Phases of the secondary structures of RNA sequences

R. BUNDSCHUH[1,2] and T. HWA[1]

[1] *Department of Physics, U.C. San Diego, La Jolla, CA 92093-0319, U.S.A.*
[2] *Physics Department, The Ohio State University, Columbus, OH 43210-1106, U.S.A.*

PACS. 87.15.Aa – Biomolecules: Theory and modeling; computer simulation.
PACS. 05.40.-a – Fluctuation phenomena, random processes, noise, and Brownian motion.
PACS. 87.15.Cc – Biomolecules: Folding and sequence analysis.

Abstract. – Formation of RNA secondary structures is an example of the sequence-structure problem omnipresent in molecular biophysics. A basic theoretical issue concerns the phase behaviour of self-attracting random RNA sequences. By studying a simplified toy model of RNA folding, we show that there are two distinct possible phases for the random RNA below its melting transition — a *molten phase* in which an exponentially large number of allowed secondary structures have comparable free energies and coexist in thermal equilibrium, and a *glass phase* in which the equilibrium ensemble is dominated by one or a few structures with much lower free energies.

Introduction: RNA is an important biopolymer critical to all living systems [1]. Like the DNA, there are four types of nucleotides (or bases) A, C, G, and U which, when polymerized can form double helical structures consisting of stacks of stable Watson-Crick (A–U or G–C) pairs. However unlike a long polymer of DNA, which is often accompanied by a complementary strand and forms otherwise featureless double helical structures, a polymer of RNA is usually single-stranded. It bends onto itself and forms elaborate spatial structures in order for bases located on different parts of the backbone to pair with each other. The structures encoded by the primary sequences often have important biological functions, much like the structures of proteins encoded by their amino acid sequences.

Understanding the encoding of structure from the primary sequence has been an outstanding problem of theoretical biophysics. Most work in the past decade has been focused on the problem of protein folding, which is very difficult analytically and numerically [2]. Here, we study the problem of RNA folding, specifically the formation of RNA *secondary structures* which is more amenable to analytical and numerical approaches due to a separation of energy scales [3]. Efficient algorithms [4,5] together with carefully measured free energy parameters [6] describing the formation of various microscopic structures (*e.g.*, stacks, loops, hairpins, etc.) allow the exact calculation of the ensemble of secondary structures formed by a given RNA molecule of up to a few thousand bases.

In this work, we are not concerned with the structure formed by a specific sequence. Instead, we will study the statistics of secondary structures formed by the ensemble of *long random* RNA sequences. At high enough temperatures, the entropy gain of a freely fluctuating polymer backbone outweighs the binding energy to be gained by any base pairing and

the RNA is in its denatured phase. As the temperature is lowered, base pairs are formed and compact structures appear. The transition between the denatured phase and compact phases is a crossover if self-avoidance of the backbone is neglected [7] but turns into a bona fide phase transition in the presence of self-avoidance [8]. Within the regime of compact structures below the denaturation temperature, it has been debated whether an ensemble of long random RNA sequences allows for only one or for two distinct thermodynamical phases [9–11]. However, the numerical results these studies are based on are not clear enough to allow unambiguous interpretation. In this letter, we provide *analytical evidence* supporting the existence of two distinct compact phases as illustrated in fig. 1, by studying some toy models of RNA folding. We characterise the behaviour of RNA in the absence of disorder (or equivalently at moderately high temperatures) and show through a two-replica calculation that disorder is perturbatively irrelevant. That implies that RNA is in a "molten phase (analogous to the "molten globule" of proteins), characterized by the coexistence of many structures which are nearly degenerate energetically. We then show that it is inconsistent to have the molten phase persist to arbitrarily strong disorder (or low temperature), leading to the notion of a distinct strong disorder phase. This strong disorder (or low temperature) phase is characterized by a small number of distinct low-energy structures, and will be referred to as the "glass phase" in analogy with studies of other disordered systems.

