Energy-level statistics in the fine conformational resolution of RNA folding dynamics

Ariel Fernández,[1,2,*] Teresita Burastero,[1] Rodolfo Salthú,[1] and Ana Tablar[1]
[1]*Instituto de Matemática, Universidad Nacional del Sur, Consejo Nacional de Investigaciones Científicas y Técnicas, Avenida Alem 1253, Bahía Blanca 8000, Argentina*
[2]*Department of Chemistry and The James Franck Institute, The University of Chicago, Chicago, Illinois 60637*
(Received 20 April 1999)

This work is aimed at determining the energy-level statistics of the fine resolution of soft-mode dynamics warranting an adiabatically simplified structural relaxation of a folding biopolymer chain. The parameters defining the intrabasin structure relaxation are specified for RNA, so that each Watson-Crick base-pairing pattern may be treated as a quasiequilibrium ensemble of substates or torsional isomers within relevant folding time scales. The temperature-dependent threshold for energy dispersion associated with the fine structure of each superbasin is determined so as to warrant the adiabatic entrainment of the torsional dynamics.
[S1063-651X(99)17510-X]

PACS number(s): 87.15.He, 87.10.+e, 87.15.Nn

I. INTRODUCTION AND MOTIVATIONS

Most theoretical underpinnings of the expediency of the biopolymer folding process [1–3] hinge upon the assumption that intrachain contact patterns (CP's) or coarsely defined folding patterns may be treated as defining equilibrated superbasins in the potential energy surface (PES). Thus it is widely assumed that the thermalization time is incommensurably shorter than the relevant folding times. Each superbasin may be resolved into an ensemble of basins, each representing a finer conformational realization of the CP. Whether or not such fine structure needs to be taken into account to simulate the folding dynamics is subject to debate, since the accuracy of such simulations is not simply determined by the time-scale window under scrutiny [3–8].

In this regard, the concept of substate has been introduced based on experimental evidence of the fine structure of the PES, which becomes dynamically relevant at very low unphysiological temperatures [5]. On the other hand, recent theoretical approaches reveal that a CP in a protein may be resolved into substates, each of which is defined by the discrete number of torsional isomeric states that the free residues—that is, those not engaged in the CP—may adopt [3,4]. In simple terms, this means that each CP could be regarded as a set of constraints imposed upon certain chain units, specifically those that are engaged in the formation of the contacts, including the concurrent loops, while the other remaining units are free to adopt more than one significant torsional isomeric state. The multiplicity of each CP basin, that is, the number of substate basins contained in it, is simply given by the number of different sequences of local torsional states. There are fixed elements in such sequences, namely those associated with the units engaged in the CP itself.

Obviously, the fine structure of the CP basin depends on the level of resolution of local conformations of the chain as significant rotamers. Since free units may adopt a discrete number of local isomeric states and each free unit is effectively uncorrelated, a reasonable random energy distribution of substates should apply within each CP basin: Correlations become effective through intrachain contacts and the concurrent local conformational demands imposed upon all residues involved in the formation of the structural motifs [3]. On the other hand, transitions between substates within a CP superbasin eventually trigger the intersuperbasin transitions. This occurs, for example, once a consensus window of ''correct'' torsional states [6] appears among the free residues of the original CP, leading to the formation of a new CP [4,7].

In the realm of RNA folding, a drastic separation of relaxation time scales has been implicitly incorporated in most—but not all [7]—predictive algorithms [8], whereby the CP superbasin is assumed to be equilibrated or thermalized within time scales incommensurably shorter than those relevant to significant folding events resolved as CP transitions [2]. Thus an adiabatic ansatz underlies this picture in which the folding dynamics have been largely assumed to be dictated by the possibilities of intrachain base pairing involving distant nucleotides (units), according to the so-called Watson-Crick complementarity map. To the best of our knowledge, the validity of this ansatz, which disregards the finer details of the torsional dynamics of the RNA chain, has not hitherto been established and thus becomes the aim of this work.

