The adiabatic ansatz in biopolymer folding dynamics

Ariel Fernández* and Alejandro Belinky**

(*) Instituto de Matemática, Universidad Nacional del Sur, Consejo Nacional de Investigaciones Científicas y Técnicas, Avenida Alem 1253, Bahía Blanca 8000;

(**) Departamento de Economía, Universidad Nacional del Sur, Avenida Alem 1253, Bahía Blanca 8000, Argentina.

Biopolymer folding is an expeditious process taking place within timescales incommensurably shorter than ergodic times. We demonstrate that this property may be accounted for by means of a variational principle which is formulated within an adiabatic approximation obtained by integrating out fast-relaxing molecular motions. The least-action relaxation is shown to yield the glassy behavior which has been previously encountered in the folding of random heteropolymers.

1. Introduction

The folding of natural biopolymers (RNA, proteins) under *in vitro* solvent conditions is an expeditious, efficient and reproducible process which represents the search in conformation space performed by a biopolymer chain that forms intramolecular contacts at the expense of losing conformational freedom. The difficulty in finding theoretical underpinnings of such phenomena is due to the fact that folding is neither an energetically-downhill process nor the result of an exhaustive random exploration of conformation possibilities [1-4], the extreme cases which would make the problem far more tractable.

Such a framework calls for an action principle [5] which should single out folding pathways resolved within a description of conformation space compatible with the means of detection [2,3,6]. In fact, a specialization of this approach to the folding of natural RNA molecules [4] has been relatively successful at reflecting the essential properties of the process while reproducing the experimentally-probed folding pathways [4,6]. The means of detection engaged in probing pathways are indeed compatible with the coarse graining of conformation space we have dealt with theoretically: The detection of folding events is essentially rooted in RNA-DNA hybridization techniques which can only resolve structure at the level of base pairing contact patterns (CP's). Within this level of resolution we intend to focus first on the following problem:

Obviously, to assume that the exploration of conformation space is dictated by a sequence of CP transitions governed by an action principle [5] implies first of all that an adiabatic approximation to folding dynamics holds in this context. Thus, in the spirit of the Born-Oppenheimer approximation in molecular physics, we have treated the folding problem adiabatically, assuming the existence of slow enslaving folding degrees of freedom: If we rely on known estimates of mean relaxation parameters [7-9], there appears to exist a separation of timescales between folding events resolved as CP transitions (10^{-4}-10^{3}s) [1-14] on one hand, and relaxation timescales for torsional dihedral motions (10^{-12}-10^{-7}s), puckering (10^{-10}-10^{-7}s), rotation about glycosidic bonds (10^{-11}s) and vibrations (10^{-15}s) on the other hand. However, there are caveats when handling this information: Spectroscopic measurements of timescales have been performed mostly for individual nucleotides (the RNA monomeric units), and often under *in vacuo* conditions.

Is the adiabatic assumption a valid one? If so, then fast microscopic motion could be integrated out as entropy, each CP would represent indeed a state of quasiequilibrium, and forward and backward activation barriers associated to CP transitions could be

evaluated respectively taking into account either the conformational entropy loss or the enthalpy loss associated to intrachain contact formation [1]. This is indeed the picture we have adopted so far for natural RNA sequences, and it appears to reproduce the phenomenology within the coarse CP description compatible with the sensitivity of the observations [6].

From a physical perspective, it would be desirable to formulate the action principle explicitly, in a fashion that makes no reference to evolutionary facts. *Thus, we shall concentrate on reproducing the glassy logarithmic relaxation of random heteropolymers [10] by means of a least action principle formulated within an adiabatic approximation.*

The so-called logarithmic relaxation is found in statistical ensembles of equivalent entities with quenched disorder whose dynamics is mapped on a complex free energy landscape [10-12]. 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 timescale. An illustration of this instance is furnished by a collection of the order of Avogadro's number ($\sim 10^{23}$) of polymer molecules of fixed length with a random primary sequence folding intramolecularly under *in vitro* solvent conditions [10-13]. The molecules undergo folding events in an asynchronous manner and only their statistical behavior is amenable of a physical treatment. Their relaxation towards foldings of increasing complexity has been accounted for by implementing a kinetic theory [14] rooted in the random energy model (REM) [10,15]. Our aim in this work is to cast this behavior in terms of a variational principle.