Fig. 1 – Putative phase diagram: The RNA is denatured above a denaturation temperature T_d. Below T_d, there is a molten phase at intermediate temperature and a glass phase at low temperature.

Model: A secondary structure S of a polymer of RNA is the set of pairings formed between all of its monomers (or bases), with each base allowed in at most one pairing. We denote the pairing between the i^{th} and j^{th} monomer by (i,j), with $1 \leq i < j \leq N$. Each such structure can be represented by a diagram like the one shown in fig. 2(a). In addition, it is common to exclude "pseudo-knots" like the one shown in fig. 2(b) from the definition of secondary structures, so that any two base pairs (i,j) and (k,l) are either independent, *i.e.*, $i<j<k<l$, or nested, *i.e.*, $i<k<l<j$. This is permissible since long pseudo-knots are kinetically difficult to form and even the short ones occur infrequently due to energetic reasons [3]. Experimentally, it is possible to "turn off" the pseudo-knots and other complicated tertiary contacts [3] and study exclusively the class of secondary structures defined above.

In order to calculate Boltzmann factors for an ensemble of secondary structures, we need to assign an energy $E[S]$ to each structure S. For the purpose of secondary structure *prediction*, it is essential to model the energy as accurately as possible [4, 5]. We are instead interested

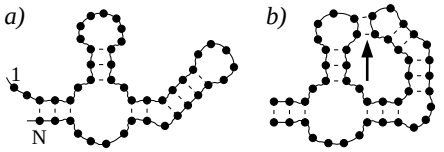

Fig. 2 – Secondary structures of an RNA: The solid and dashed lines represent the backbone and base pairings respectively. *a)* shows a valid secondary structure while *b)* contains a pseudo-knot as indicated by the arrow.

in *universal properties* of long, random RNA sequences. Since accurate energy models for the RNA are much too complex for the quantitative analysis of universal properties, we resort to simplified models used in earlier studies [9,10,12,13]. Only after the possible phases of these simplified models are characterised can we address the phase behavior of RNA with realistic energy parameters at physiological conditions. We will defer the latter to future studies.

We associate an interaction energy $\varepsilon_{i,j}$ with every pairing (i,j) and assign $E[S]=\sum_{(i,j)\in S}\varepsilon_{i,j}$ as the total energy of the structure S. To retain the spirit of Watson-Crick pairing, we choose random sequences $b_1 \ldots b_N$ of the four bases A, U, C, and G and then assign

$$\varepsilon_{i,j} = \begin{cases} -u_m & (b_i, b_j) \text{ is a Watson-Crick base pair} \\ u_{mm} & \text{otherwise} \end{cases} \tag{1}$$

with $u_m, u_{mm} > 0$. Here, u_m represents the typical energy of base-pair matching. It can range from close to zero right below the denaturation transition, to $2 \sim 4\ k_BT$'s under physiological conditions. The value of u_{mm} is not essential as long as it is repulsive, since there is always the option to not bind at all (with energy "0"). In our study, we will use the value of u_m as a tuning parameter to control the strength of "sequence disorder".

To study the weak disorder behaviour, it will be convenient to take the $\varepsilon_{i,j}$'s to be *independent Gaussian* random variables specified by the mean $\overline{\varepsilon}$ and variance

$$\overline{(\varepsilon_{i,j} - \overline{\varepsilon})(\varepsilon_{k,l} - \overline{\varepsilon})} = D\,\delta_{i,k}\delta_{j,l}. \tag{2}$$