In order to study this problem, the fine structure of CP superbasins needs to be elucidated for the RNA folding process. Unfortunately, unlike in the case of proteins, where the so-called Ramachandran maps define in a clearcut way the small number—two to four—of local torsional isomeric states that each unit may adopt throughout the folding process [9], the data on the fine conformational structure of CP's in RNA is scant. While in proteins we may visualize local torsional isomers as basins in the Ramachandran maps that define the local torsional dynamics of the chain, the RNA backbone is far too complex (seven torsional degrees of freedom in contrast to the two in each protein residue) to allow for such a simplified resolution of the CP fine structure or a straightforward computation of the CP degeneracy. For instance, the number $M(i)$ of substates of a protein chain of length N in a specific CP denoted i is readily computed as $M(i)=\Pi_{n=1,\ldots,N} q_i(n)$, where $q_i(n)$ is the number of Ramachandran basins available to residue n within CP i. Thus

*Author to whom correspondence should be addressed.

$q_i(n)=1$ if n is engaged (that is, not free) in CP i and $2 \leqslant q_i(n) \leqslant 4$, otherwise, depending on the type of residue at contour position n.

These facts prompt us to resort to a statistical treatment to deal with the fine resolution of each CP superbasin in the PES for RNA. The reader should note that, in spite of the wanton complexity of the RNA backbone, the torsional dynamics has almost invariably been considered enslaved or subordinated by the adiabatic dynamics resolved as transitions between Watson-Crick intrachain patterns (CP's), at least within the time scale (1 μs to seconds) relevant to folding [2,7,8].

In statistical terms, soft torsional degrees of freedom of the flexible chain may be integrated out as conformational entropy of the CP state for relatively long time scales. Thus the adiabatic treatment resolves the folding pathway as a sequence of elementary or single-step CP transitions, each one regarded as an activated process with a single kinetic barrier, and thermal equilibration within CP basins leads to an Arrhenius-type kinetics at the coarse level of resolution [2,3]. This adiabatic treatment hinges upon the hypothesis that CP's are quasiequilibrium states. As such, it will be subject to scrutiny in this work, where we shall provide an estimation of the semiempirical statistical parameters for intra-CP-superbasin relaxation that warrants the conventional treatment of CP transitions using unimolecular Arrhenius-like rate constants.

II. LOCAL VALIDITY OF THE RANDOM ENERGY MODEL

The random energy model (REM) distribution of energy levels [5,10] within a superbasin has been successfully introduced to explain the phenomenological recombination kinetics of biopolymer-ligand interactions [5,11]: The x-ray structure of a folded protein is regarded as an average over a statistically large number of substates [5] that can exchange faster than the observation time scale inherent in x-ray determination. Thus it may be assumed that specific folding patterns may be viewed as superbasins containing a number of energy levels that are uncorrelated within the same superbasin. This local randomness implies that the source of correlation materializes only at the level of resolution at which structural changes involve changes in the enthalpic content, in turn determined by the contact pattern, and not at the finer level of structural detail in different torsional realizations of a CP with the same enthalpic content [11]. This generic idea enables us to statistically characterize the relaxation within a CP basin i by a single parameter: the dispersion μ_i of the Gaussian distribution of energy levels around the thermal average $\langle E\rangle_i$ due to the slightly different energies of torsional isomers, all belonging to the same CP class i.