Focusing on relaxation processes with steps of increasing difficulty, it becomes intuitively obvious that the fastest process would correspond to a logarithmic growth of activation barriers, since such a growth is the slowest possible. Then, the following

question arises: *Is it possible to derive a least action principle holding for processes of increasing difficulty such that the glassy logarithmic relaxation described above is singled out as a brachistochrone or least over-all time relaxation pathway ?*

In order to answer the question posed, we shall first develop a formal adiabatic scheme with minimal detail that effectively reproduces the statistics of the folding dynamics of an ensemble of random heteropolymers. Then, we shall prove that the expected relaxation leading to increasing structural complexity obeys a variational principle which encompasses a broad range of phenomena involving systems with quenched disorder.

2. Basic tenets of the folding statistics

To determine the statistics upon which the dynamics are built within an adiabatic approximation, we pick the enthalpy H ($H \leq 0$) of a folded state as the relevant coordinate, assigning $H=0$ to the random coil (RC) conformation. This choice is appropriate since enthalpy changes result from heat released and transferred to the statistical bath (the solvent) due to intramolecular contact formation and, consequently, the enthalpic content of a specific state depends directly on the CP to which the state is associated. Thus, our theory aims at defining the statistical dynamics along a single coordinate, H, as a projection of the adiabatic dynamics within the CP space for random copolymers in the long chain limit.

Within the adiabatic approximation, the dynamics are determined following a general scheme [1,4]: The kinetic barrier B associated to a contact formation, B=B(loop), is entropic in nature since the transition state entails a loop closure with the concurrent loss in conformational freedom [1]: $B(\text{loop}) \approx -T\Delta S = -T\Delta S(\text{loop})$, where $\Delta S(\text{loop})$ is the entropy loss associated to loop closure, already computed for any size loop [5]. On the other hand, the kinetic barrier associated to contact dismantling, B=B(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: $B(del.) \approx -\Delta H$.

We introduce the density of microscopic realizations 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, within the adiabatic approximation 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 exists within the adiabatic approximation 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 to 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 the folding is mostly delayed due to dismantling of "misfolded" structure, in accord with current observations [2-6,16]. Eq. 3 yields a straightforward relation to determine the unimolecular rate constant k for intra-chain pairing. Thus, the generic Arrhenius expression [1] valid within the adiabatic approximation: $k=\mu \exp(-B/RT)$, with μ=effective collision frequency, becomes proportional to $B\exp(-B/RT)$. This is so since $\mu=10^5 s^{-1} n$, with n=number of contacts formed once the loop closure has taken place [1,6], and $-\Delta H=sB/RT$ is proportional to n.

Within the adiabatic approximation described above, and since we are dealing with ensemble-averaged relaxation, we shall consider a generic situation under the following

constraints which apply in particular to the statistics of the folding dynamics of random heteropolymers: a) The free energy landscape is rugged because of the high degeneracy of enthalpy levels; b) the expected barrier B(t) encountered at time t grows monotonically with t, as is the case for a system of increasing complexity in which successive relaxation steps become increasingly difficult; and c) the progress of relaxation may be monitored by a single-valued function H(t)=H(B(t)).

3. The variational formulation

Under the tenets expounded above, and given the simple form of the unimolecular rate constant within the adiabatic approximation, it becomes convenient to define a path integral functional G{H(B)} giving a magnitude proportional to the overall relaxation time $\int dt$:

$$G\{H(B)\} = \int B^{-1}\exp(B/RT)\,(1+H'^2)^{1/2}\,dB, \qquad (4)$$

where $H'=dH/dB$ and $(1+H'^2)^{1/2}dB$ is the arc differential in the plot H=H(B) with integration extremes B=0 and B=B. Making use of canonical tools of calculus of variations we obtain the following set of coupled ordinary differential equations:

$$dB/dt = [B^2\exp(-2B/RT) - B^4\exp(-4B/RT)]^{1/2} \qquad (5)$$

$$dH/dt = [B^2\exp(-2B/RT) - (dB/dt)^2]^{1/2} \qquad (6)$$

The numerical integration of eqs. 5 and 6 has been performed elsewhere [17] and need not be repeated here. It reveals a logarithmic relaxation. The logarithmic time-dependent behavior of H(t) fits into the physical picture of general relaxation dynamics for glassy disordered materials [10,11], thus corroborating the validity of the approach

