ISSN 1068–3372, Journal of Contemporary Physics (Armenian Academy of Sciences), 2013, Vol. 48, No. 2, pp. 98–102. © Allerton Press, Inc., 2013.
Original Russian Text © G.N. Hayrapetyan, H.L. Tsaturyan, Sh.A. Tonoyan, Y.Sh. Mamasakhlisov, 2013, published in Izvestiya NAN Armenii, Fizika, 2013, Vol. 48, No. 2, pp. 144–150.

Thermodynamics of Single-Stranded RNA with Random Sequence: Constrained Annealing Approach

G. N. Hayrapetyan, H. L. Tsaturyan, Sh. A. Tonoyan, and Y. Sh. Mamasakhlisov

Yerevan State University, Yerevan, Armenia

Received November 2, 2012

Abstract—The effect of sequence disorder on thermodynamics of ssRNA is studied on the basis of constrained annealing approach. A random sequence with bimodal disorder is considered. The temperature behavior of specific heat and helicity degree is examined. A reasonable agreement with numerical results is obtained. In the presence of competing interactions the model exhibits not only partial high-temperature melting, but also partial cold denaturation.

DOI: 10.3103/S1068337213020084

Keywords: RNA, thermodynamics, constrained annealing, melting, cold denaturation

1. INTRODUCTION

Single-stranded RNA plays central role in all live systems. In addition to transmission of genetic information RNA participates actively in different cellular processes [1]. RNA-strand consists of four different types of nucleotides, A, C, G, and U, and can form double helical structures consisting of successive stable A–U or G–C Watson–Crick pairs. Nucleatide sequence affects the three-dimensional native structure of RNA. At the same time, RNA-sequence is the result of evolution and, hence, it is expected that the stability of the secondary structure is important for natural selection [2]. There exist efficient algorithms for exact calculation of partition functions and corresponding thermodynamical parameters of secondary structures of RNA [3, 4].

Stability of the secondary structure can experimentally be measured in the process of denaturation when RNA molecule is losing its tertiary and secondary structures. It is common, like in proteins, to distinguish thermal and cold denaturation. The former occurs at heating [5] while the latter at cooling. The cold denaturation of RNA is studied in [6].

Phase behavior of ssRNA was studied carefully in the context of glass states [7–11], long-loop effects [12], tension [13], and so on. From the point of view of statistical physics, a decisive obstacle on the way to quantitative analytical description of ssRNA is the quenching mean when logarithm of the partition function is disorder averaged, whereas in case of annealing the partition function is itself disorder averaged. The quenching approximation corresponds physically to thermodynamical equilibrium between the degrees of freedom relating to the sequence and structure. Comparative analysis of quenched and annealed ensembles is given [14–17]. Work [16] proposes a technique of constrained annealing which we use in the present work. So far we know it is the first case where the constrained annealing technique is applied to thermodynamics of ssRNA. We evaluate analytically thermodynamic parameters of ssRNA using approach proposed in [16]. The heat capacity we obtain has two peaks indicating two structural transitions. Our analytical results are in good agreement with numerical results obtained on the basis of methods proposed in [3, 4]. We obtained cold denaturation when the RNA molecule loses essentially its secondary structure.

2. MODEL

We propose to consider for simplicity a random sequence of ssRNA consisting of only A and U nucleotides. Topological rules which determine the allowed structures are important for effective numerical calculation of the free energy of secondary structure. The basic rule consists in forbidding of formation of so-called pseudonodes from many accessible secondary structures, as in majority of other works concerning the physics of ssRNA. Thus, for any two pairs of bases (i,j) and (k,l) with $i<j$, $k<l$, and $i<k$ we have either $i<k<l<j$ or $i<j<k<l$ [7]. The secondary structure is by definition

the set of all pairs of bases and each base bay is a part of no more than one pair. Partition function of an arbitrary substrand of ssRNA molecule without pseudonodes is calculated recursively [3, 4] as

$$Z_{i,j} = Z_{i,j-1} + \sum_{k=1}^{j-1} Z_{i,k-1} Q_{ij} Z_{k+1,j-1}, \tag{1}$$

where $Z_{i,j}$ is the partition function of substrand between nucleotides i and j and $Q_{ij} = \exp\left(-\beta\varepsilon_{ij}\right)$ is statistical weight of formation of hydrogen bonds between nucleotides i and j. The Hamiltonian of the model looks like