Thus the adiabatic condition may be given as $\tau_i/\tau_{ij} \ll 1$ for each CP superbasin i and all CP superbasins j in the $J(i)=$class of all CP's connected to i by an elementary step. Here τ_i is the mean equilibration time within basin i and τ_{ij} is the mean first passage time for the transition $i \rightarrow j$ viewed within a renormalized bistable potential with minimum well energies $\langle E\rangle_i$ and $\langle E\rangle_j$. Explicitly, if $E_i^{\neq}$ represents the expected energy of the generic transition state [12] for substate transitions within basin i, $E_{ij}^{\neq}$ denotes the adiabatic energy of the transition state corresponding to the $i \rightarrow j$ transition, and f indicates the effective preexponential frequency [3], we get

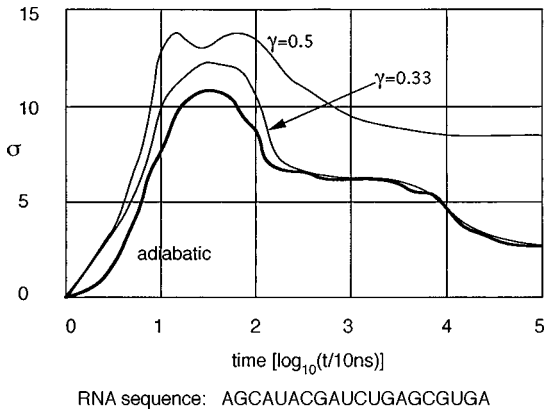

FIG. 1. Plots of the Shannon information entropy σ relative to the CP partition of conformation space for a randomly generated RNA sequence of length $N=20$. The letters A, U, G, and C denote the four types of units with Watson-Crick complementarity A-U, G-C. The reduced working temperature is $T'=|T-T_c/T_c|=0.03$. The thick line plot corresponds to the adiabatic computation and the thin lines to semiempirical REM computations with different upper bounds for the intra-CP-superbasin dispersions.

$$\tau_i = \max_{(E\text{'s in CP superbasin } i)}\{f^{-1}\exp[(E_i^{\neq}-E)/RT]\}, \tag{1}$$

$$\tau_{ij} = f^{-1}\exp\{[E_{ij}^{\neq}-\langle E\rangle_i]/RT\}, \tag{2}$$

$$\tau_{ij}(\text{eq.}) = f^{-1}\exp\{[E_{ij}^{\neq}-(\langle E\rangle_i-\mu_i^2/RT)]/RT\}. \tag{3}$$

Equation (3) defines the thermalized transition time and is valid since the equilibration energy within the CP superbasin i is $E(\text{eq.})_i=\langle E\rangle_i-\mu_i^2/RT$ (cf. [5,11]). Thus the question addressed in this work is: How can the statistical microscopic parameters μ_i's be determined so that the adiabatic approximation becomes a valid projection in the long-time limit for the detailed dynamics resolved at the finer substate level?

To answer this question, we shall first consider a marker indicating the time range of validity of the adiabatic ansatz [7]. One such marker is the information entropy σ taken with respect to the fixed partition of conformation space in mutually disjoint CP classes. This coarse information entropy is an indicator of the spreading in the CP-population dynamics and serves as a marker of the robustness and expediency of the folding process [3,7]. This is so since the CP-population dynamics may be generated by adiabatically integrating the underlying torsional motion or, alternatively, incorporating it in a REM-like model. Thus we may gauge the validity of the adiabatic approximation by comparing both resulting time-dependent σ plots (cf. Fig. 1).

The entropy σ measures the spreading of the probability distribution vector $P(t)=P_1(t),\ldots,P_M(t)$, where $P_j(t)$, $j=1,\ldots,M$, indicates the probability that a chain is folded into the CP j at time t, and $M \sim O(\exp N)$ is the total number of *a priori* possible CP's for a fixed RNA sequence of length N. These probabilities should be interpreted in a Gibbsian sense, as we have a statistically large number ($\sim 10^{20}-10^{23}$ per unit volume) of replicas of our system given by actual RNA molecules that are folding onto themselves as soon as renatur-

ation conditions are established or recovered in the environment. Thus the information entropy associated with the folding process resolved at the CP level is

$$\sigma(t) = -\sum_{j=1,\ldots,M} P_j(t) \ln P_j(t). \tag{4}$$

An adiabatically defined stochastic process governs the flow of probability [2,3]. This process is determined by the activation energy barriers required to produce or dismantle interactions that stabilize the CP's. Thus, at each instant, the partially folded chain undergoes a series of disjoint elementary events with transition probabilities dictated by the unimolecular rates of the events. This coarsely resolved stochastic process is Markovian, since the choice of the set of disjoint events at each stage of folding is independent of the history that led to the particular state [3].

