

Adiabatic ansatz in RNA folding dynamics

Ariel Fernández,¹ Gustavo A. Appignanesi,¹ and Ruben A. Montani²

¹*Instituto de Matemática de Bahía Blanca (INMABB-CONICET), Universidad Nacional del Sur, Avenida Alem 1253, 8000-Bahía Blanca, Argentina,*

²*Laboratorio de Fisicoquímica, Departamento Química e Ingeniería Química, Universidad Nacional del Sur, Avenida Alem 1253, 8000-Bahía Blanca, Argentina*

(Received 6 January 1997)

A statistical theory of the folding dynamics for an ensemble of random heteropolymers is developed. The theory hinges upon an adiabatic approximation which is validated by the broad separation between significant folding events recorded at the level of contact patterns and comparatively fast dihedral motion of the chain backbone. The results reveal the existence of a metastable folded phase as dynamic equilibrium in accord with experimental findings and a form of relaxation in the same universality class of other systems with quenched disorder. [S1063-651X(97)12807-0]

PACS number(s): 87.15.-v

I. INTRODUCTION

The so-called logarithmic relaxation is found in statistical ensembles of equivalent units with quenched disorder whose dynamics is mapped on a complex free energy landscape [1–4]. By logarithmic relaxation we mean an averaged behavior characterized by a logarithmic growth of the expected encountered activation barrier, which at time t is of the order of $\ln(t/\tau)$, with τ being a characteristic time scale. An illustration of this instance is furnished by a collection of the order of Avogadro's number of polymer molecules of fixed length with a random primary sequence folding intramolecularly under *in vitro* solvent conditions [1–5]. The molecules undergo folding events in an asynchronous manner and only their statistical behavior is amenable to a physical treatment. Their relaxation towards foldings of increasing complexity has been accounted for by implementing a kinetic theory [6] rooted in the random energy model [7]. Instead of introducing models, the aim of this work is to develop a generic statistical theory for the folding of RNA and other frustration-free model biopolymers which rationalizes this mode of relaxation with minimal judicious assumptions.

The description of the RNA folding dynamics resolved at the contact pattern (CP) level entails the underlying assumption that the evolution of the system is determined by a series of transitions between quasiequilibrium states, the CP's themselves. Such an assumption originates in the time scale separation between the different degrees of freedom of the biopolymer chain. Typically, for the various motions of the biopolymer chain a sketch of time scales gives: 10^{-15} s for bond vibrations, 10^{-12} – 10^{-8} s for isolated dihedral torsions, 10^{-9} s for hinge motions, 10^{-6} s for helix-coil transitions and 10^{-4} – 100 s for folding events. Then it can be assumed that there is a vast separation of time scales between folding events and isolated or correlated dihedral torsions.

On the basis of this time scale separation, the faster degrees of freedom can be integrated out in the spirit of an adiabatic approximation: within the biopolymer multidimensional free energy landscape, each state is characterized by a definite CP containing many conformational substates. These arise from the different relative orientations of the backbone

dihedrals within a single specified CP. In this way, each local minima or valley of the free-energy landscape is structured into many subvalleys. Such fine structure could only be revealed by lowering the temperature well below denaturation conditions, as in Frauenfelder's endeavors leading to an elucidation of low T macroscopic kinetic laws of protein relaxation within a class of conformational substates [1,5]. The quasiequilibrium hypothesis holds if local fluctuations representing substate transitions are fast compared with transitions between CP's. In this situation, the barriers for CP transitions are regarded as averages over all substates within the CP's involved. This hypothesis leads to the implementation of an Arrhenius-like kinetic approach to deal with significant folding pathways resolved within the adiabatic approximation. In this approach, the dihedral degrees of freedom are integrated out as conformational entropy which is lost upon a loop closure leading to contact formation [8], while the kinetic barrier corresponding to contact dismantling is taken as the corresponding enthalpy change due to intrachain contact formation [8].