$$H\left(\hat{m},\{h_i\}\right) = \sum_{i<j} m_{ij}\varepsilon_{ij}, \tag{2}$$

where the interaction constant is $\varepsilon_{ij} = \varepsilon_0 + \varepsilon h_i h_j$, the sum is taken over all non-repeating pairs of bases, $m_{ij} = 1$ if bases i and j are a pair, and $m_{ij} = 0$ otherwise. The variables $\{h_i\}$ describe the type of nucleotides; $h_i = \pm 1$ with $+1$ corresponding to A nucleotide and -1 to U nucleotide. Partition function for ssRNA strand of N nucleotides is written as

$$Z_N\left(\{h_i\}\right) = \sum\nolimits_{\hat{m}}{}' \exp\left[-\beta H\left(\hat{m},\{h_i\}\right)\right], \tag{3}$$

where $\beta = 1/k_B T$ and the sum is taken over all realizations of the $\hat{m}$ without pseudonodes. The matrix $\hat{m}$ contains no more than a unity in every line or column. The last condition describes saturation of base-pairing. A random sequence $\{h_i\}$ is generated in correspondence with the distribution function

$$P\{h\} = \prod_{i=1}^{N} \rho\left(h_i\right), \tag{4}$$

where $\rho\left(h_i\right) = q\delta\left(h_i - 1\right) + \left(1-q\right)\delta\left(h_i + 1\right)$ and $0 < q < 1$. Due to the property of self-averaging, the reduced free energy becomes in the thermodynamic limit $N \to \infty$ not random quantity and

$$f\{h_i\} = f = -\left(1/N\right)\overline{\ln Z_N\left(\{h\}\right)}, \tag{5}$$

where f is the reduced quenched free energy [18], $\overline{O}$ is the average with the distribution function of the sequence (4). According to [16], the free energy of ssRNA with random fixed sequence of nucleotides satisfies the following conditions:

$$f \geq g\left(\beta,\mu\right) \geq f_a, \tag{6}$$

where f_a is the reduced quenched free energy

$$g\left(\beta,\mu\right) = -\frac{1}{N}\ln Z_N = -\frac{1}{N}\ln \overline{Z_N\left(\{h_i\}\right) e^{-N\mu\alpha\left(\{h_i\}\right)}}. \tag{7}$$

Here $\alpha\left(\{h_i\}\right)$ is some self-averaged quantity depending on the sequence. So, $g\left(\beta,\mu\right)$ determines the lower boundary of quenched free energy f. According to inequality (6), the best lower boundary of quenched free energy is given by $\max_\mu g\left(\beta,\mu\right)$, and we can evaluate the free energy of ssRNA molecule with fixed random sequence as

$$f \approx \max_{\mu} g\left(\beta,\mu\right). \tag{8}$$

The simplest constraint imposed on variables $\{h_i\}$, describing the sequence, is given by the expression $\alpha\left(\{h_i\}\right) = \left(1/N\right)\sum_{i=1}^{N}\left[h_i - \left(2q-1\right)\right]$, which does not specify the types of separate monomers h_i, but only the mean value of the sum $\sum_i h_i$. It may be shown that the partition function Z_N defined in formula (7) can after some transformations be written as

$$Z_N = e^{N\mu(2q-1)}\Omega^N Z_N^0\left(\varepsilon_0 + \overline{\varepsilon}\right), \tag{9}$$

where $Z_N^0\left(\varepsilon_0 + \overline{\varepsilon}\right)$ is the partition function (3) for homopolymer ssRNA with effective interaction constant $\varepsilon_{ij} = \varepsilon_0 + \overline{\varepsilon}$. Here

$$\begin{aligned} &\overline{\varepsilon} = -\left(1/\beta\right)\ln\left[W\left(\mu,\beta,\varepsilon\right)/\Omega\left(\mu\right)^2\right], \quad \Omega\left(\mu\right) = qe^{-\mu} + \left(1-q\right)e^{\mu}, \\ &W\left(\mu,\beta,\varepsilon\right) = e^{-\beta\varepsilon}\left[q^2 e^{-2\mu} + \left(1-q\right)^2 e^{2\mu}\right] + 2q\left(1-q\right)e^{\beta\varepsilon}. \end{aligned} \tag{10}$$