In order to compute the probability distribution at any given time and the resulting behavior of σ, we first discretize time t, coarse-graining it by multiples of u, the shortest mean time for a CP transition: $t = t'u$, $u = \min_{(i,j)} \tau_{ij}$. Then, if $\mathbf{U}$ represents the stochastic transition matrix at the CP level, we get

$$P(t) = [\mathbf{U}]^{t'} P(0) \quad \text{with} \quad t = t'u, \tag{5}$$

where the matrix element $[U]_{ij}$ is given by

$$[U]_{ij} = \left[k_{ij} \Big/ \sum_{j' \in J(i)} k_{ij'} \right]. \tag{6}$$

In Eq. (6), k_{ij} denotes the adiabatic unimolecular rate constant for the CP transition $i \to j$, $J(i)$ is the set of CP's accessible from i through elementary transition steps involving surmounting a single kinetic barrier (see below), the factor $[k_{ij} / \Sigma_{j' \in J(i)} k_{ij'}]$ represents the probability for the transition $i \to j$ dictated by kinetic control within a time span of the order of $\tau_{ij} = k_{ij}^{-1}$.

Explicit values of the unimolecular rate constants require an updated compilation of the thermodynamic parameters at renaturation conditions [8]. These parameters are used to generate the set of kinetic barriers associated with the formation and dismantling of stabilizing interactions, the elementary events in our context of interest. Thus the activation energy barrier $[E_{ij}^{\neq} - \langle E \rangle_i]$ for the rate-determining step in the formation of a stabilizing $i \to j$ interaction is $-T\Delta S_{\text{loop}}$, where ΔS_{loop} indicates the loss of conformational entropy associated with closing a loop. Such a loop might be of any of four admissible classes: bulge, hairpin, internal, or pseudoknotted [2,3,8]. For a fixed number L of unpaired bases in the loop, we shall assume the kinetic barrier to be the same for any of the four possible types of loops. This assumption is warranted, since the loss in conformational entropy is due to two overlapping effects of different magnitude: the excluded volume effect, meaningful for relatively large L ($L \geqslant 100$) and the orientational effect that tends to favor the exposure of phosphate moieties towards the bulk solvent domain for better solvation. Since both effects are independent of the type of loop, we may conclude, in relatively good agreement with calorimetric measurements, that the kinetic barriers are independent of the type of loop for fixed L [8,13]. On the other hand, the activation energy barrier associated with dismantling a stem is $-\Delta H(\text{stem})$, the amount of heat released due to base pairing and stacking when forming all contacts in the stem.

For completion we shall give the analytic expressions for the adiabatic unimolecular rate constants k_{ij}'s. For clarity of notation we shall drop the subindexing, since we shall focus each time on a specific CP transition. If the transition happens to be a helix decay process, we obtain

$$k = fn \exp[G_h / RT], \tag{7}$$

where n is the number of base pairs in the helix formed in the $i \to j$ step, $f \approx 10^6\ \text{s}^{-1}$ is the fixed effective frequency of successful collisions [2,10,13], and G_h is the (negative) free energy contribution resulting from stacking of the base pairs in the helix. Thus the essentially enthalpic term $-G_h = -\Delta H(\text{stem})$ should be regarded as the activation energy for helix disruption. On the other hand, if the transition happens to be formation of a stabilizing interaction, the inverse of the mean time for the transition will be given by

$$k = fn \exp[-\Delta G_{\text{loop}} / RT], \tag{8}$$

where $\Delta G_{\text{loop}} \approx -T\Delta S_{\text{loop}}$ is the change in free energy due to the closure of the loop concurrent with helix formation. Thus, the kinetic barrier is in this case due to the cost in conformational freedom demanded by the need to bring together the portions of the chain to be engaged in the putative favorable interaction.