Although the hypothesis formulated above seems reasonable and its implementation in an algorithmic form to elucidate the kinetics of folding of RNA led to results in agreement with experimental findings [8], further work will be necessary to validate it from first principles. At present we are focusing on this problem by incorporating all significant internal backbone motions within a dynamic hierarchy of relaxation time scales, reflecting the time-dependent transference of mutual constraints between folding modes and dihedral motion.

The aim of this paper is to implement a statistical theory of the folding dynamics for an ensemble of random heteropolymers within an adiabatic treatment. In this context, a description at the CP level with quasiequilibrium states separated by Arrhenius-like kinetic barriers of thermodynamic origin holds valid. Consequently, to determine the statistics upon which the dynamics are built, we pick the enthalpy H ($H < 0$) of a folded state as the relevant coordinate, assigning $H = 0$ to the random coil conformation (RC). This choice is appropriate since enthalpy changes result from heat released and transferred to the statistical bath—here the

solvent—due to intramolecular contact formation and, consequently, the enthalpic content of a specific state depends in frustration-free models directly on the CP to which the state is associated. Furthermore, the dynamics at the level of transitions between CP's are understood and have been effectively modeled [8,9]. Thus, our theory aims at defining the statistical dynamics along a single coordinate H as a projection of the dynamics within the CP space for random copolymers in the long chain limit.

II. BASIC TENETS

Our theory is inspired by preexisting models, such as the random energy model [1,7]. However, it is built upon the level distribution of a different thermodynamic variable compatible with the appropriate coarse CP description of conformation space. Within the CP space, the dynamics are determined following a general scheme [8,10]. The kinetic barrier B associated with a contact formation $B=B(\text{loop})$, is entropic in nature since the transition state entails a loop closure with the concurrent loss in conformational freedom $B(\text{loop}) = -T\Delta S(\text{loop})$, where $\Delta S(\text{loop})$ is the entropy loss associated with loop closure, already computed for any size loop [8]. On the other hand, the kinetic barrier associated with contact dismantling $B=B(\text{del})$ is of enthalpic origin, since deletion of an intramolecular contact requires heat absorption in the same amount as that released, ΔH , upon formation of the contact. Thus, we get $B(\text{del}) = -\Delta H$.

The level of coarse graining adopted in our kinetic theory overlooks conformational detail beyond that necessary to specify the thermodynamic quantities indicated above. Therefore, the validity of the adiabatic approximation is tantamount to warranting the validity of our theory.

The description of conformation space resolved at the CP level together with the thermodynamic identification of the kinetic barriers is valid strictly in a frustration-free context since frustration, when present, may be reflected in the free-energy landscape. In our treatment, the enthalpy content can only be varied by means of contact formation or contact disruption events, therefore neglecting frustrating interactions: For instance, the enthalpy may *only* be lowered by increasing the number of intrachain contacts, thus the possibility of eliminating frustration by dismantling intrachain bonds is not contemplated as a means of lowering enthalpy. Thus our results are relevant to the folding of RNA and frustration-free model biopolymers. By contrast, the protein folding context, where frustration is believed to play an important role, falls outside the scope of our theory.

We introduce the density of conformations with enthalpy H , $\Omega(H)/\Omega = F(H)$, where $\Omega(H)$ and Ω are, respectively, the number of conformations compatible with enthalpy H and the total number of conformations. Thus the entropy $S=S(H)$ of a state with enthalpy H is $S(H) = R\ln F(H)$. In order to determine $F(H)$, we make use of the fact that there must exist a denaturation temperature T_c . Thus at $T=T_c$, $G(H) = \Delta G(H)$ is identically zero irrespective of H . The quantity $\Delta G(H)$ is the free-energy change associated with the transition from the RC to the folded state with enthalpy H . Then, the following relations hold:

$$-RT_c \ln F(H) + H = 0, \quad (1)$$

$$F(H) = \exp(-|H|/s), \quad (2)$$

$$\Delta S = (R/s)\Delta H, \quad (3)$$

where $s = RT_c$. Since $RT/s < 1$ for $T < T_c$ and given the nature of the kinetic barriers involved in formation and dismantling of intramolecular contacts, Eq. (3) implies that folding is mostly delayed by dismantling of “misfolded” structure, in accord with pertinent observations [11,12,9,13]. In order to determine the statistical dynamics, we define the probability $P(H,t)$ of enthalpy H at time t , satisfying $P(H,t) = N(H,t)/N$, where an ensemble of copies of the system, each one consisting in an individual polymer molecule is assumed, and $N(H,t)$ and N indicate, respectively, the number of molecules with enthalpy H at time t and the total number of molecules in the ensemble. Thus we define a master equation for P as follows:

$$\begin{aligned} \partial P(H,t)/\partial t = & \Omega(H) \int_{-\infty}^H \exp[(H'-H)/RT] P(H',t) dH' \\ & - P(H,t) \int_{-\infty}^H \exp[(H'-H)/s] \Omega(H') dH' \\ & + \Omega(H) \int_H^0 \exp[(H-H')/s] P(H',t) dH' \\ & - P(H,t) \int_H^0 \exp[(H-H')/RT] \Omega(H') dH', \end{aligned} \quad (4)$$

where $f \approx 10^6 \text{ s}^{-1}$ is the rate constant for contact formation once the nucleating event of loop closure has taken place [8]. The two sources of probability represented by the first and third terms on the right-hand side of Eq. (4) give the positive rate of probability change due to elementary transitions $H' \rightarrow H$, from levels with enthalpy H' below and above H , respectively. The barrier associated with the former transition is $B = H - H'$ ($H' \leq H$), while the latter transition requires surmounting a barrier of entropic origin $B = RT(H' - H)/s$ ($H' \geq H$). The rate contributions follow directly from Eq. (3), the computation of the kinetic barriers for elementary transitions within the space of CP's, and the general form of the unimolecular rate constant r for a folding transition. This rate is computed as $r = f \exp(-B/RT)$, where B is the kinetic barrier involved in the transition. On the other hand, the two sinks of probability, given by the second and fourth terms, correspond to transitions $H \rightarrow H'$. In this situation, whenever $H' \leq H$ (second term) the barrier is entropic and it becomes enthalpic if $H' \geq H$ (fourth term).

Equation (4) may be integrated numerically with the appropriate initial condition $P(H,0) = \delta(H-0)$. This condition holds because folding is assumed to take place when a renaturation temperature $T < T_c$, is reestablished and thus the starting point of the process is the RC with $H=0$. In order to monitor the dynamics, we follow the expected enthalpy $\langle H(t) \rangle$ at time t :

$$\langle H(t) \rangle = \int_{-\infty}^0 H P(H,t) dH. \quad (5)$$

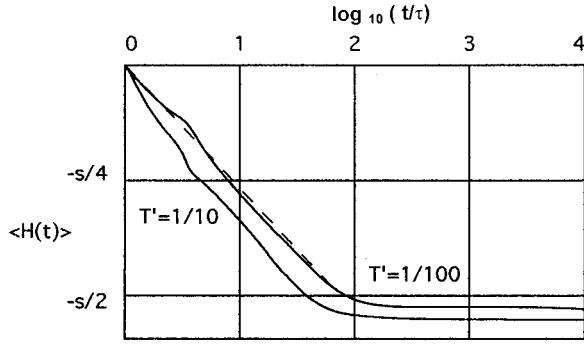


FIG. 1. The time-dependent behavior of the expected enthalpy $\langle H(t) \rangle$ obtained by numerical integration of the master equation. The ideal logarithmic relaxation is indicated by the dashed lines. Abscissas are given in logarithmic form using a scaling constant $\tau = 1 \mu\text{s}$. The reduced temperatures were fixed at 0.1 and 0.01.