As the partition function of homopolymer single-strand RNA may be written in the form [9] $Z_N^0(\varepsilon) \simeq A_0(Q) N^{-3/2} \left(1+2\sqrt{Q}\right)^N$, where $Q = \exp(\beta\varepsilon)$, the variational free energy $g(\beta,\mu)$ for long chains ($N \gg 1$) acquires the form

$$g(\beta,\mu) = -\mu(2q-1) - \ln\Omega(\mu) - \ln\left(1+2\sqrt{\bar{Q}}\right), \tag{11}$$

where $\bar{Q} = e^{-\beta(\varepsilon_0+\bar{\varepsilon})}$. Maximization of the potential $g(\beta,\mu)$ with respect to μ gives the solution $\mu_0(\beta)$, determined by the equation

$$2q-1 = \left[\frac{2\sqrt{\bar{Q}}}{1+2\sqrt{\bar{Q}}} - 1\right]\frac{d\ln\Omega(\mu)}{d\mu} - \frac{1}{2}\frac{2\sqrt{\bar{Q}}}{1+2\sqrt{\bar{Q}}}\frac{\partial \ln W(\mu,\beta,\varepsilon)}{\partial\mu}. \tag{12}$$

3. RESULTS AND DISCUSSION

Entropy per monomer has the form

$$s(\beta) = -g(\beta) + \beta\, dg(\beta)/d\beta, \tag{13}$$

and the heat capacity

$$c_V(\beta) = -\beta^2 \frac{d^2 g(\beta)}{d\beta^2}. \tag{14}$$

Figure 1 compares the behavior of heat capacity obtained by the technique of constrained annealing with that calculated numerically using McCaskill algorithm [4]. The mean value of heat capacity calculated numerically agrees well with that determined by the method of constrained annealing. Temperature dependence of heat capacity exhibits two peaks which is the evidence of two structural transitions.

In order to ascribe the behavior of heat capacity to structural transformations of ssRNA, we define the degree of helicity as a mean fraction of the Watson–Crick base-pairs:

$$\theta = \frac{2}{N}\overline{\left\langle \sum_{i<j} m_{ij} \right\rangle} = \frac{2\sqrt{\bar{Q}}}{1+2\sqrt{\bar{Q}}}, \tag{15}$$

where $\langle \ldots \rangle$ stands for thermodynamical average. It may be shown that the rhs of equation (15) is given by the expression for the degree of helicity of homopolymer ssRNA obtained directly from the partition function of the latter [9]. So, in the approximation of constrained annealing the degree of helicity is written as for homopolymer RNA, but with effective statistical weight $\bar{Q}$.

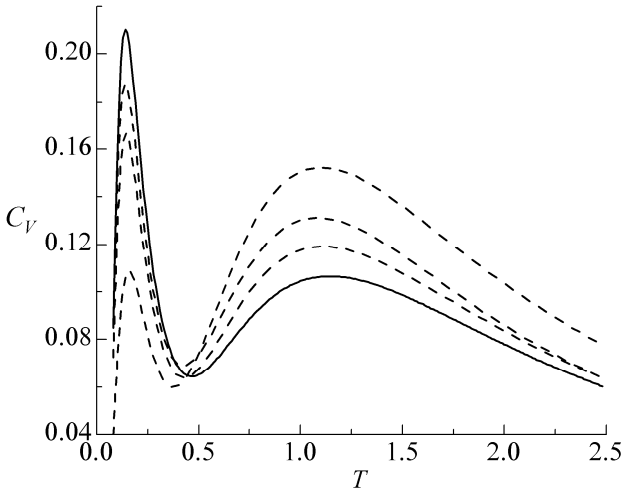

Fig. 1. Dependence of heat capacity per nucleotide (C_V) on the temperature $T = 1/\beta$. Dashed lines are obtained by means of McCaskill algorithm for 3 random realizations for $N = 150$ nucleotides with parameters $\varepsilon_0 = -1$, $\varepsilon = 1.5$, and $q = 0.75$. Solid lines are obtained in variational approximation (8) in thermodynamical limit $N \to \infty$.

