Interbasin motion approach to dynamics of conformationally constrained peptides

Florin Despa, Ariel Fernández, and R. Stephen Berry[a)]
Department of Chemistry, The University of Chicago, Chicago, Illinois 60637

Yaakov Levy and Joshua Jortner
Department of Chemistry, Tel Aviv University, Ramat Aviv 69978, Israel

(Received 14 August 2002; accepted 31 December 2002)

In this paper, the interbasin motion (IBM) approach is applied to studying dynamics of conformationally constrained peptides, being extended to a nonideal contact of the system with a thermal bath. The coupling of the system with the thermal bath is expressed in terms of a memory function. The aim of the present study is twofold. First, we present a dynamical diagnosis of the three hexapeptide variants with the main focus on the transitions between basins rather than between individual states. Second, the present study is intended to pinpoint a way for extracting useful information about the strength of the system–solvent coupling and how this interaction affects the propensity of relaxation towards the native state. We show that a slight variation of the value of the memory friction parameter may induce a sizable modification of the relaxation time. In addition, the change of the memory friction parameter produces alterations on short time scales among the population distributions. Especially, high energy basins seem affected the most. In the Markovian limit, the basin populations computed within the IBM model are compared to those obtained by using state-to-state transition rates in the full master equation approach. The two methods yield similar results when the separation of time scales between intra- and interbasin dynamics is completely achieved. © *2003 American Institute of Physics.* [DOI: 10.1063/1.1554393]

I. INTRODUCTION

Recently in the study of peptides and proteins, it was indicated that the overall topography of the potential surface (PS) changes as a result of a conformation constraint and portrays manifestations of the relation between bioactivity and conformation space.[1–5] An important step toward understanding these aspects has been achieved by sampling the characteristic potential surfaces of three alanine–hexapeptide analogs, alanine hexapeptide with neutral terminals (Ala_6), alanine hexapeptide with charged terminals (chrg–Ala_6), and a backbone cyclized alanine hexapeptide (cyc–Ala_6).[3–5] These systems vary in the degree of constraint imposed on their conformational motion. The effect of the conformational constraints on the PS has been quantified and compared to the base of topography, connectivity, and order parameters. In particular, it has been found, not surprisingly, that the internal flexibility of a linear alanine hexapeptide diminishes upon either adding opposite charges or introducing a covalent bond between the two terminals; either constraint decreases the accessible conformation space.[3] The constriction in the conformation space may frequently culminate with splitting the energy landscape in multiple competing basins. In addition, the behavior of the order parameter ρ measuring the similarity based on the dihedral angles to the native structure shows a diminishing structural variability at low energies with a decrease of the volume in the conformation space. The low energy structures apparently become more nativelike for conformationally constrained analogs.

The significantly different topographies of the potential surfaces of the hexapeptide analogs stimulate the question of how to infer the system's dynamics from these specific patterns, using the methods at hand to study dynamics on the PS. Stochastic simulation may offer a practical approach to simulating and understanding dynamics on multidimensional potentials, at least it seems now, in contrast to the impracticability of any approach based on knowing the full topography of the system. Statistical methods have largely been used to study complex kinetics[6–8] and frequently invoked to illuminate the relation between topography and kinetics.[9–14] Transition state theory (TST) can be used to compute a matrix of well-to-well transition probabilities, and from this matrix one can construct a master equation whose solution reveals the population density in each state of the characteristic PS.[9] Thereafter, one may find the probability of finding the system in a certain basin by summing over the probabilities of the individual states belonging to that basin. This strategy seems very informative in the problem of studying dynamics of conformationally constrained peptides. It has been shown that each of the three polyalanine systems has its own energy landscape and relaxation time scale that can be traced to its topographical pattern.[5] The relaxation kinetics of these systems are hierarchical and characterized by multiple time scales of fast and slow events which correspond to parallel processes.[5,15]

Recently, Despa and Berry developed the interbasin motion (IBM) approach,[14,16] which recasts the dynamics on a multidimensional PS in terms of basin-to-basin rather than well-to-well transitions. The leading assumption is that the

[a)]Electronic mail: berry@uchicago.edu

intra- and interbasin dynamics are of different scales. The IBM approach uses a general prescription to compute, directly, the escape rate of the system from each basin of the characteristic PS with full consideration of the topographical fingerprint of that basin. The theoretical foundation is provided by a non-Markovian treatment of the system's behavior, which allows one to study inter-relationship between the general characteristics of the topography of the energy landscape and dynamics. Moreover, the IBM method enables the study of the relation between the dynamics and the coupling with the thermal bath.

This strategy for studying dynamics on complex potential surfaces presents several important steps forward from traditional methods.[6–13] Some are matters of practicality. The IBM approach offers a way to speed up the computation and to reduce the dimension of the system of kinetic equations without losing information about the large-scale topography and configuration entropy. Also, it can be used to infer kinetic properties of relatively large regions (basins) of the PS with a minor computational effort. Others are reasons of purpose. An example is the above-described issue of estimating the effect of modifications in the system's structure on its overall dynamics. Because any structural change affects relatively large areas on the characteristic PS, it then might be much more relevant for studying dynamics to focus on the ability of the system to survey well-defined basin regions, rather than to monitor single state occupancies.[5] It has been demonstrated that the values of the rates of passage among the basins on the characteristic PS are sensitive to the topographical modifications and constitute valuable indicators of the system's kinetic properties.[16]

To illustrate the utility of the concepts and techniques aforementioned, we present a few examples of their application in dynamical studies of polyalanine. In the following, we assume that the basins composing characteristic potential surfaces of Ala_6, chrg–Ala_6, and cyc–Ala_6 underlie non-Markovian kinetic properties, and employ the IBM approach to explore the dynamical behavior of these systems. The basins forming each PS of the three hexapeptides are outlined by a partition scheme based on structural similarity and kinetic connectivity.[3] We computed the basin populations both by applying the IBM model and by using state-to-state TST rates in the full master equation approach. The two methods yielded very similar results for chrg–Ala_6. The agreement is preserved only at a qualitative level for the two other systems, Ala_6 and cyc–Ala_6. For these last two instances, we found from analyzing their transition matrices that the separation of the intra- and interbasin dynamics is not completely achieved. Some of the basins of Ala_6 and cyc–Ala_6 cannot be delineated properly by the present partition scheme. Hence, assuming non-Markovian dynamics inside a basin and Markovian dynamics between basins cannot be done trivially, and always relies on a suitable subset of the basins on the characteristic PS.

Furthermore, we speculated that the IBM approach introduces in the system's dynamics features of nonideal contact with the thermal bath. In this regard, it goes beyond the TST theory and allows us to extract useful information about the strength of the system–solvent coupling and how this interaction affects the propensity of folding. Since the energy landscapes of the three hexapeptide analogs were constructed to simulate vacuum, this strategy proved useful for estimating the effect of the environment on the dynamics.

