nat. Struct. Biol. **5**, 385 (1998).
[20] Y. Bai, T. R. Sosnick, L. Mayne, and S. W. Englander, Science **269**, 192 (1995).
[21] T. Pan, X. Fang, and T. Sosnick, J. Mol. Biol. **286**, 721 (1999).
[22] H. Frauenfelder and P. G. Wolynes, Science **229**, 337 (1985).
[23] H. Frauenfelder and P. G. Wolynes, Phys. Today **47**, 58 (1994).
[24] H. Frauenfelder, in *Physics of Biological Systems* (Springer, Berlin, 1998).
[25] H. Frauenfelder and B. McMahon, Ann. Phys. (Leipzig) **9**, 655 (2000).
[26] M. Gruebele, Annu. Rev. Phys. Chem. **50**, 485 (1999).

[27] J. Sabelko, J. Ervin, and M. Gruebele, Proc. Natl. Acad. Sci. U.S.A. **96**, 6031 (1999).
[28] T. Metzler, J. Klafter, J. Jortner, and M. Volk, Chem. Phys. Lett. **293**, 477 (1998).
[29] K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, and H. S. Chan, Protein Sci. **4**, 561 (1995).
[30] M. Karplus, J. Phys. Chem. B **104**, 11 (2000).
[31] H. S. Chan and K. A. Dill, Proteins: Struct., Funct., Genet. **30**, 2 (1998).
[32] R. Du, V. S. Pande, A. Y. Grosberg, T. Tanaka, and E. S. Shakhnovich, J. Chem. Phys. **108**, 334 (1998).
[33] F. B. Sheinerman and C. L. Brooks, Proc. Natl. Acad. Sci. U.S.A. **95**, 1562 (1998).
[34] P. Ferrara and A. Caflisch, Proc. Natl. Acad. Sci. U.S.A. **97**, 10780 (2000).
[35] O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).
[36] Y. Levy and O. M. Becker, Phys. Rev. Lett. **81**, 1126 (1998).
[37] Y. Levy and O. M. Becker, J. Chem. Phys. **114**, 993 (2001).
[38] R. Metzler, J. Klafter, and J. Jortner, Proc. Natl. Acad. Sci. U.S.A. **96**, 11085 (1999).
[39] A. L. Blatz and K. L. Magleby, J. Physiol. (London) **378**, 141 (1986).
[40] J. Klafter, G. Zumofen, and A. Blumen, J. Phys. A **24**, 4835 (1991).
[41] A. R. Fersht, Phil. Trans. Biol. Sci. Soc. Lond. B **348**, 11 (1995).
[42] R. A. Broglia and G. Tiana, J. Chem. Phys. **114**, 7267 (2001).
[43] N. G. v. Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam (1981).
[44] R. Czerminski and R. Elber, J. Chem. Phys. **92**, 5580 (1990).
[45] K. D. Ball and R. S. Berry, J. Chem. Phys. **109**, 8557 (1998).
[46] M. A. Miller, J. P. K. Doye, and D. J. Wales, Phys. Rev. E **60**, 3701 (1999).
[47] D. J. Wales, J. P. K. Doye, M. A. Miller, P. N. Mortenson, and T. R. Walsh, Adv. Chem. Phys. **115**, 1 (2000).
[48] P. N. Mortenson and D. J. Wales, J. Chem. Phys. **114**, 6443 (2001).
[49] O. M. Becker, J. Comput. Chem. **19**, 1255 (1998).
[50] Y. Levy, J. Jortner, and O. M. Becker, Proc. Natl. Acad. Sci. U.S.A. **98**, 2188 (2001).
[51] A. E. Garcia, Phys. Rev. Lett. **68**, 2696 (1992).
[52] R. Abagyan and P. Argos, J. Mol. Biol. **225**, 519 (1992).
[53] J. M. Troyer and F. E. Cohen, Proteins: Struct., Funct., Genet. **23**, 97 (1995).
[54] A. Fernandez and G. Appignanesi, Phys. Rev. Lett. **78**, 2668 (1997).
[55] L. S. D. Caves, J. Evenseck, and M. Karplus, Protein Sci. **7**, 649 (1998).
[56] Y. Levy, E. Hanan, B. Solomon, and O. M. Becker, Proteins: Struct., Funct., Genet. **45**, 382 (2001).
[57] O. M. Becker, Proteins: Struct., Funct., Genet. **27**, 213 (1997).
[58] O. M. Becker, J. Mol. Struct.: THEOCHEM **398**, 507 (1997).
[59] O. M. Becker, Y. Levy, and O. Ravitz, J. Phys. Chem. B **104**, 2123 (2000).
[60] N. Elmaci and R. S. Berry, J. Chem. Phys. **110**, 10606 (1999).
[61] A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, Proteins: Struct., Funct., Genet. **17**, 412 (1993).
[62] K. D. Ball and R. S. Berry, J. Chem. Phys. **111**, 2060 (1999).
[63] R. S. Berry and R. E. Kunz, Phys. Rev. Lett. **74**, 3951 (1995).
[64] R. E. Kunz and R. S. Berry, J. Chem. Phys. **103**, 1904 (1995).
[65] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[66] B. R. Brooks and M. Karplus, Proc. Natl. Acad. Sci. U.S.A. **80**, 6571 (1983).
[67] F. H. Stillinger and T. A. Weber, Science **225**, 983 (1984).