

## Statistical folding dynamics for random heteropolymers

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1996 J. Phys. A: Math. Gen. 29 L523

(<http://iopscience.iop.org/0305-4470/29/20/004>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.70

The article was downloaded on 02/06/2010 at 04:02

Please note that [terms and conditions apply](#).

## LETTER TO THE EDITOR

# Statistical folding dynamics for random heteropolymers

Ariel Fernández

Instituto de Matemática—INMABB, Universidad Nacional del Sur, Consejo Nacional de Investigaciones Científicas y Técnicas, Avenida Alem 1253, Bahía Blanca 8000, Argentina

and

The Frick Laboratory, Princeton University, Princeton, NJ 08544, USA

Received 3 July 1996

**Abstract.** The existence of a denaturation temperature,  $T_c$ , in the folding of random heteropolymers allows us to determine the distribution of enthalpy levels within the space of contact patterns. The resulting statistics yield a metastable ensemble of foldings dynamically dominant below  $T_c$ , and reproduce the relaxation dynamics of a disordered glassy material, in accord with previous findings.

The biopolymer folding problem is central to molecular biophysics: the search for the active or functional conformation performed by a biopolymer chain that forms intramolecular interactions under *in vitro* solvent conditions is neither a downhill process nor the result of a random exploration in conformation space [1–4]. Functional conformations are obtained within a short biologically-relevant timeframe generally *incompatible* with thermodynamic timescales [5–7]. Moreover, an accumulation of counterexamples challenges the notion that global free-energy minimization criteria might yield a generic predictive algorithm to determine the active conformation [1–7].

Thus, a statistical theory of folding dynamics must at least warrant the existence of a metastable folded phase regarded as a sharply concentrated ensemble of structures which are the destinations of the manifold pathways realized in physically relevant timescales [1–3]. Furthermore, such a theory must account for kinetic control in the folding process [1–7], while making the kinetics compatible with the denaturation–renaturation thermodynamics. This implies that the dynamically dominant ensemble of folded states of the system must actually undergo a phase transition, coexisting with the random coil (RC) at the denaturation critical temperature  $T_c$ .

Within this context, the study of random heteropolymers is justified from a physicist's perspective, as the folding process may be cast in this case as the relaxation of a disordered glassy material [8, 9]. Furthermore, the choice of a random sequence enables us to examine the worst-performance limit case in the folding process and separate the properties inherent to the process itself from those arising from natural selection.

To determine the statistics upon which the dynamics are built, we pick the enthalpy  $H$  ( $H \leq 0$ ) of a folded state as the relevant coordinate, assigning  $H = 0$  to the RC. 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 contact pattern (CP) to which the state is associated. Furthermore, the dynamics at the level of transitions between CPs

are understood and have been effectively modelled [3,5]. 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.

Our theory is inspired by pre-existing models, such as the random energy model (REM) [7, 10]. 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 [3, 6]: the kinetic barrier  $B$  associated to 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 [5]:  $B(\text{loop}) \approx -T\Delta S(\text{loop})$ , where  $\Delta S(\text{loop})$  is the entropy loss associated with loop closure, already computed for any size loop [5]. 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,  $\Delta H$ , upon formation of the contact. Thus, we get  $B(\text{del}) \approx -\Delta H$ .

We introduce the density of conformations with enthalpy  $H$ :  $\Omega(H)/\Omega = F(H)$ , where  $\Omega(H)$  and  $\Omega$  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, equation (3) implies that the folding is mostly delayed due to dismantling of 'misfolded' structure, in agreement with current observations [1, 2, 4, 6].

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 consisting of 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:

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

The two sources of probability represented by the first and third terms in the right-hand side of equation (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 equation (3), the computation of the kinetic barriers for elementary transitions within the space of CPs 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  $f \approx 10^6 \text{ s}^{-1}$  is the rate constant for contact formation once the nucleating event of loop closure has taken place [5] and  $B$  is the kinetic barrier involved in the transition. On the other hand, the two sinks of probability, given by the second and fourth terms, corresponds 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).