II. INTERBASIN DYNAMICS ON THE CHARACTERISTIC POTENTIAL SURFACES OF THE THREE HEXAPEPTIDE VARIANTS

The three alanine–hexapeptide analogs we study here are: linear alanine hexapeptide with neutral terminal groups (Ala_6), alanine hexapeptide with a positive charge at the *N*-terminus and a negative charge at the *C*-terminus (chrg–Ala_6), and a backbone cyclized alanine hexapeptide (cyc–Ala_6). These analogs differ in the constraints imposed on their flexibility. A comprehensive characterization of the PSs of these three alanine–hexapeptide analogs, highlighting various aspects such as connectivity, topography, and order parameters, has been published elsewhere.[1,3,5]

A. Partition of the basins on the characteristic potential surfaces

In Figs. 1(a), 1(b), and 1(c) the topological disconnectivity graphs are displayed for each of these three peptides. We notice that all these PSs have staircase topographics in their regions of low energy.[9(c)] The mean saddle heights are small between high energy minima and grow somewhat larger as the energies of the minima decrease. However, the decreases in energy between successive minima are significantly larger than the increases in saddle height along the downward sequences. We can also observe a strikingly large degree of structural variability present at the low energy level, indicating that the native structure of each peptide is an ensemble of conformations and not a single well-defined structure.

For these small peptides there are no real tertiary native contacts[3] and the characterization of their relaxation dynamics in terms of finding a specified native fold is inappropriate. However, the analogy with the protein folding problem can be useful in the following dynamical diagnosis with the specification that, here, folding defines the system's predilection for populating low energy basins. Thus, by looking at Fig. 1(a) and recalling a common criterion of foldability,[17,18] which requires a large energy gap between the energy in the native conformation and the lowest energy of the conformations structurally dissimilar to the native conformation, we can infer that Ala_6 has a low ability to relax to native structures. Also, this surface has a very deep, narrow basin ($B13$) dropping from the slab at 21 kcal/mol to a minimum at 4 kcal/mol above the global minimum, which may act as a kinetic trap. Figures 1(b) and 1(c) show that the constraints imposed on the flexibility of those molecules, chrg–Ala_6 and cyc–Ala_6, induce higher propensities of relaxation towards native structures, as indicated by the important large energy steps on corresponding potential surfaces. Moreover, it was shown that although the conformation spaces of the constrained hexapeptides are much smaller than that of the linear hexapeptide, their relaxation time scales are longer, indicating that the connectivity (i.e., topology) of the PS is an important factor in the relaxation rate.[5]

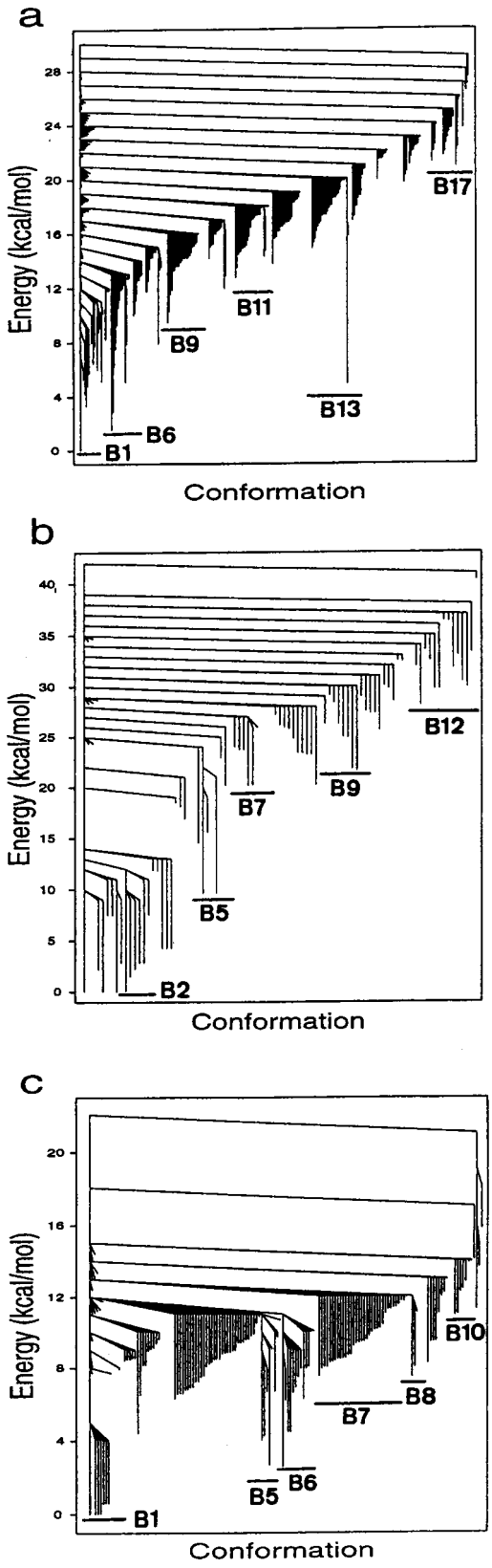


FIG. 1. Topological disconnectivity graphs for (a) Ala_6; (b) chrg–Ala_6; and (c) cyc–Ala_6. Each disconnectivity graph was partitioned into basins and numbered based on the energy where they branch off from the main branch starting from the basin which contains the global minimum. Several basins of each system are marked.

Of central importance in applying the IBM approach is the partition of the basins on the PS at hand. As defined by Berry and Kunz,[19] and Becker and Karplus,[11] a basin is a set of configurations that maps to a single lowest local minimum by direct minimization or by annealing. Particular interest for studying the system's dynamics lies in the simplified interpretation of the PS in terms of different basins that include similar, closely related conformations. For example, in the protein folding case we may classify together all systems that have reached some state with the native physiological activity, but these states may have structural differences in parts of the system not essential to the protein's function. For the specific case of alanine–hexapeptide analogs, the values of the effective order parameter ρ measuring the similarity to the native structure have a slight variation in any branch (basin) of the disconnectivity graph,[3] which implies that the structural variability inside a basin is generally low. This observation may justify a partition scheme based on the kinetic connectivity, i.e., the heights of the saddle points. However, one may frequently encounter a situation in which the minima constituting a certain basin with a similar kinetic connectivity are quite different energetically and geometrically. The PSs of the three-hexapeptide analogs (Fig. 1) reveal that the minima in each basin are energetically different. Moreover, two basins with similar connectivities [see, for example, basins $B5$ and $B6$ in Fig. 1(c)] and that contain conformations with similar potential energies have different values of the average order parameter ρ, indicating different geometries. For tetra-alanine it was found that clustering conformations based on structural similarity and kinetic connectivity yield different pictures, suggesting that each basin cannot be represented by a single structure.[20]

In the following, the region corresponding to each set of minima that branches off from the main (central) branch in Figs. 1(a), 1(b), and 1(c) is considered a basin on the characteristic PS. The distribution of local minima and saddles with respect to the position of the lowest minimum determines the topographical pattern of that basin region. The analysis of the values of the effective order parameter ρ reveals structural similarities among the minima constituting a branch of the disconnectivity graph.[3] We mention here that, although the basin partition based on the kinetic connectivity seems very suggestive, just the regularity of the structural order parameter ρ in each basin, as we noticed above, might not be adequate to ensure complete separation of the time scales of the intra- and interbasin dynamics. As we shall see, the transition matrix can help to identify regions on a PS where this time-scale separation is not fully achieved. For those basins the IBM approach becomes inapplicable.

In the following, we rely on these PSs and focus on dynamics, especially on the modifications in relaxation behaviors induced by conformation constraints.