Working equations (4)–(8) define the adiabatic folding dynamics in a statistical sense. The expeditious nature of the folding process is monitored through the time dependence of the information entropy, as displayed in Fig. 1. As direct inspection of the thick line plot (the adiabatic computation) in Fig. 1 reveals, incipient or imperfect helices formed in structure-nucleation events in the range 5×10^{-9}–5×10^{-7} s are easily dismantled (a bubble is more easily formed than in fully developed structures). This fragility of incipient structures causes the large CP fluctuations marked by a large σ in the time range 10^{-7}–5×10^{-6} s. After all ephemeral misfoldings have been dismantled, a large plateau starting at 2.4×10^{-6} s marks the formation of a relatively stable kinetic intermediate that contains all structural motifs that may form noncooperatively [14]; that is, those motifs whose associated $L = L(\text{loop})$ lies within the favorable ranges of low conformational entropy cost: $3 \leqslant L(\text{loop}) \leqslant 10$ [9,10]. At 10^{-5} s, cooperative events lead to other helices whose loops have favorable renormalized sizes, while their sizes relative to the random coil are unfavorable [3,14]. On the other hand, the increase in base-pair stacking beyond the formation of the kinetic intermediate stabilizes the patterns determining its survival. This determines the relatively low fluctuations beyond 10^{-4} s.

In estimating the time range where CP's may be treated as quasiequilibrium states, we face the following problem: How do we compare the data generated by the adiabatic approximation in Fig. 1 with a semiempirical REM computation of torsional dynamics?

The REM computation requires a Gaussian distribution peaked at the expected energy $\langle E \rangle_i$ of the density $\beta_i(E)$ of torsional substates with energy E contained in the generic CP basin i. Thus we get

$$\beta_i(E)=(2\pi\mu_i)^{-1}\exp[-(E-\langle E\rangle_i)^2/2\mu_i^2]. \quad (9)$$

Then, the conditional probability $p_i(E,t)$ of finding a sample molecule in a substate with energy E, given that we know that it belongs to CP class i at time t, obeys the REM master equation

$$\partial p_i(E,t)/\partial t=-p_i(E,t)\int K_i(E',E)dE' + \int K_i(E,E')p_i(E',t)dE', \quad (10)$$

where the kinetic kernel $K_i(E,E')$ is given by

$$K_i(E,E')=\beta_i(E)f\exp[-(E_i^{\neq}-E')/RT]. \quad (11)$$

The stationary probability distribution $(\partial p_i(E,t)/\partial t=0)$ within CP basin i obtained from Eq. (10) for t in the range $\tau_{ij}\geqslant t\gg\tau_i$ begets a stationary distribution of kinetic barriers for the $i\rightarrow j$ transition: Using the relation $h_{ij}=h_{ij}(E)=f\exp[-(E_{ij}^{\neq}-E)/RT]$, where $h_{ij}(E)$ is the unimolecular rate constant for the CP transition $i\rightarrow j$ from a substate in CP basin i with energy E, we obtain a stationary distribution $H_{ij}=H_{ij}(h_{ij})$ of unimolecular rate constants h_{ij}'s associated with the CP transition $i\rightarrow j$ from different substates contained in CP basin i. Thus

$$H_{ij}(h_{ij})=\left\{\int[\tau_{ij}(\text{eq.})h_{ij}]^{q_i}\times\exp[-\tau_{ij}(\text{eq.})h_{ij}]dh_{ij}\right\}^{-1}[\tau_{ij}(\text{eq.})h_{ij}]^{q_i}\times\exp[-\tau_{ij}(\text{eq.})h_{ij}] \quad (12)$$