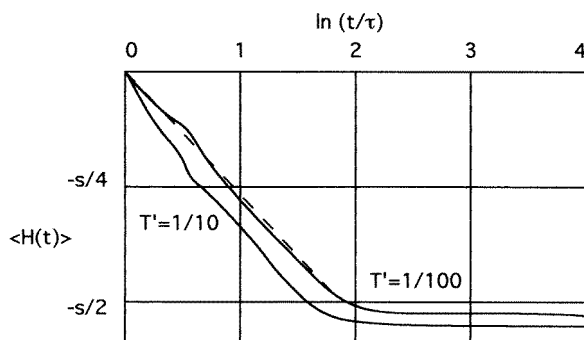
Equation (4) may be integrated numerically with the appropriate initial condition:  $P(H, 0) = \delta(H - 0)$ . This condition holds because folding is assumed to take place when a renaturation temperature  $T < T_c$  is re-established and thus, the starting point of the process is the RC with  $H =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

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

The results are displayed in figure 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  $T_c = 318 \text{ }^\circ\text{K}$ . The logarithmic time-dependent behaviour of  $\langle H(t) \rangle$  fits into the physical picture of general relaxation dynamics for glassy disordered materials [7, 10], 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  $|\langle H(t) \rangle - H(\infty)| \leq 10^{-5} H(\infty)$ .

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  $r \uparrow(H)$  is the rate of upwards transition:

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



**Figure 1.** Time-dependent behaviour of the expected enthalpy  $\langle H(t) \rangle$ , as obtained by numerical integration of equation (4). The broken line indicates ideal logarithmic relaxation. The abscissas are dimensionless and given in logarithmic form with scaling constant  $\tau = 1 \mu\text{s}$ . The ordinates are given in units of  $s = RT_c$ . Two plots are given, corresponding to two choices of the reduced temperatures:  $T' = 1/10$  and  $T' = 1/100$ .

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

$$H(\infty) = \ln[(3RT - s)/2RT]/(1/RT - 1/s). \quad (7)$$

Since we get

$$\lim_{T \rightarrow T_c} H(\infty) = -s/2 \quad \text{while } H(\infty) = 0 \text{ if } T > T_c \quad (8)$$

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 [11].

The time-dependent behaviour of the expected enthalpy displayed in figure 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 may 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  $273 \text{ }^\circ\text{K} < T < T_c$ . Actually, no analysis is required below the freezing point of the solvent ( $273 \text{ }^\circ\text{K}$ ), since the folding process cannot take place.

To summarize, our model is justified and corroborated on the following grounds:

(1) The relaxation dynamics below criticality reproduce the known relaxation behaviour of disordered glassy materials.

(2) There exists a dynamically dominant *metastable* folded phase which undergoes a first-order phase 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.

This work was financially supported by the J S Guggenheim Memorial Foundation of New York City through a fellowship awarded to the author.

## References

- [1] Jaenicke R 1984 *Angew. Chem. Intl. Ed. Engl.* **23** 295
- [2] Creighton T 1985 *J. Phys. Chem.* **89** 2452; 1990 *Protein Folding* ed L M Gierasch and J King (Washington: Am. Assoc. Adv. Sci.) pp 157–70
- [3] Fernández A 1995 *Ann. Physik* **4** 600
- [4] Zarrinkar P and Williamson J 1994 *Science* **265** 918
- [5] Fernández A, Arias H and Guerin D 1995 *Phys. Rev. E* **52** R1299
- [6] Shakhnovich E I, Farztdinov G, Gutin A and Karplus M 1991 *Phys. Rev. Lett.* **67** 1665
- [7] Fernández A and Shakhnovich E I 1990 *Phys. Rev. A* **42** 3657
- [8] Fernández A 1990 *Phys. Rev. Lett.* **64** 2328
- [9] Fernández A 1994 *Ber. Bunsenges. Phys. Chem.–Rapid Commun.* **98** 260
- [10] Derrida B 1981 *Phys. Rev. B* **24** 2613
- [11] Privalov P L 1988 *Protein Structure and Protein Engineering* ed E-L Winnacker and R Huber (Berlin: Springer) pp 6–15