B. Interbasin motion approach

To provide further details of the computation, we briefly review the basic concepts of the IBM approach. In addition to the previously published material,[14] particular attention will be paid here to revealing the main aspects of the system–solvent coupling. This information is intended to of-

fer a way to understand how the system–solvent interaction affects the propensity of relaxation towards the native state.

1. Basic equations of the IBM model

Under circumstances assumed in the present approach, the basin dynamics is governed by energy accumulation and relaxation processes among the local wells. Moreover, due to the change of curvature of the effective potential between the bottom and the top of each local well, the typical rate of energy exchange between the reaction mode and other modes may become comparable with the characteristic vibrational period of the thermal bath. Therefore, the foundation of the IBM approach starts with the assumption that, inside the basin, the system's motion has a non-Markovian behavior. This behavior is characterized by a time-dependent friction kernel $Z=Z(t)$ at the contact between the system and the thermal bath, which usually is assumed to have a Gaussian form

$$Z(t)=\frac{\gamma}{t_c\sqrt{\pi}}\exp\left(-\frac{t^2}{4t_c^2}\right). \tag{1}$$

$Z(t)$ expresses the time correlation function of the random force $F(t)$ exerted by the thermal bath on the reactive mode and is associated with the correlation time t_c and the Markovian friction γ; t_c is a measure of the extent or rate of equipartition of the reaction mode ω with the thermal bath.

The above statement on the frictional coupling of the reaction system to the thermal bath enables a formulation of a stochastic, one-degree-of-freedom process for the energy E in the unstable mode $\omega(E)$. The approach leads to an expression for the mean first passage time $\tau_{i,\alpha}$ along each "escape channel" i that leads the system to reach the top of the basin $\alpha(E_{b\alpha})$ starting from E_s[21]

$$\tau_{i,\alpha}(E_{b\alpha},E_s)=\int_{E_s}^{E_{b\alpha}}dE\,\frac{\exp(\beta E)}{\mu_{i,\alpha}(E)}\int_0^E dE'\,\frac{\exp(-\beta E')}{\omega_{i,\alpha}(E')}. \tag{2}$$

μ is the energy diffusion coefficient and $\beta=1/k_BT$ has the usual meaning. Following the IBM approach, we now define an effective value of the mean first passage time $\tau_\alpha(E_{b\alpha},E_s)$, which is the summation over independent contributions $\tau_{i,\alpha}(E_{b\alpha},E_s)$ of all the individual states $w_{i,\alpha}$ ($i=\overline{0,N_\alpha}$) of the basin α

$$\frac{1}{\tau_\alpha(E_{b\alpha},E_s)}=\sum_{i=0}^{N_\alpha}\frac{f_{i,\alpha}(E_i,E_{bi})}{\tau_{i,\alpha}(E_{b\alpha},E_s)}, \tag{3}$$

where E_i and E_{bi} are the positions on the energy spectrum of the minimum $w_{i,\alpha}$ and the saddle between the minima $w_{i,\alpha}$ and $w_{i+1,\alpha}$, uphill. Here, $\tau_\alpha(E_{b\alpha},E_s)$ is the average time in which the system escapes from the basin α; N_α is the number of conformation states in the basin α and accounts for the geometric entropy dimension of that basin; f is the weight of each individual contribution to the average (3) given by

$$f_{i,\alpha}(E_i,E_{bi})=\vartheta(E_{bi}-E_s)\,\vartheta(E_s-E_i), \tag{4}$$

where ϑ is the unit-step function. We then use the following expression for computing the escape rate from the basin α:

$$k_\alpha=\left[\int_0^{E_{b\alpha}}dE_s\;p_{ss,\alpha}(E_s)\left(\sum_{i=0}^{N_\alpha}\frac{f_{i,\alpha}(E_i,E_{bi})}{\tau_{i,\alpha}(E_{b\alpha},E_s)}\right)^{-1}\right]^{-1}, \tag{5}$$

which is the average of the effective value of the mean first passage time τ_α over the steady-state distribution $p_{ss}(E_s)$. We assume that the system rapidly achieves thermal equilibrium in the well of the starting point E_s, so that the steady-state distribution of probability p_{ss} is of a Boltzmann form

$$p_{ss,\alpha}(E)\simeq\frac{e^{-\beta E}}{\beta(1-e^{-\beta E_{b\alpha}})}. \tag{6}$$

As can be seen from above, Eq. (5) contains all the requisite information about the main properties of the topographical pattern, particularly the relative positions of minima and saddles along the monotonic sequences. Moreover, general properties of the thermal bath also intervene in system's dynamics.

The set of equations (2)–(6) is used to compute escape rates from the basins α corresponding to each PS of the three alanine–hexapeptide analogs (see Fig. 1). Those rates are incorporated into a master equation which is then solved. The description of flow of a distribution on the surface is then extracted by analyzing the eigenvectors of the master equation in terms of the pattern of the topography of that PS.

2. Escape rates and the relation with thermal bath properties

In the following, the set of minima in a given branch α of each disconnectivity graphs in Figs. 1(a), 1(b), and 1(c) represents the basin α on the corresponding PS. The global minimum (GM) and the first branch of the disconnectivity graph form together the basin of the global minimum, $B1$.

In order to compute the escape rate k_α [Eq. (5)] corresponding to that basin, we first have to account for the average value of $\tau_\alpha(E_{b\alpha},E_s)$. According to Eq. (3), this quantity can be obtained if one knows the contributions $\tau_{i,\alpha}$ of the individual minima. The independent contributions $\tau_{i,\alpha}$ [Eq. (2)] to the average value of the mean first passage time τ_α depend on both vibrational frequency function $\omega_{i,\alpha}(E)$ and corresponding energy diffusion coefficient $\mu_{i,\alpha}(E)$. The key frequency is derived from the specific form of the potential V_{eff} in the region of the basin divide. Here, this is approximated by a harmonic oscillator of frequency $\omega_\alpha^{(0)}$. Accordingly, the corresponding energy diffusion coefficient is[14]

$$\mu_\alpha(E)=\frac{E\gamma}{\beta\omega_\alpha^{(0)}}\exp(-\rho_\alpha^2), \tag{7}$$

where

$$\rho_\alpha=\omega_\alpha^{(0)}t_c \tag{8}$$

determines the memory friction parameter.

If the system is embedded in a solvent, the friction constant γ (the damping rate) depends on the viscosity of the environment and on the characteristic geometry of the system inside the particular basin α. Therefore, we may write the following approximation:

$$\gamma_\alpha \simeq \pi \langle R \rangle_\alpha \eta \frac{n}{M}, \tag{9}$$

where $\langle R \rangle_\alpha$ is, for example, the radius of gyration of all the peptide atoms averaged over the conformations in the basin α, η is the viscosity, and n is an integer depending on the (slip or stick) boundary condition at the interface. On the other hand, the frictional relaxation time t_c and the Markovian friction constant γ are bound by

$$t_c = a\gamma, \tag{10}$$

where a is essentially the inverse of the infinite frequency shear modulus of the solvent. In this context, we may determine the corresponding friction correlation time $t_{c\alpha}$ and damping rate γ_α for each basin on the PS. This fact implies that the relaxation behavior depends intrinsically on the way the various conformations of the system couple to the thermal bath.