Throughout the text, we will use the overline to denote averages over the ensemble of random pairing energies. Here, the parameter D serves as the measure of the strength of the randomness. With $\overline{\varepsilon} = -\frac{1}{4}u_m + \frac{3}{4}u_{mm}$ and $D = \frac{1}{4}u_m^2 + \frac{3}{4}u_{mm}^2 - \overline{\varepsilon}^2$, it is an approximation to the model (1) in two respects: First, it replaces the discrete distribution of energies by a Gaussian distribution. Moreover, it neglects the *correlations* between $\varepsilon_{i,j}$ and $\varepsilon_{i,k}$ induced by the shared base b_i. A similar approximation is commonly used in the context of protein folding [14]; it has also been investigated extensively in the context of the related sequence alignment problem [15]. As in ref. [15], we do not anticipate universal quantities to depend on the subtle differences in the statistics of the $\varepsilon_{i,j}$'s. This will be tested numerically by comparing the scaling behaviour produced by the two models and an intermediate *discrete disorder* model in which the $\varepsilon_{i,j}$'s are chosen independently but with the discrete values

$$\varepsilon_{i,j} = \begin{cases} -u_m & \text{with probability } p = 1/4 \\ u_{mm} & \text{with probability } 1-p = 3/4 \end{cases} \tag{3}$$

Given the energy of each secondary structure, we can study the partition function $Z(N) = \sum_{S\in\mathcal{S}(N)} \exp\{-\beta E[S]\}$ where $\mathcal{S}(N)$ comprises all valid secondary structures of a molecule of length N. This function can be conveniently computed in terms of the *restricted* partition function $Z_{i,j}$ for the substrand $b_i \ldots b_j$. For the simple energy model $E[S]$ described above, $Z_{i,j}$ can be split up according to the possible pairings of position j to yield a recursion

$$Z_{i,j} = Z_{i,j-1} + \sum_{k=i}^{j-1} Z_{i,k-1} \cdot \exp(-\beta\varepsilon_{k,j}) \cdot Z_{k+1,j-1} \tag{4}$$

which is exact [7,9,16]. This equation can then be iterated to compute the full partition function $Z(N) = Z_{1,N}$ for *arbitrary* interaction energies $\varepsilon_{i,j}$'s in $O(N^3)$ time [4,16]. It also forms the basis of analytical approaches to the problem.

The molten phase: If sequence disorder does not play an important role, we may describe the RNA molecule by replacing all the binding energies $\varepsilon_{i,j}$ by some effective value ε_0. As we

will see later, this is an adequate description of random RNA at sufficiently weak disorder. We first briefly review the properties of RNA in this molten phase. Since the $Z_{i,j}$'s become translationally invariant in the absence of disorder, it is straightforward to solve eq. (4) for the molten phase partition function $Z_0(N)$ using the z-transform, yielding [7, 12, 16]

$$Z_0(N) \overset{N\gg 1}{\approx} N^{-\theta_0} \exp[-\beta N f_0(q)] \tag{5}$$

for large N where $q \equiv \exp(-\beta\varepsilon_0)$ is the only parameter, $\theta_0 = 3/2$ is a universal exponent, and $f_0(q) = -k_B T \ln(1+2\sqrt{q})$ is the free energy per length. A useful observable characterising the state of the RNA is the free energy cost $\Delta F(N)$ of *dividing* the polymer into two non-interacting halfs. In the entropy dominated molten phase it simply reflects the loss of *configurational entropy* of the secondary structures due to the division and is given by

$$\Delta F = -k_B T \ln[Z_0^2(N/2)/Z_0(N)] \approx \tfrac{3}{2} k_B T \ln N. \tag{6}$$

Numerics: Before the main analysis we present numerical evidence for the existence of two distinct phases for the ensemble of random RNA. We generate configurations of interaction energies $\varepsilon_{i,j}$'s according to the three models (1), (2), and (3), with $\beta u_m = \frac{1}{2}$ and $\beta u_{mm} = 1$ for weak disorder and $\beta u_m = 10$ and $\beta u_{mm} = 1$ for strong disorder. Then, we calculate the quantities $Z_{i,j}$ in eq. (4), and from that the division free energy $\Delta F = -k_B T \ln[Z_{1,N/2} Z_{N/2+1,N}/Z_{1,N}]$. The result is averaged over 1000 to 10000 disorder configurations and illustrated in fig. 3.