III. RESULTS AND DISCUSSION

The results are displayed in Fig. 1 for specific reduced temperatures T' . The following notation has been adopted: $\tau = f^{-1}$; $T' = (T_c - T)/T_c$ ($T \leq T_c$). The critical temperature has been fixed at $T_c = 318$ K. The logarithmic time-dependent behavior of $\langle H(t) \rangle$ fits into the physical picture of general relaxation dynamics for glassy disordered materials [1,7], thus corroborating the validity of the approach presented in this work. This relaxation regime is invariably followed by a sudden asymptotic relaxation to a saturation enthalpy value $H = H_\infty$. Within this new regime, $\langle H(t) \rangle$ remains almost constant, satisfying $|\langle H(t) \rangle - H_\infty| < 10^{-5} H_\infty$.

The saturation enthalpy may be easily determined: For a given enthalpy H we may determine the ratio $Y(H) = r_\downarrow(H)/r_\uparrow(H)$, where $r_\downarrow(H)$ is the rate of downwards transition in the enthalpy spectrum with starting point H , and r_\uparrow is the rate of upward transition:

$$Y(H) = \frac{\int_{-\infty}^H \exp[(H' - H)/s] \Omega(H') dH'}{\int_H^0 \exp[(H - H')/RT] \Omega(H') dH'} \quad (6)$$

In general $Y(H) > 1$ if $H > H_\infty$, that is, starting at the RC ($H = 0$), there is, on average, a tendency to increase the folding complexity by forming contacts until the saturation enthalpy $H = H_\infty$ is reached. The saturation enthalpy is defined as satisfying the equation $Y(H) = 1$. This gives

$$H_\infty = \frac{\ln[(3RT - s)/2RT]}{(1/RT - 1/s)} \quad (7)$$

This behavior is illustrated in Fig. 2 for a temperature range of biological interest. Since we get

$$\lim_{T \rightarrow T_c} H_\infty = -\frac{s}{2}, \quad (8)$$

while $H_\infty = 0$ for $T > T_c$, we obtain at T_c a first-order phase transition with latent heat $s/2$, in qualitative agreement with experimental findings rooted in calorimetric measurements of denaturation [14].

This phase transition is characterized by

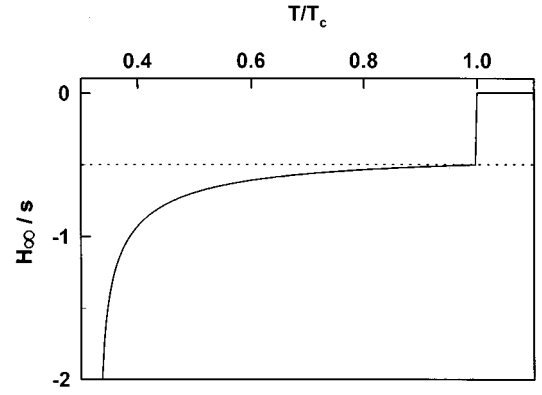


FIG. 2. The saturation enthalpy H_∞ as defined by Eq. (7) plotted as a function of reduced temperature T/T_c . The dashed line indicates the asymptotic limit $s/2$.

$$\frac{H_\infty(T) - H_\infty(T_c)}{H_\infty(T_c)} = \left| \frac{T - T_c}{T_c} \right|^\alpha, \quad (9)$$

where $\alpha = 1$. The heat capacity at constant pressure C_p is given explicitly by the following expression:

$$C_{p^\infty} = \frac{RT_c^2 \ln\left(\frac{3T - T_c}{2T}\right)}{(T - T_c)^2} - \frac{RT_c^2}{(3T - T_c)(T - T_c)}, \quad (10)$$

then

$$\lim_{T \rightarrow T_c} C_{p^\infty} = \frac{R}{8}. \quad (11)$$

The time-dependent behavior of the expected enthalpy displayed in Fig. 1 reveals the existence of a metastable folded phase emerging as a dynamic equilibrium. This is so since the minimum free energy realized is $G = H_\infty - (RT/s)H_\infty$. On the other hand, since $G = (1 - RT/s)H$, the free energy could, in principle, decrease boundlessly in the limit of long chains considered. Therefore, we may conclude that the metastable phase becomes dominant as a dynamic equilibrium in the range $273 \text{ K} < T < T_c$. Actually, no analysis is required below the freezing point of the solvent (273 K), since the folding process cannot take place.