One more thing should be stressed at this point. The Gaussian memory function (1) represents a reasonable model for both short and intermediate relaxation time scales, such as those describing a single molecule in a dense fluid or characterizing internal couplings between the reactive mode and the remaining nonreactive modes. During the passage, dynamic coupling may drive the degrees of freedom perpendicular to the reaction coordinate out of equilibrium, or into a mutual equilibrium that does not encompass the reaction coordinate.[22,23] Essentially, this drag effect can be seen as a viscoelastic response of the vacuum. For studying dynamics of isolated systems, the use of internal coupling is consistent with vacuum potential surfaces. Therefore, the memory function (1) fits the aim of the present investigation.

We are aware, from results reported recently,[4] that the effects of solvation may have dramatic consequences for the PS of polyalanine. This fact requires special attention in assessing thermodynamics and kinetics from those surfaces. The Gaussian approximation for the memory function might not be suitable to account for system's couplings with strong dielectric and/or ionic environments. The basic assumption used in the following is that of invariance of the energy landscape for different regimes of the memory friction.[24] Under such circumstances, we calculate the escape rates corresponding to all basin regions displayed on the characteristic PS of each of the three alanine hexapeptide analogs. These escape rates are computed for three given values of the memory friction parameter [see Eq. (10)]: $\rho_\alpha^{(1)}=0.1$, $\rho_\alpha^{(2)}=0.4$, and $\rho_\alpha^{(3)}=0.7$. For all these values, the correlation time t_c associated with the thermal bath is still smaller than the typical molecular vibrational period $(\omega_\alpha^{(0)})^{-1}$ but the system starts approaching non-Markovian behavior ($\rho_\alpha > 1$) from the ideal, Markovian limit ($\rho_\alpha = 0$). In the limit $\rho_\alpha \to 0$, we are able to compare our results with those obtained previously by the mean of the full master equation approach.[5]

As usual, we express the values of the escape rates in units of γ_α, which renders the scale of relaxation time $\sim \gamma_\alpha^{-1}$. Actually, each variety of kinetics features its own multiple time scale ($\cup_{\alpha=1}^{N} \gamma_\alpha^{-1}$), a consequence of the characteristic coupling in each basin α, as we discussed above.

C. Dynamics

In the following, we would like to see the relaxation of each of these three alanine hexapeptide analogs on their characteristic PS. Specifically, we would like to know how the probability evolves as the system surveys basin regions on each PS, with its own different energetic states and structural categories. Also, we are interested in finding how the topography governs the relaxation and kinetic traps. We compare their relaxation behaviors and infer from them the effect of the conformation constraints on the systems' dynamics. Finally, we point out the connection between the thermal bath properties and the ability of the system to relax on its characteristic PS. This can be inferred from the change of the tendency of relaxation of the system with the variation of the memory friction parameter ρ_α.

Prior to applying the IBM approach for the hexapeptide systems, the transition matrices, W_{ij}, were mapped to examine the validity of the separation of time scales between intra- and interbasin dynamics (Fig. 2). In these figures, the diagonal elements represent intrabasin dynamics and the off-diagonal elements represent interbasin dynamics. The white regions correspond to slow rates. It can be seen that for Ala_6 [Fig. 2(a)] as well as cyc–Ala_6 [Fig. 2(c)] both diagonal and off-diagonal elements include fast transition rates. However, for chrg–Ala_6 [Fig. 2(b)], only one diagonal element includes fast processes, indicating that intra- and interbasin motions occur on different time scales.

To study the state-to-state versus basin-to-basin dynamics, the master equation was solved for both situations by introducing the transition probability W_{ij} from state j to state i

$$\frac{\partial P_i}{\partial t} = \sum_j (W_{ij}P_j - W_{ji}P_i),$$

where P_i is the probability for finding the system in a state i, and by the transition probability $k_{\alpha\lambda}$ from basin α to basin λ

$$\frac{\partial P_\alpha}{\partial t} = \sum_\lambda (k_{\lambda\alpha}P_\lambda - k_{\alpha\lambda}P_\alpha), \tag{11}$$

where P_α is the probability that the system visits basin α. Solving the master equation for Ala_6, chrg–Ala_6, and cyc–Ala_6 by the state-to-state approach involves 280, 86, and 148 states, respectively. Within the IBM approach, these states are compressed to 17, 12, and 11 basins and all the topographical information (the positions on the energy spectrum of the minima and saddle) is transferred into the basin escape rates as described via Eqs. (3)–(5). The partition of the disconnectivity graphs in basins can be seen in Fig. 1.

The initial conditions for the stochastic simulations, both configurations and temperature, were similar to those set in the previous study.[5] Thus, the population probabilities for each system were calculated at 400 K, with the initial population all in the highest energy basin, which includes unfolded conformations. Figure 3 shows the time evolution of