At weak disorder ($u_m < k_B T$), the division free energy follows the molten phase behaviour expected according to eq. (6). Thus, at weak disorder the molten phase description is applicable even if the interaction energies $\varepsilon_{i,j}$ are not uniform. Moreover, we find no difference between the three models of disorder. At strong disorder ($u_m \gg k_B T$), the picture is different. The length dependence of ΔF can still be fit by a logarithm (for $N \geq 160$). A power law with a small exponent is also consistent with the data. The latter was reported by recent numerical studies [11] using a different scheme to force the system away from its ground state. As shown in fig. 3, the prefactors of the logarithms differ from one choice of disorder to another. Had we fitted these to power laws, the observed exponents would also be different for the different disorders. Since different exponents are inconsistent with the expected universality while different prefactors of a logarithm are not, we prefer the interpretation of the data as a logarithmic behaviour. The prefactors of the logarithm in all three models are by far larger

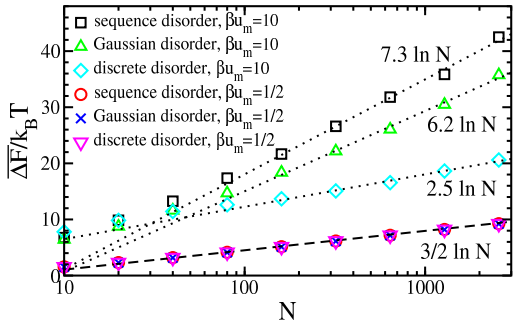

Fig. 3 – Ensemble averaged division free energy $\overline{\Delta F}(N)/k_B T$ for the three models of disorder given by eqs. (1), (2), and (3), at $\beta u_m = \frac{1}{2}$ and $\beta u_m = 10$. The statistical error is no larger than the size of the symbols. At $\beta u_m = \frac{1}{2}$, the data follows the expected $\frac{3}{2} \ln N$ behaviour (dashed line) for all 3 models. At $\beta u_m = 10$, the data is still best fitted (see text) to a logarithmic form $\Delta \cdot \ln N$(dotted lines), with $\Delta \approx 7.3$, 6.2, and 2.5 for the three models.

than the prefactor $\frac{3}{2}k_BT$ expected in the molten phase. This suggests that there is a distinct glassy phase; this finding is reinforced by similar changes detected in other observables [17].

Weak disorder behaviour: We now investigate the stability of the molten phase to weak disorder. This is usually done in statistical mechanics by computing some physical observable perturbatively, and see whether the result provides a finite or diverging correction in the limit of large N. Here we choose to compute the free energy itself, since various observables can be obtained from the thermodynamic derivatives, e.g., the fraction of paired bases from the derivative with respect to $\ln q$. The leading order correction to the free energy due to disorder is obtained from the two-replica partition function $\overline{Z^2(N)}$. For convenience, we will use the uncorrelated Gaussian disorder characterised by eq. (2) for the perturbative study, since the universal properties are expected to be independent of the choices of disorder as demonstrated by the numerical results shown in fig. 3. The ensemble average for $Z^2(N)$ can be explicitly performed in this case, yielding after some algebra

$$\overline{Z^2(N)} = \sum\nolimits_{\{S_1,S_2\in\mathcal{S}(N)\}} q^{|S_1|}q^{|S_2|}\widetilde{q}^{|S_1\cap S_2|} \tag{7}$$

where $q \equiv \exp\left(-\beta\overline{\varepsilon} + \frac{1}{2}\beta^2 D\right)$ and $\widetilde{q} \equiv \exp\left(\beta^2 D\right)$ are the two relevant "Boltzmann factors", and $|S_i|$ and $|S_i \cap S_j|$ are the number of bases contained in structure S_i or common to S_i and S_j respectively. This effective partition function has a simple physical interpretation: It describes two RNA molecules subject to a *homogeneous* attraction with effective interaction energy $\varepsilon_0 \equiv -\beta^{-1}\ln q = \overline{\varepsilon} - \frac{1}{2}\beta D$ between any two bases of the same molecule. In addition, there is an inter-replica attraction characterised by the factor $\widetilde{q}$ for each bond *shared* between the two replicas. The inter-replica attraction is induced by the same disorder shared by the replicas. It can potentially force the replicas to "lock" together, *i.e.*, to become correlated.