and

$$\int H_{ij}h_{ij}dh_{ij}=\langle h_{ij}\rangle\approx h_{ij}^*=q_i/\tau_{ij}(\text{eq.}), \quad (13)$$

where h_{ij}^* denotes the most probable value of h_{ij}. The basin exponent q_i is obtained from the adiabatic approximation $q_i/\tau_{ij}(\text{eq.})\approx k_{ij}$, whose validity depends on the relative size of μ_i. Thus Eq. (13) yields the estimation

$$q_i=\exp[-(\mu_i/RT)^2]. \quad (14)$$

At this point, we may construct a CP transition probability matrix, **U**(REM), resulting from a REM semiempirical treatment of substate dynamics within CP basins:

$$[U(REM)]_{ij}=\int H_{ij}h_{ij}dh_{ij}. \quad (15)$$

Thus the validity of the adiabatic ansatz may be probed by comparing the time-dependent behavior of the Shannon information entropies with respect to the CP partition obtained respectively from the stochastic processes defined by transition matrices **U** and **U**(REM). Since the latter matrix is parametrically dependent on the family $\{\mu_i\}$, we need to establish upper bounds $\mu_i(\max)$ on the μ_i's that still warrant the validity of the adiabatic ansatz. The following fit is shown to be adequate (cf. Fig. 1):

$$\mu_i(\max)=6.8\times10^{-3}|T_c/(T-T_c)|^{\gamma}\langle E\rangle_i, \quad (16)$$

where $T_c\approx318$ K denotes the denaturation temperature (cf. [8]), and γ is estimated at 0.33 (cf. Fig. 1). Thus a relatively large temperature T gives more latitude in the choice of the dispersion parameters within which the adiabatic ansatz remains valid (alternatively, large μ_i's demand a large T to reach superbasin equilibration required by the adiabatic ansatz). This fact stemming from direct examination of Eq. (16) reflects the faster thermalization within CP superbasins that takes place as T is raised.

III. RESULTS

At this point we must compare the adiabatic statistical dynamics with the rigorous REM computation in order to establish the validity of the former *vis-à-vis* our estimation of the semiempirical microscopic parameters μ_i's. For short time scales $10^{-8}-10^{-6}$ s, fast-evolving internal degrees of freedom simulated as yielding uncorrelated substates are not yet enslaved or entrained by CP transitions that evolve within larger time scales of the order of $2\times10^{-6}-10^{-2}$ s for the chain length $N=20$. For this reason, within the range $10^{-8}-10^{-6}$ s, the level of exploration of conformation space due to uncorrelated or short-range correlated torsional excitations must be vastly larger than that resulting from an adiabatic process, as shown in Fig. 1. However, as soon as the stabilized kinetic intermediate is formed [14], the long-range correlations, coupling distant units in the RNA chain, begin to develop. Thus cooperative effects occur upon short-range nucleating interactions. These long-range correlations are in turn induced by CP transitions. Thus initial structure-nucleating steps involving uncorrelated or locally correlated motions do not demand as much enslavement of fast-evolving torsions reflected as substate hopping as cooperative events, which entail long-range correlations. For this reason, we expect the adiabatic approximation to fit the REM results as soon as long-range correlations governed by CP transitions occur. This is indeed what takes place, as the almost perfect coincidence of both the adiabatic and REM plots beyond 2.4×10^{-6} s reveals (Fig. 1).

Both the adiabatic and the REM-based plots reveal an almost perfect coincidence with higher than 92% agreement beyond 3×10^{-6} s. The discrepancy between the adiabatic and REM $\sigma(t)$ plot rises to an upper bound of 12% within the range $10^{-7}-10^{-6}$ s. This is clearly due to the microscopic origin of fluctuations that becomes apparent at shorter time scales and is therefore only effectively captured by the REM dynamics. An inspection of Fig. 1 reveals that the REM dynamics become effectively entrained over the longer time scales relevant to folding.