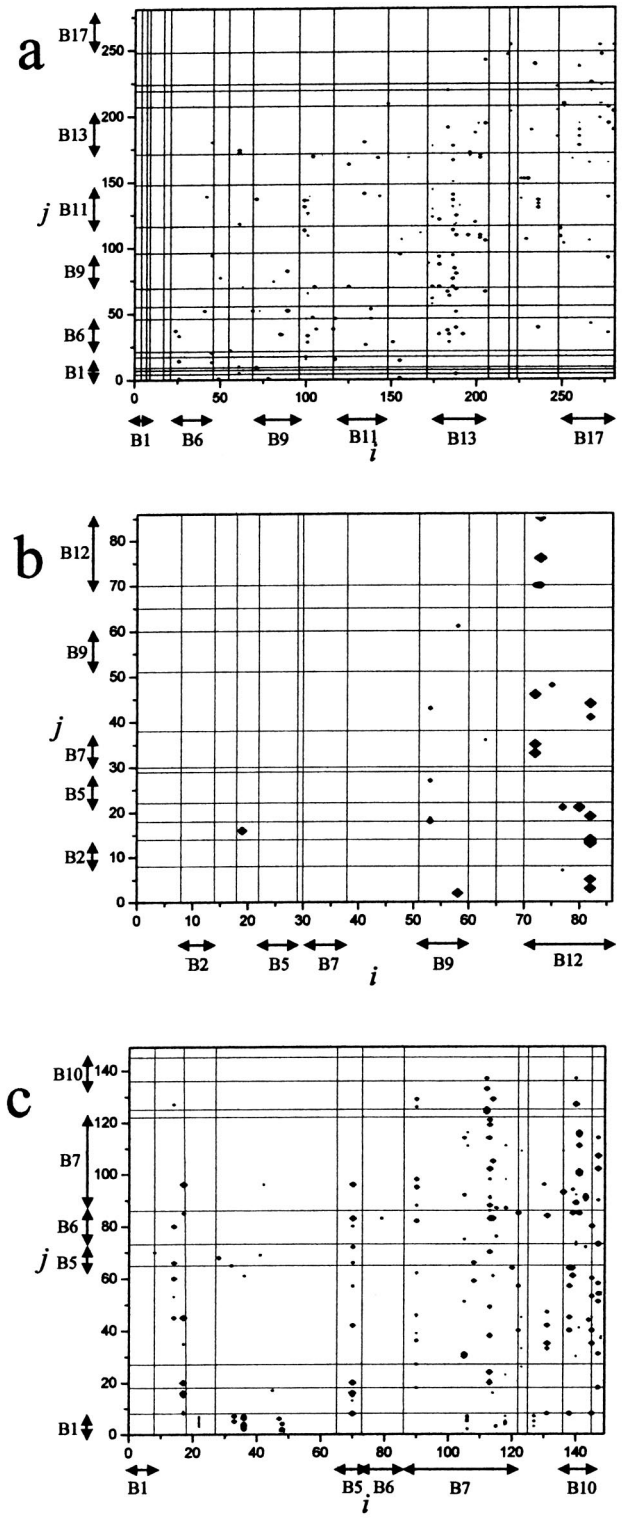


FIG. 2. The transition matrix W of (a) Ala_6; (b) chrg–Ala_6; and (c) cyc–Ala_6. Each matrix reflects the state-to-state transition probability (calculated by the TST) among the 280, 86, and 148 states of Ala_6, chrg–Ala_6, and cyc–Ala_6, respectively. The states are grouped according to the basin partitions as described by the disconnectivity graphs (Fig. 1). While the diagonal elements correspond to intrabasin dynamics, the off-diagonal elements describe interbasin dynamics. Above the diagonal are transition rates from low energy basins to a basin located higher on the graph. Below the diagonal are transition rates to a basin that branched off at a lower energy mode on the disconnectivity graphs.

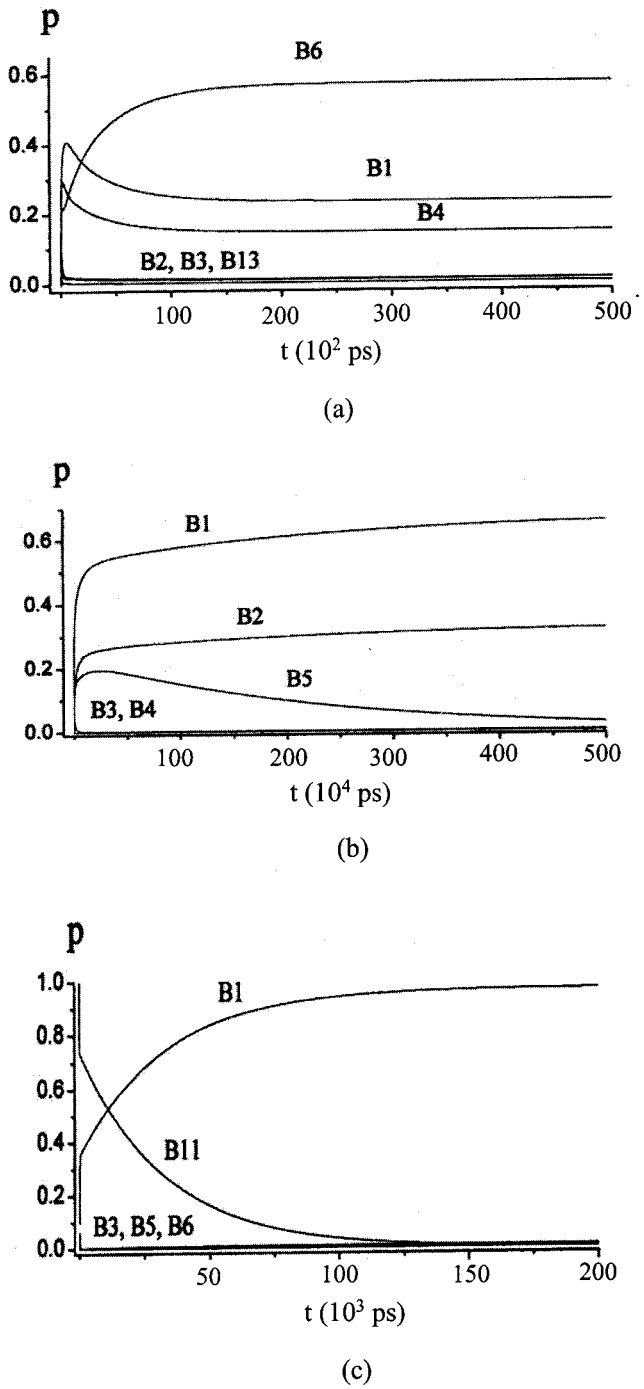


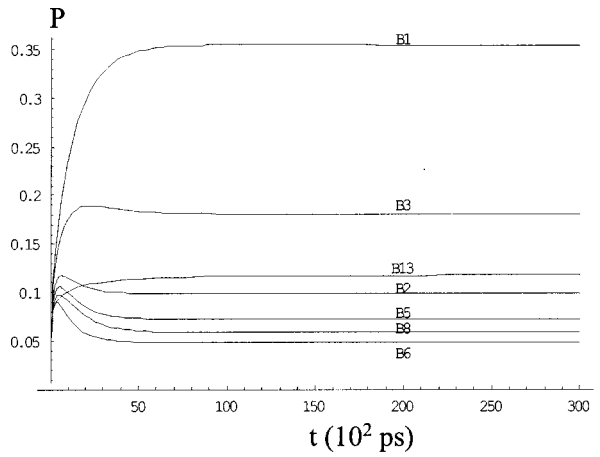
FIG. 3. The time evolution of the population probabilities of finding the (a) Ala_6; (b) chrg–Ala_6; and (c) cyc–Ala_6 in particular basins on their characteristic PS. The kinetic data were obtained by solving the master equation using the state-to-state approach and summing all the populations of the states that constitute each basin.

the basin populations as generated by the state-to-state movements. The population of the basin α, $P_\alpha(t)$, is obtained by summing the probability densities corresponding to all its states [i.e., $P_\alpha(t)=\Sigma_{i\in\alpha}P_i(t)$].

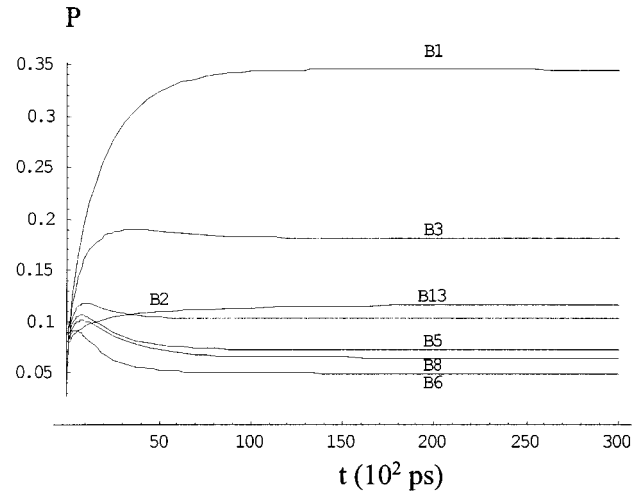
III. RESULTS

A. Ala_6 molecule

We first investigate the relaxation of Ala_6 under the conditions established above. Figure 4(a) depicts the time evo-



(a)



(b)

FIG. 4. (a) Time evolution of probability densities in basins (marked B_j, $j=1-3$, 5, 6, 8, 13 in order of increasing energy) on the characteristic PS of the linear Ala_6 polypeptide for $\rho_\alpha=0.1$. (b) For $\rho_\alpha=0.7$.