The sum in eq. (7) can be performed exactly [17], by first summing over the bonds which are *common* to the two structures, noting that the possible configurations of the common bonds are themselves the set $\mathcal{S}(N)$ of valid secondary structures. Within a given configuration of common bonds, all possibilities to place non-common bonds in the two individual structures can then be summed over, leading to an effective single RNA problem. However, the necessary algebra is quite involved [17]; here we just quote the results. The solution has the form $\overline{Z^2(N)} \sim N^{-\theta}\exp[-\beta N f(q,\widetilde{q})]$, with two different expressions for θ and f depending on whether $\widetilde{q}$ is above or below a critical value $\widetilde{q}_c = 1+1/[q^2\sum_{N=1}^{\infty} N g^2(N)]$, where $g(N) = Z_0(N)/(1+2\sqrt{q})^{N-1} \sim N^{-3/2}$ for large N.

For $\widetilde{q} > \widetilde{q}_c$, we have $\theta = 3/2$ and f given by a complicated function of $Z_0(N)$ [17]. Here, the two-replica partition function has the *same* asymptotic form as that of the single-replica system in (5), implying that the disorder coupling *locks* the two replicas together. The weak disorder limit of interest to the present perturbative study lies in the regime $\widetilde{q} < \widetilde{q}_c$, since as $D \to 0$, $\widetilde{q} \to 1$ while $\widetilde{q}_c$ converges to a (q-dependent) number larger than 1. In this regime, $f(q,\widetilde{q})$ is still different from $2f_0(q)$, indicating the existence of an extensive correction to the free energy f_0 due to weak disorder. However, this in itself does not indicate the relevance of disorder. More important is the value of the exponent, $\theta = 2\theta_0 = 3$, which indicates that in this regime, the two-replica partition function $\overline{Z^2(N)}$ is essentially a product of two single-replica partition functions $Z_0(N)$. Without a coupling between the two replicas, there is no divergent correction to physical observables and we conclude that disorder is irrelevant in this regime.

Strong disorder behaviour: Next, we determine whether the molten phase persists for all disorder strengths. We will consider the sequence disorder model (1) which includes the correlations among the $\varepsilon_{i,j}$'s neglected during the study of the weak disorder regime. In the following, we will *assume* that long random RNA is characterized by the molten phase for *all*

values of u_m, *i.e.*, that the partition function for any long substrand is given by eq. (5), with some effective value of q. We will show that this assumption leads to a contradiction beyond some large but finite value of u_m, implying the existence of a distinct glass phase.

We will again focus on the division free energy ΔF. Under the assumption that the random sequences are described by the molten phase, it is given by eq. (6) for large N *independently* of the effective value of q. On the other hand, we can study this division free energy for each randomly chosen sequence of bases. For each such sequence, we can look for a continuous segment of $\ell \ll N$ Watson-Crick pairs $(b_i, b_j)(b_{i+1}, b_{j-1}) \dots (b_{i+\ell-1}, b_{j-\ell+1})$ where the bases $b_i \dots b_{i+\ell-1}$ are within the first half of the molecule and the bases $b_{j-\ell+1} \dots b_j$ are in the second half. For random sequences, there are roughly N^2 ways to choose the two segments and the probability for each pair to be exactly complementary is $4^{-\ell}$. Thus, the largest ℓ in a sequence of length N is given by $N^2 4^{-\ell} \approx 1$ or $\ell = \ln N / \ln 2$ as also proven rigorously [18].