presented in this work. This relaxation regime is followed by a sudden asymptotic relaxation to a saturation enthalpy value H=H(∞). Thus, the time-dependent behavior of the expected enthalpy marks the existence of a metastable folded phase emerging as a dynamic equilibrium. This is so since the minimum free energy realized is G=H(∞)-(RT/s)H(∞). This is a local and not the global minimum, since being 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 within an adiabatic approximation, the metastable phase becomes dominant as a dynamic equilibrium in accord with previous findings [1,14]. The asymptotic relaxation to a finite H(∞) merits discussion: Eq. 3 reveals that at renaturation conditions, the activation barrier associated to a decrease in enthalpy is actually smaller than the barrier for the backwards transition. However, the tendency to decrease in enthalpy is curbed by an opposite trend due to the progressive scarcity of enthalpy levels, as indicated by Eq. 2. The occurrence of a saturation enthalpy is therefore a consequence of two opposing effects, one of enthalpic and the other of entropic origin.

The growth in time of the encountered activation barrier resulting from numerical integration of eqs. 5 and 6 reveals again a signature of glassy relaxation while casting the result in a new light, as it leads to an interpretation of the logarithmic growth: Since such growth is actually the slowest possible growth, it must yield the brachistochrone relaxation pathway within a process of increasing level of difficulty, as evidenced by the minimization of the functional presented in Eq. 4.

The variational formulation described by Eq. 4 not only validates the adiabatic statistical theory of folding presented in this work, but also reveals that the logarithmic law which governs the glassy relaxation of several materials with quenched disorder actually represents the fastest relaxation amongst pathways which entail a monotonic increase with time in the size of the expected activation barrier. Because of the generic

nature of these results, our endeavors are now directed at assessing whether the variational formulation holds beyond the context of folding of disordered copolymers.

This work was financially supported by the J. S. Guggenheim Memorial Foundation of New York City through a fellowship awarded to A.F.. Part of this research was carried out during a visit to the Courant Institute of New York University. We thank Prof. J. K. Percus for his kind hospitality and insightful discussions.

References

[1] A. Fernández, *Ann. Physik* **4**, 600 (1995)
[2] R. Jaenicke, *Angew. Chem. Intl. Ed. Engl.* **23**, 295 (1984)
[3] T. Creighton, *J. Phys. Chem.* **89**, 2452 (1985); T. Creighton, in "Protein Folding", L. M. Gierasch and J. King, eds., pages 157-170, Am. Assoc. Adv. Sci. Publishers, Washington (1990)
[4] A. Fernández, H. Arias and D. Guerín, *Phys. Rev. E* **52**, R1299 (1995)
[5] A. Fernández and H. Cendra, *Biophysical Chemistry* **58**, 335 (1996)
[6] P. Zarrinkar and J. Williamson, *Science* **265**, 918 (1994)
[7] C. Cantor and P. Schimmel, "Biophysical Chemistry", Freeman and Co., New York (1980)
[8] W. Saenger, "Principles of Nucleic Acid Structure", Springer, New York (1984)
[9] C. Brooks, M. Karplus and M. Pettitt, "Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics", Wiley, New York (1988)
[10] A. Fernández and E. I. Shakhnovich, *Phys. Rev. A* **42**, 3657 (1990)
[11] J. D. Bryngelson and P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **84**, 7524 (1987); A. Fernández, *Ber. Bunsenges. Phys. Chem. - Rapid Comm.* **98**, 260 (1994)
[12] A. Fernández, *Phys. Rev. Lett.* **64**, 2328 (1990)
[13] H. Frauenfelder, S. Sligar and P. G. Wolynes, *Science* **254**, 1598 (1991)
[14] A. Fernández, *Physica A* **201**, 557 (1993)
[15] B. Derrida, *Phys. Rev. B* **24**, 2613 (1981)
[16] E. I. Shakhnovich, G. Farztdinov, A. Gutin and M. Karplus, *Phys. Rev. Lett.* **67**, 1665 (1991)
[17] A. Fernández and G. Appignanesi, *Phys. Rev. Lett.*, in press (1997).