lution (in γ^{-1} units) of the probability densities in several of the most populated basins (branches) of its characteristic PS. In Table I we displayed the equilibrium probability densities corresponding to those basins. Our computation shows that these basins are $B1$, corresponding to the GM, $B2$, $B3$, $B5$, $B6$, $B8$, and $B13$. (Here, $\rho_\alpha=0.1$.) We observe that the increase of the population of the primary basin, containing the global minimum, is sharp and the probability reaches a plateau of about 35% (see the curve labeled $B1$). But, this is not the only basin dominating the PS. The basins $B2$ and $B3$ display also, in the asymptotic regime, relatively high weights, comparable to that of $B1$. The basin $B13$, a very deep narrow basin on the energy slab of 21 kcal/mol, acts as a significant kinetic trap. The accumulation regime of $B13$ is long (about 10^2 γ^{-1} units), predominating over the other probability densities, and the asymptotic limit has a high value ($\sim$12%), above the saturation limit of the low energy basin $B2$, for example. The Ala_6 linear hexapeptide is relatively inefficient at finding the region of the global minimum. Other higher energy basins ($B5$, $B6$, and $B8$) acquire significant populations in their long-time limits. One more thing needs to be stressed at this point. As we can see from Fig. 4, the basins $B2-B6$ undergo sequential kinetics, with probability profiles that first increase and then decrease with time. Therefore, dynamics in these regions of the PS is governed by a hierarchical relaxation provided by gentle funnel properties of the PS. In contrast, the basin $B13$, which is deep and narrow, has a pronounced unistep kinetics.

TABLE I. The equilibrium probability densities in basins on the characteristic PS of for the linear Ala_6 polypeptide.

B_1	B_2	B_3	B_5	B_6	B_8	B_{13}
0.34	0.1	0.17	0.07	0.05	0.06	0.12

The relaxation behavior of Ala_6 changes with changes of the properties of the thermal bath. Precisely, by acting on the parameters of the memory function (1), we may tune the strength of coupling between the system and the thermal bath. For example, we increase the value of the memory friction parameter from $\rho_\alpha=0.1$ to $\rho_\alpha=0.7$ and look at the relaxation behavior [see Fig. 4(b)]. The folding time t_f, defined as the starting point of the asymptotic regime of the probability density inside the basin of the GM, is increased by about 70% in this way. In addition, the change of the memory friction parameter induces variations on short time scales among the population distributions. Especially, high energy basins seem affected the most.

The state-to-state approach to dynamics of Ala_6 [Fig. 3(a)] supports the aforementioned results in the sense that, besides $B1$, several other basins ($B5$, $B6$, and $B8$) are significantly populated at equilibrium. However, the basin populations are different in these two results. This result was already expected since the diagonal and off-diagonal elements of the corresponding transition matrix [Fig. 2(a)] showed comparable values. This discrepancy suggests that the basin partition based on the kinetic connectivity and structural similarity is not the most relevant for studying the dynamics of Ala_6.

B. Chrg–Ala_6 molecule

In Fig. 5(a) we display the evolution in time of the probability densities inside the basins $B1-B5$ corresponding to the PS of chrg–Ala_6. Table II contains the values of the equilibrium probability densities in those basins. The population distributions over the other remaining basins are insignificantly small and are not shown. Here, the conformations corresponding to the energy slab 20–22 kcal/mol are taken together. As in the previous case, the memory friction parameter is set first to $\rho_\alpha=0.1$. We see that this system has a greater propensity for folding, which defines the system's predilection for populating low energy basins, than Ala_6. An important energy gap, about 6 kcal/mol [see Fig. 1(b)], separates the low energy basins from those of high energy. The presence of this energy gap might induce a higher accumulation in low energy basins. The probability density in the primary basin ($B1$), of the GM, shows a sharp increase in

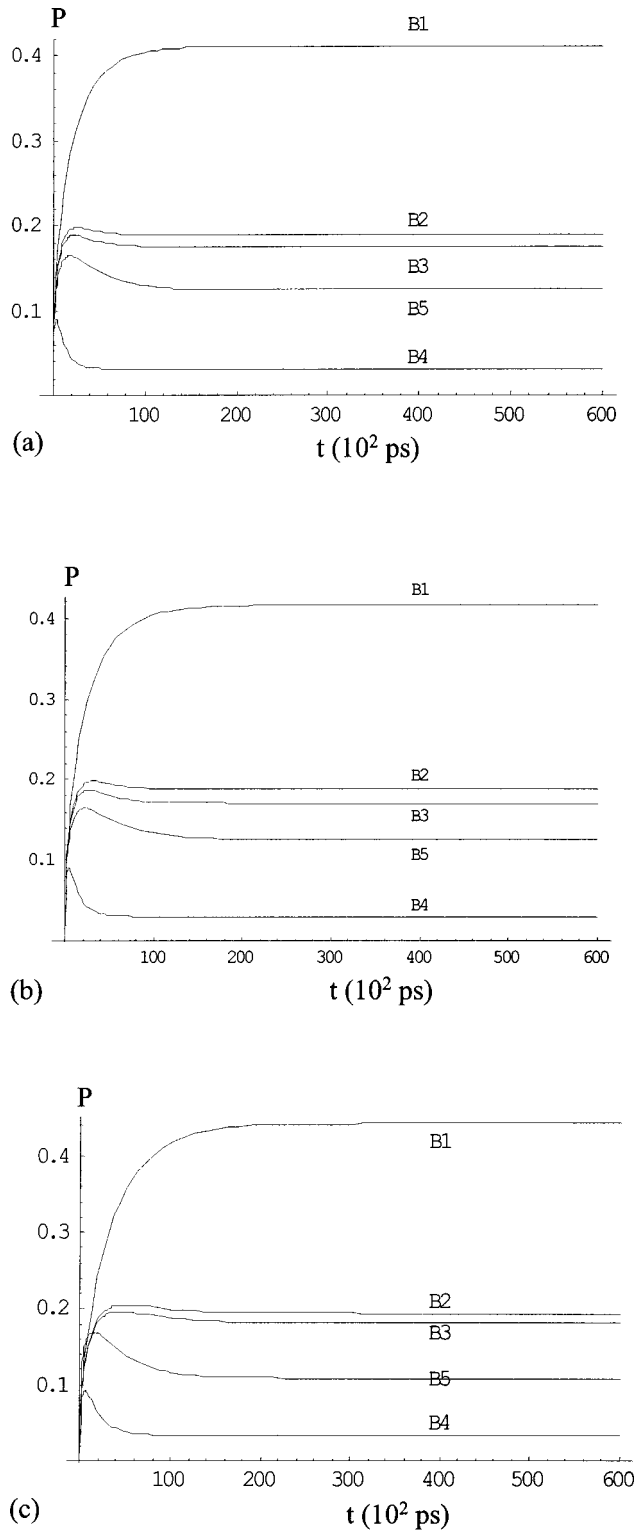


FIG. 5. (a) Time evolution of probability densities in basins (marked B_j, $j=1-5$ in order of increasing energy) on the characteristic PS of the chrg–Ala_6 polypeptide for $\rho_\alpha=0.1$. (b) For $\rho_\alpha=0.4$. (c) For $\rho_\alpha=0.7$.

time and reaches, in the asymptotic regime, a population of about 42%. The basins $B2$ and $B3$, which are below the energy gap, are highly populated in the asymptotic regime. The probability density in the basins below the energy gap is about 80%. We can also perceive a slight trend for sequential kinetics corresponding to these two low energy basins. The other basins, of higher energy, display more prominent peaks in their first increase of probability density distributions. For each of them, the subsequent decrease ends, however, with a low depletion limit. This sort of dynamical behavior may sign the presence of intermediates in the folding process. Also, we can see that the constraints on the flexibility of the molecule induced by electrostatic forces led to preclude the kinetic trap ($B13$) we evidenced in the previous case of the linear molecule.