Now we evaluate the two terms of the division free energy $\Delta F = F_{\rm div} - F_{\rm free}$ separately. The partition function yielding the unconstrained free energy $F_{\rm free}$ contains *at least* all the configurations in which the two complementary segments $b_i \dots b_{i+\ell-1}$ and $b_{j-\ell+1} \dots b_j$ are completely paired. Thus,

$$F_{\rm free} \leq F_{\rm paired} \tag{8}$$

where $F_{\rm paired}$ is the free energy of the ensemble of structures in which the two complementary segments are paired. The latter is the sum of the energy of the paired segments and those of the two remaining substrands of lengths $L_1 = j - i - 2\ell + 1$ and $L_2 = N + i - j - 1$, *i.e.*,

$$F_{\rm paired} = -\ell u_m + (N - 2\ell) f_0 + \tfrac{3}{2} k_B T \big[\ln(L_1) + \ln(L_2)\big]. \tag{9}$$

The free energy $F_{\rm div}$ is, by the assumption of the molten phase, the free energy of the two substrands $b_1 \dots b_{N/2}$ and $b_{N/2+1} \dots b_N$. According to eq. (5), this is $F_{\rm div} = f_0 N + 2 \times \frac{3}{2} k_B T \ln N$ up to terms independent of N. Combining this with eqs. (8) and (9), and noting that ℓ is typically of the order $\ln N / \ln 2$ and L_1, L_2 are typically proportional to N, we obtain $\overline{\Delta F} \geq [u_m + 2 f_0] \ln N / \ln 2$ for very large N. This is only consistent with eq. (6) if

$$\tfrac{3}{2} k_B T \geq [u_m + 2 f_0] / \ln 2. \tag{10}$$

For $u_m \gg k_B T$, the free energy per length f_0 approaches its zero temperature value of u_m times the average number of base pairs per monomer in the minimal energy structure of each sequence. If *all* bases in a sequence could form Watson-Crick base pairs, the free energy per length would be $f_0 = -\frac{1}{2} u_m$ for large u_m. However, if sequence disorder leads to frustration, there is always a finite fraction of bases which cannot be incorporated into Watson-Crick base pairs ([1]) and thus $f_0 \to -c \cdot u_m$ with some $c < \frac{1}{2}$ for $u_m \gg k_B T$. Therefore, the right hand side of the consistency condition (10) becomes arbitrarily large as u_m becomes larger. It follows that there is some value u_m^* above which the consistency condition (10) breaks down, implying the inconsistency of the molten phase assumption in this regime. Beyond this disorder strength the energy gain due to the configuration in which the two exactly complementary substrands of length ℓ are permanently bound outweighs the entropic contribution of an exponential number of coexisting configurations which dominates the molten phase. Thus, the system is in a different phase where the energetically most favourable configuration dominates the ensemble. A detailed characterisation of the glass phase will be discussed elsewhere [17].

Conclusions: We studied the statistical properties of secondary structures formed by random RNA sequences below the denaturation transition using a simplified energy model. A two-replica calculation shows that disorder is perturbatively irrelevant, *i.e.*, an RNA molecule

([1])This is not the case for the sequence disorder model at an alphabet size 2 unless a minimum hairpin length constraint is introduced as in ref. [10].

with weak sequence disorder is in the molten phase where many secondary structures with comparable total energy coexist. By further considering the rare regions of strong sequence complementarity, we show that the molten phase cannot exist for arbitrarily strong disorder, and thereby establish the existence of a distinct glass phase. Our analysis follows the approach used in ref. [19] to show the relevance of disorder in the denaturation of double-stranded DNA. It will be interesting to see whether a renormalization group theory along the line of ref. [19] may be constructed to elucidate the glass phase of the random RNA problem.

We expect that random RNA with realistic energy parameters at physiological conditions to belong to the glass phase since the variation in binding energies is substantially larger than k_BT. On the other hand, long sequences repeats, *e.g.*, AUAUAU... are expected to be in the molten phase. It would be interesting to understand quantitatively how much sequence disorder corresponds to a certain disorder strength D and whether the molten phase description may be applicable to random RNA close to the denaturation temperature.