Figure 2 makes apparent the robustness of the metastable phase since H_∞ approaches asymptotically its limit value $s/2$ in a very smooth way as T approaches T_c within the temperature range of biological interest. The ensemble average enthalpy of the destination state is relatively invariant over a broad range of temperatures below and within the neighborhood of T_c . This suggests a highly robust metastable folded phase in the sense that the ensemble average of the degree of folding, as measured by $\langle H(t) \rangle$ in a frustration-free context, is resistant to temperature change.

IV. CONCLUDING REMARKS

The relaxation dynamics below criticality predicted by our statistical theory reproduces the known relaxation behavior of disordered glassy materials. Our results also predict the existence of a dynamically dominant metastable folded phase

which undergoes a first-order transition with latent heat at the critical temperature, in accord with calorimetric experiments and with mounting evidence pointing towards the need for an alternative predictive tool radically different from free-energy minimization algorithms.

A description of the folding dynamics of random heteropolymers within a coarse-grained conformation space resolved at the CP level hinges upon the assumption that the system evolves by means of transitions between quasiequilibrium states (adiabatic approximation). The validity of this ansatz is sustained by the broad time scale separation between relevant folding events and the fast-evolving dihedral motions of the chain which can therefore be integrated out as entropy. The thermodynamic nature with which we endow the kinetic barriers associated to folding events arises from

this assumption. Therefore, the adiabatic ansatz determines the context of applicability of our kinetic theory for the folding problem. Although the results obtained in this and previous works [8,10] are consistent with experimental findings, further work will be necessary to state the validity of the broadly used adiabatic ansatz within the folding context.

ACKNOWLEDGMENTS

A.F. is principal investigator of CONICET, the National Research Council of Argentina. Financial support from the John Simon Guggenheim Memorial Foundation (New York City) to A.F. is gratefully acknowledged. G.A.A. thanks CONICET for financial support.

-
- [1] A. Fernández and E. I. Shakhnovich, *Phys. Rev. A* **42**, 3657 (1990).
- [2] J. D. Bryngelson and P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **84**, 7524 (1987).
- [3] A. Fernández, *Ber. Bunsenges. Phys. Chem.* **98**, 260 (1994).
- [4] A. Fernández, *Phys. Rev. Lett.* **64**, 2328 (1990).
- [5] S. Sligar, H. Frauenfelder, and P. G. Wolynes, *Science* **254**, 1598 (1991).
- [6] A. Fernández, *Physica A* **201**, 557 (1993).
- [7] B. Derrida, *Phys. Rev. B* **24**, 2613 (1981).
- [8] A. Fernández, H. Arias, and D. Guerin, *Phys. Rev. E* **52**, R1299 (1995).
- [9] P. Zarrinkar and J. Williamson, *Science* **265**, 918 (1994).
- [10] A. Fernández, *Ann. Phys. (Leipzig)* **4**, 600 (1995).
- [11] R. Jaenicke, *Angew. Chem. Int. Ed. Engl.* **23**, 295 (1984).
- [12] T. Creighton, *J. Phys. Chem.* **89**, 2452 (1985).
- [13] E. I. Shakhnovich, G. Farztdinov, A. Gutin, and M. Karplus, *Phys. Rev. Lett.* **67**, 1665 (1991).
- [14] P. L. Privalov, *Protein Structure and Protein Engineering*, edited by E. L. Winnacker and R. Hubert (Springer, Berlin, 1988), pp. 6–15.