TABLE II. The equilibrium probability densities in basins on the characteristic PS of for the linear chrg–Ala_6 polypeptide.

B_1	B_2	B_3	B_4	B_5
0.42	0.18	0.17	0.02	0.12

Again, the increase of the memory friction parameter induces different time scales for the survey of the configuration space of chrg–Ala_6. Roughly, the time scale becomes about 20 γ^{-1} units larger at each 0.3 increment of the memory friction parameter. In Figs. 5(a), 5(b), and 5(c), we see that the folding time increases from $t_f \simeq 160\ \gamma^{-1}$ to $t_f \simeq 200\ \gamma^{-1}$ when the value of ρ_α increases from 0.1 to 0.7. Also, the high energy population densities change on short time intervals. The memory friction effect alters the way in which the system approaches equilibrium.

The basin populations displayed in Fig. 5 are very similar to those obtained by applying the TST approach [Fig. 3(b)]. This indicates that, for chrg–Ala_6, the basin partition outlined from the kinetic connectivity is consistent with the complete separation of the intrabasin and interbasin dynamics, which is required by the condition for validity of the IBM theory. The result is in agreement also with the interpretation of the transition matrix [Fig. 2(b)].

C. Cyc–Ala_6 molecule

As we already noticed in the introductory part of this paper, the topography of the PS of cyc–Ala_6 is significantly different from those of the other two peptides. A dissimilarity in its relaxation behavior in comparison with previous cases is obviously expected. [For visualization, see Fig. 6(a).] The two competitive basins $B5$ and $B6$ marked at high energies on the disconnectivity graph act as kinetic traps. All three competing basins display exponential kinetics with saturation regimes and the corresponding time scales are similar. In the asymptotic regime, population distributions of basins $B5$ and $B6$ achieve high densities, about 20%. However, the probability density in the primary basin ($B1$) stays at its higher value, about 40%. (For the equilibrium probability densities in the most populated basins on the characteristic PS of this molecule, see Table III.) We also notice a striking behavior of the unfolded conformations in basin $B11$, which are highly populated even at long times. The depletion limit (12%) of the corresponding probability density in this basin is comparable with those of low energy basins (15% in $B3$, for example). The fact originates with the topographical pattern of that region on the characteristic PS that displays a relatively deep basin of unfolded states at the energy slab of 22 kcal/mol. The folding time is about 30 γ^{-1} units, which is

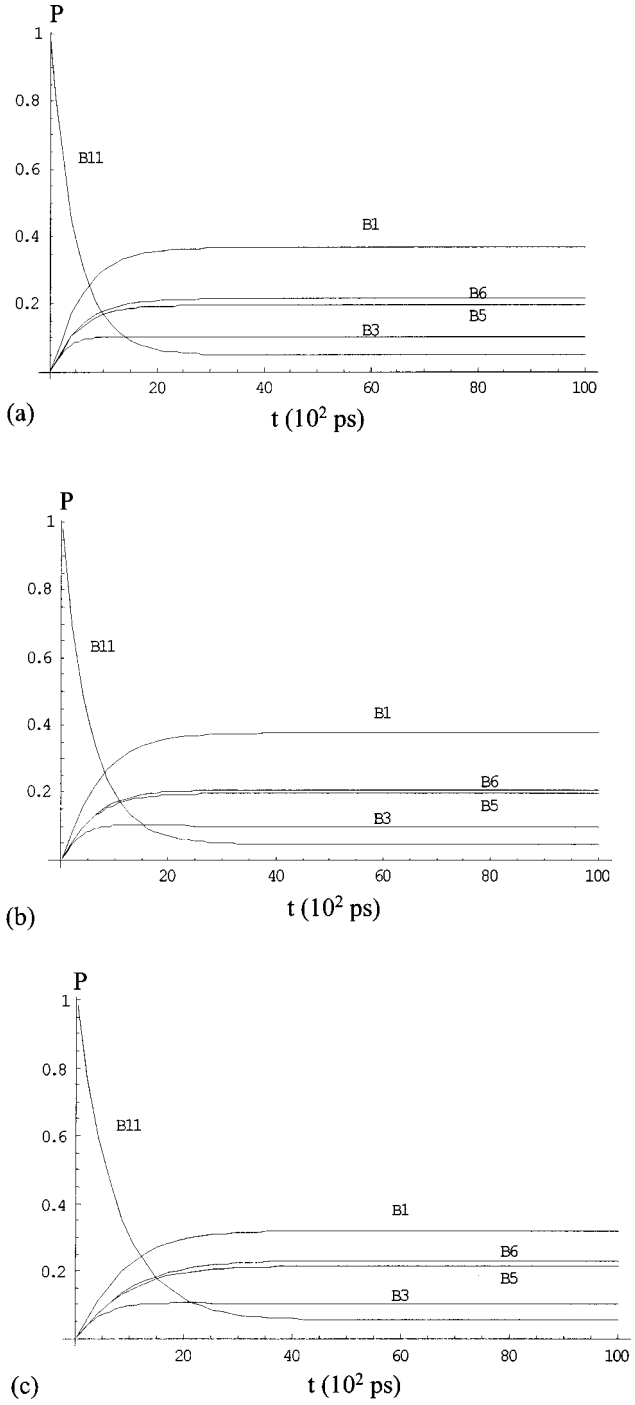


FIG. 6. (a) Time evolution of probability densities in basins (marked B_j, $j=1, 3, 5, 6, 11$ in order of increasing energy) on the characteristic PS of the cyc–Ala_6 polypeptide for $\rho_\alpha=0.1$. (b) For $\rho_\alpha=0.4$. (c) For $\rho_\alpha=0.7$.

much shorter than that we found in the previous case of chrg–Ala_6. This fact is a consequence of the decrease in depth of the low energy basins (from 10 to 5 kcal/mol). The trend of the folding time with changing the value of the memory friction parameter is similar to two other cases [see Figs. 6(b) and (c)].

The cyc–Ala_6 basin populations, as were found by the state-to-state dynamics with the aid of the TST [Fig. 3(c)], are in partial agreement with the basin populations found above. While both results agree that the global minimum is highly populated, the TST approach shows that two competing basins, $B5$ and $B6$, are negligibly populated, in contrast to their significant populations found by the IBM approach.

TABLE III. The equilibrium probability densities in basins on the characteristic PS of for the linear cyc–Ala_6 polypeptide.