Summarizing, the initially large fluctuations and discrepancies observed in both the adiabatic and REM computation of the Shannon entropy correspond to the formation of noncooperative ''misfolded'' structures, that is, CP's that are ephemeral enough so that their lifetimes are comparable with thermalization or equilibration times scales within CP superbasins. Thus substate hopping during these early stages of folding becomes important and determines the difference between the adiabatic and REM dynamics. The effect of this hopping vanishes as soon as there exists a sharp separation between thermalization and CP transition time scales. This requires formation of ''better'' CP's with long enough lifetimes. Most of these CP's yield a kinetic intermediate made

up of a fairly stable cluster of kinetically related structures [7,10]. The existence of such a dynamic intermediate state is confirmed by the existence of a plateau sustained within the 2–100 μs time-scale range.

As displayed in Fig. 1, σ does not tend to zero in the long-time dynamics relevant to the folding time-scale frame. Rather, the coarse entropy decreases asymptotically to a plateau value $\sigma = 2.4$ valid for $N = 20$. This reflects the fact that folding into a unique structure, reaching a sharply peaked probability distribution within biologically relevant time scales is not a generic feature of the long-time chain dynamics. However, the almost perfect coincidence between the long-time adiabatic and the projected soft-mode torsional behavior captured by the semiempirical REM model is indeed a generic feature of RNA folding because it was obtained irrespective of natural selection, revealing the inherently REM-like structure of the intrabasin soft-mode dynamics and the validity of a quasiequilibrium assumption on base-pairing patterns in the long-time limit.

IV. DISCUSSION

The vast gap in time scales separating molecular dynamics simulations (1 ps to 10 ns) from meaningful folding events (1 μs to 100 s) imposes a formidable challenge to the rigorous theoretical underpinnings of such basic properties as the robustness and expediency of the folding process when viewed from a microscopic perspective. Such difficulties are all the more apparent as one attempts to elucidate the bearing of the torsional dynamics upon the folding process in the long-time limit. Thus two different coarse grainings of torsional conformation space have been adopted in this work, and their validity *vis-à-vis* the purported objective has been tested within relevant time scales.

The crudest coarse graining is represented by the adiabatic approximation whereby contact patterns (or intramolecular base-pairing patterns) are treated as quasiequilibrium states [3,12–14]. In spite of its crudeness and limitations, this level of approximation is endowed with predictive value, as revealed by the computation of dominant folding pathways resolved at this coarse level and contrasted with experimentally probed kinetic bottlenecks [2,14]. Of course, one would ideally wish to validate this approach from first principles. This requires incorporating the torsional motion of the chain, at least in the long-time limit [15]. A first step in this direction has been undertaken in this work: The coarse description has been refined by resolving each CP superbasin as a statistical ensemble of basins, each representing a crudely defined torsional state of the chain, or substate. In turn, each torsional state is given by specifying the torsional isomeric state (rotamer) of each uncorrelated unit in the RNA chain. That is, the existence of a statistical ensemble of uncorrelated energy levels stems from the fact that the torsional isomeric states of different free nucleotides within a CP are uncorrelated, as dictated by the absence of significant long-range interactions that would lower the enthapic content below its value for the given CP. This is precisely why we can introduce a REM-like description of the fine structure of the CP superbasin by regarding each energy level as being associated with a sequence of isomeric states, one for each of the units that are not engaged in the CP. This hierarchical resolution of the basin structure is in the spirit of the generic principles put forth by Becker and Karplus [16].

The results of this work reveal that the master equation dynamics drawn upon the REM local refinement may be coarse grained to a simpler master equation, since REM thermalization within CP superbasins holds valid for adequate energy dispersions in the long-time limit relevant to the folding process. Analogous reductions in hierarchical topographies have been generically introduced [16] but without reference to a local REM-like fine structure or within the RNA folding context.