Acknowledgements: The authors benefitted from discussions with U. Gerland, D. Moroz, and especially L.-H. Tang regarding rare regions. This work is supported by the NSF through grants no. DMR-9971456 and DBI-9970199. The authors are grateful to the hospitality of the Institute for Theoretical Physics (Santa Barbara) where this work was completed.

REFERENCES

[1] See, *e.g.*, GESTELAND R.F. and ATKINS J.F. *eds.*, *The RNA world: the nature of modern RNA suggests a prebiotic RNA world* (Cold Spring Harbor Laboratory Press) 1993.
[2] DILL K.A. *et al.*, *Protein Sci.*, **4** (1995) 561; ONUCHIC J.N., LUTHEY-SCHULTEN Z. and WOLYNES P.G., *Annu. Rev. Phys. Chem.*, **48** (1997) 545; GAREL T., ORLAND H., PITARD E., *J. Phys. I*, **7** (1997) 1201; SHAKHNOVICH E.I., *Curr. Op. Struct. Biol.*, **7** (1997) 29.
[3] TINOCO JR I. and BUSTAMANTE C., *J. Mol. Biol.*, **293** (1999) 271 and references therein.
[4] MCCASKILL J.S., *Biopolymers*, **29** (1990) 1105.
[5] ZUKER M. and STIEGLER P., *Nucl. Acid. Res.*, **9** (1981) 133; HOFACKER I.L., FONTANA W., STADLER P.F., BONHOEFFER S., TACKER M. and SCHUSTER P., *Monatshefte f. Chemie*, **125** (1994) 167; CHEN S.J. and DILL K.A., *Proc. Natl. Acad. Sci. USA*, **7** (2000) 646.
[6] FREIER S.M., KIERZEK R., JAEGER J.A., SUGIMOTO N., CARUTHERS M.H., NEILSON T. and TURNER D.H., *Proc. Natl. Acad. Sci. USA*, **83** (1986) 9373.
[7] DE GENNES P.-G., *Biopolymers*, **6** (1968) 715.
[8] MOROZ J.D. and HWA T., in preparation.
[9] HIGGS P.G., *Phys. Rev. Lett.*, **76** (1996) 704.
[10] PAGNANI A., PARISI G., and RICCI-TERSENGHI F., *Phys. Rev. Lett.*, **84** (2000) 2026, *ibid.*, **86** (2001) 1383; HARTMANN A.K., *ibid.*, **86** (2001) 1382.
[11] KRZAKALA F., MÉZARD M., and MÜLLER M., *Europhys. Lett.* **57** (2002) 752; MARINARI E., PAGNANI A., and RICCI-TERSENGHI F., *Phys. Rev. E* **65** (2002) 041919.
[12] BUNDSCHUH R. and HWA T., *Phys. Rev. Lett.*, **83** (1999) 1479.
[13] HIGGS P.G., *Quart. Rev. Biophys.*, **33** (2000) 199.
[14] GAREL T. and ORLAND H., *Europhys. Lett.*, **6** (1988) 307; SHAKHNOVICH E.I. and GUTIN A.M., *J. Phys. A*, **22** (1989) 1647.
[15] HWA T. and LÄSSIG M., *Phys. Rev. Lett.* **76** (1996) 2591; DRASDO D., HWA T. and LÄSSIG M., *J. Comp. Biol.*, **7** (2001) 115.
[16] WATERMAN M.S., in *Advances in Mathematics, Supplementary Studies*, edited by G.-C. ROTA, Vol. **1** (Academic Press, New York) 1978, pp. 167–212.
[17] BUNDSCHUH R. and HWA T., submitted to *Phys. Rev. E.*
[18] ARRATIA R., GORDON L. and WATERMAN M.S., *Ann. Stat.*, **18** (1990) 539.
[19] TANG L.-H. and CHATÉ H., *Phys. Rev. Lett.*, **86** (2001) 830.