B_1	B_3	B_5	B_6	B_{11}
0.36	0.1	0.19	0.21	0.05

One more thing calls for attention. If we assume the typical range for values of the friction constant, $\gamma \sim 10^{12}-10^{14}\ s^{-1}$, we then obtain folding times in the interval observed in the previous study.[5] If we assign a specific friction constant γ_α to each basin on the PS, the use of γ^{-1} units to scale the folding time is no longer appropriate. Obviously, for the present situation, with a unique, universal friction constant, a better description of the dynamical behavior of the system would be a combined representation of the probability densities obtained for the three cases ($\rho_\alpha=0.1$, 0.4, and 0.7). There are, however, certain limitations in our work which are due mainly to uncertainty of the data on viscoelastic response of vacuum.[25] We believe that the approximations we have been forced to use do not affect the qualitative results of this paper. Particularly significant is the result that the values of our rate constants are of the same order of magnitude as those from TST theory.

IV. CONCLUSIONS

We have illustrated, by using the mean solution of the reduced master equation for interbasin motion, the connection between topography and dynamics for the particular case of conformationally constrained peptides. Generally, we observed that the differences in PSs yield differences in dynamics as well. These kinetic implications have been revealed in the results of the calculations. Particularly, most of the information we extracted here with regard to the population distributions over the corresponding conformation space agrees reasonably well with that obtained by using the mean of the full master equation.[5]

In addition, the present study indicates that modifications of the relaxation behavior can be quantified in some cases and compared by an adequate partitioning of the PS. Focusing on the transitions between basins rather than between individual states appears, in principle, more informative. This strategy can be useful for comparing with experimental results.[26] Although the partition scheme of the PS based on the kinetic connectivity and structural similarity seems suggestive at first glance, it proved deficient in delineating correctly some basins on the PS of Ala_6 and cyc–Ala_6. Other routines might provide a subset of the characteristic PS with the required evidence for the complete separation of the intrabasin and interbasin dynamics.[19]

It has been suggested that the relaxation behavior depends not only on the topography of the PS but also on the way the various conformations of the system couple to the thermal bath. Each basin on the PS is characterized by a specific coupling between the system and the surrounding thermal bath. The strength of this coupling depends on both

solvent properties (viscosity, shear effects) and the structural variety of that basin, such as the mean overall radius of gyration or the extent of exposure of polar groups. We intend to use additional data to better understand the solvent effect and advance a suitable thermal bath description. Specifically, variational simulations of polyalanine embedded in various (implicit) solvents may yield useful information.[27] Implicit solvent treatments can still account for the local environments created by a chain as it folds by means of three-body correlations. Such correlations effectively allow us to deal with the participation of a third body in altering the dielectric environment surrounding a pairwise interaction. This effect is especially important when dealing with backbone desolvation as a means of stabilizing amide–carbonyl hydrogen bonds in spite of the net gain in self-energy of the intervening polar groups. The structuring, immobilization, and ultimate exclusion of water surrounding such bonds has a stabilizing effect that may be accounted for by rescaling the strength of such bonds according to the number of three-body correlations with vicinal hydrophobic groups.[28,29] Further implementation of this sort of information in the kinetic scheme enriches the precision with which we can characterize the ability of proteins to relax, preferentially, to only a limited number of geometrical structures from the vastly larger variety that the protein might exhibit.

ACKNOWLEDGMENTS

The authors would like to thank the National Science Foundation for its support of the work done at The University of Chicago.

[1] Y. Levy and O. M. Becker, Phys. Rev. Lett. **81**, 1126 (1998).
[2] O. M. Becker, Y. Levy, and O. Ravitz, J. Phys. Chem. B **104**, 2123 (2000).
[3] Y. Levy and O. M. Becker, J. Chem. Phys. **114**, 993 (2001).
[4] Y. Levy, J. Jortner, and O. M. Becker, Proc. Natl. Acad. Sci. U.S.A. **98**, 2188 (2001).
[5] Y. Levy, J. Jortner, and O. M. Becker, J. Chem. Phys. **115**, 10533 (2001).
[6] J. C. Light, J. Chem. Phys. **40**, 3221 (1995).
[7] P. M. Rodger and M. G. Sceats, J. Chem. Phys. **83**, 3358 (1985); M. G. Sceats, Adv. Chem. Phys. **70**, 357 (1988).
[8] A. Nitzan, Adv. Chem. Phys. **70**, 489 (1988).
[9] (a) R. S. Berry and R. E. Kunz, Phys. Rev. Lett. **74**, 3951 (1995); (b) R. E. Kunz and R. S. Berry, J. Chem. Phys. **103**, 1904 (1995); (c) K. D. Ball, R. S. Berry, R. E. Kunz, F.-Y. Li, A. Proykova, and D. J. Wales, Science **271**, 963 (1996).
[10] R. E. Kunz, *The Dynamics of First-Order Phase Transitions. In Mesoscopic and Macroscopic Equilibrium and Nonequilibrium Systems* (Harri Deutsch, Frankfurt/M, 1995).
[11] O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).
[12] D. J. Wales, M. A. Miller, and T. Walsh, Nature (London) **394**, 758 (1998); M. A. Miller and D. J. Wales, J. Chem. Phys. **111**, 6610 (1999).
[13] F. Despa and R. S. Berry, Eur. Phys. J. D **16**, 55 (2001).
[14] F. Despa and R. S. Berry, J. Chem. Phys. **115**, 8274 (2001).
[15] Y. Levy, J. Jortner, and R. S. Berry, Phys. Chem. Chem. Phys. **4**, 5052 (2002).
[16] F. Despa and R. S. Berry, in Proceedings of the 11th International Symposium on Small Particles and Inorganic Clusters, Strasbourg, 9–13 Sept. 2002; Eur. Phys. J. D (to be published).
[17] E. I. Shakhnovich, Phys. Rev. Lett. **72**, 3907 (1994).
[18] P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, Science **267**, 1619 (1995).
[19] R. S. Berry and R. E. Kunz, Phys. Rev. Lett. **74**, 3951 (1995).
[20] O. M. Becker, Proteins **27**, 213 (1997).
[21] G. H. Weiss, Adv. Chem. Phys. **13**, 1 (1966).
[22] T. Komatsuzaki and M. Nagaoka, J. Chem. Phys. **105**, 10838 (1996); Chem. Phys. Lett. **265**, 91 (1997); T. Komatsuzaki and R. S. Berry, J. Chem. Phys. **110**, 9160 (1999); J. Mol. Struct. **506**, 55 (2000); Phys. Chem. Chem. Phys. **1**, 1387 (2000).
[23] We have shown that, if we drop the assumption that the thermal bath responds instantly at the frequency of the reactive mode, at least inside the region with a rapid variation of the potential (the transition zone), then the memory effects change the barrier width [F. Despa and R. S. Berry (unpublished)]; R. S. Berry, J. Chem. Phys. **112**, 5223 (2000).
[24] G. R. Fleming, *Chemical Applications of Ultrafast Spectroscopy* (Oxford University Press, Oxford, 1986).
[25] See, for example, R. Golestanian and M. Kardar, Phys. Rev. Lett. **78**, 3421 (1997).
[26] See, for example, L. Pollak, M. W. Tate, A. C. Finnefrock *et al.*, Phys. Rev. Lett. **86**, 4962 (2001).
[27] A. Fernández and G. Appignanesi, Phys. Rev. Lett. **78**, 2668 (1997); A. Fernández, K. S. Kostov, and R. S. Berry, J. Chem. Phys. **112**, 5223 (2000).
[28] A. Fernández, J. Chem. Phys. **114**, 2489 (2001).
[29] A. Fernández, A. Colubri, and R. S. Berry, Physica A **307**, 235 (2002).