The coarsening of torsional RNA dynamics implemented in this work is semiempirical insofar as it does not follow from a systematic application of a projection technique required to rigorously obtain the subordinating or enslaving modes from the full torsional dynamics (cf. [15]). On the other hand, our approach is justified because it provides crucial information on the way in which the PES topography for RNA torsional dynamics is explored: All torsional states realizing a specific CP may be grouped in a single basin that is thermalized in the long-time limit relevant to the folding process. This resolution has been elucidated by adopting the REM ansatz.

A more detailed study using projection operator techniques would be required to validate this intuitive assumption. Such a treatment is beyond the scope of this work, as it requires a correct mode-coupling theory and an estimation of the torsional state energies, followed by a thorough statistical analysis, in order to single out uncorrelated clusters of energy levels within a correlated long-range hierarchy. Thus, although a rigorous treatment probably represents a daunting task, this work provides a heuristic way to coarse grain torsional dynamics so that an essential feature becomes apparent: The adiabatic simplification whereby CP's are treated as quasiequilibrium states is valid precisely in the time limit that is relevant to the folding process. Furthermore, this crude level of description implies a sequence of local thermalizations in the PES topography accessible to the system. This renormalization of the PES is a generic feature insofar as it is independent of whether or not the RNA chain is a product of natural selection.

A. F. acknowledges financial support from the Fulbright U.S.-Argentina Bilateral Commision, the J. S. Guggenheim Memorial Foundation, and the Alexander-von-Humboldt Stiftung. The kind hospitality of R. S. Berry at the University of Chicago is gratefully acknowledged.

[1] O. B. Ptitsyn and G. V. Semisotnov, in *Conformations and Forces in Protein Folding,* edited by B. Nall and K. Dill (American Association for the Advancement of Science, Washington, DC, 1991), pp. 155–168.

[2] P. Zarrinkar and J. Williamson, Science **265**, 918 (1994); A. Fernández, Ann. Phys. (Leipzig) **4**, 600 (1995).

[3] A. Fernández and A. Colubri, Physica A **248**, 336 (1998); J. Math. Phys. **39**, 3167 (1998).

[4] A. Fernández, Phys. Rev. E **59**, 5928 (1999).
[5] H. Frauenfelder, S. G. Sligar, and P. Wolynes, Science **254**, 1598 (1991); H. Frauenfelder, G. Petsko, and D. Tsernoglou, Nature (London) **280**, 558 (1979).
[6] R. Zwanzig, Proc. Natl. Acad. Sci. USA **92**, 9801 (1995).
[7] A. Fernández, J. Stat. Phys. **92**, 237 (1998).
[8] A. E. Walter, D. H. Turner, J. Kim, M. H. Lyttle, P. Müller, D. H. Mathews, and M. Zuker, Proc. Natl. Acad. Sci. USA **91**, 9218 (1994); J. A. Jaeger, D. H. Turner, and M. Zuker, *ibid.* **86**, 7706 (1989).
[9] J. M. Thornton, in *Protein Folding*, edited by T. Creighton (W. Freeman, New York, 1992), pp. 59–64.
[10] E. I. Shakhnovich and A. M. Gutin, Europhys. Lett. **9**, 569 (1989); A. Fernández, Phys. Rev. Lett. **65**, 2259 (1990).
[11] A. Fernández, Ber. Bunsenges. Phys. Chem. **98**, 260 (1994).
[12] A. Fernández and E. I. Shakhnovich, Phys. Rev. E **42**, R3657 (1990).
[13] A. Fernández, H. Arias, and D. Guerín, Phys. Rev. E **52**, R1299 (1995).
[14] G. Appignanesi and A. Fernández, J. Phys. A **29**, 6265 (1996).
[15] C. J. Camacho and D. Thirumalai, Proc. Natl. Acad. Sci. USA **90**, 6369 (1993).
[16] O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495 (1997).