The degree of helicity can also be calculated numerically using the probability of formation of base pair between nucleotides i and j [9]:

$$p_{ij} = \langle m_{ij} \rangle = \left(Q_{ij} Z_{i+1,j-1} Z_{j+1,N+i-1} \right) \Big/ Z_{1,N} . \tag{16}$$

The rhs of equation (16) was calculated on the basis of expression (1) and the degree of helicity for a specific realization of nucleotide sequence is determined as

$$\theta = \left(2/N \right) \sum_{i<j} p_{ij} . \tag{17}$$

Figure 2 compares the degree of helicity obtained by the method of constrained annealing with the results obtained on the basis of equations (1), (16), and (17) for a collection of randomly generated sequences. As in case of heat capacity, the mean value of the degree of helicity calculated numerically is in good agreement with that determined by means of the method of constrained annealing. It is seen in Fig. 2 that the degree of helicity increases sharply with the increase in temperature and then, above some temperature near 0.5, decreases. Such behavior of the degree of helicity indicates high- and low-temperature melting and, possibly, denaturation.

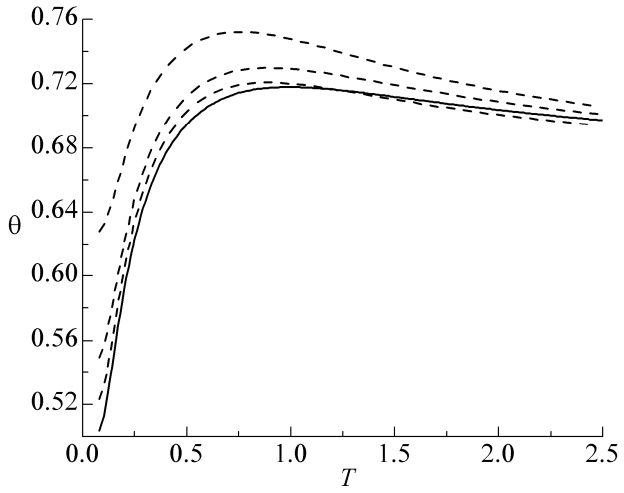

Fig. 2. Dependence of the degree of helicity θ on the temperature $T = 1/\beta$. Dashed lines are obtained by means of McCaskill algorithm for 3 random realizations for $N = 150$ nucleotides with parameters $\varepsilon_0 = -1$, $\varepsilon = 1.5$ and $q = 0.75$. Solid lines are obtained in variational approximation (8) in thermodynamical limit $N \to \infty$.

The high-temperature limit corresponds to the homopolymer case where the contribution from internucleotide interactions is negligible. We neglect for simplicity the temperature dependence of the (free) energy of formation of base-pairs and $\lim_{T\to\infty} \theta = 2/3$. For more realistic choice, for example $\varepsilon_0 = \Delta H - T\Delta S$, the high-temperature limit of the degree of helicity will mainly be determined by entropy loss ΔS in formation of one base pair. ΔH is the enthalpy per base-pair. When comparing with Fig. 1, the low-temperature peak of heat capacity may be attributed to the low-temperature (cold) denaturation and the high-temperature peak to usual thermal denaturation.

Thus, the technique of constrained annealing provides results in good agreement with numerical simulation.

ACKNOWLEDGMENT

The work was performed thanks to support of Volkswagen foundation, grant "Equilibrium and non-equilibrium behavior of single- and double-stranded biological molecules".

REFERENCES

1. *The RNA World*, Gesteland, R.F.and Atkins, J.F., Eds., 1993, vol. 2.
2. Higgs, P.G., *J. Phys.* (*Fr.*), 1993, vol. 1, p. 45.

3. Zuker, M. and Stiegler, P., *Nucleic Acids Res.*, 1981, vol. 9, p. 133.
4. McCaskill, J.S., *Biopolymers*, 1990, vol. 29, p. 1105.
5. Tinoco, I. and Bustamante, C., *J. Mol. Biol.*, 1999, vol. 293, p. 271.
6. Mikulecky, P.J. and Feig, A.L., *J. Am. Chem. Soc.*, 2002, vol. 124, p. 890.
7. Higgs, P.G., *Phys. Rev. Lett.*, 1996, vol. 76, p. 704.
8. Pagnani, A., Parisi, G., and Ricci-Tersenghi, F., *Phys. Rev. Lett.*, 2000, vol. 84, p. 2026.
9. Bundschuh, R. and Hwa, T., *Phys. Rev. E*, 2002, vol. 65, p. 031903.
10. Krzakala, F., Mezard, M., and Mueller, M., *Europhys. Lett.*, 2002, vol. 57, p. 752.
11. Lassig, M. and Wiese, K.J., *Phys. Rev. Lett.*, 2006, vol. 96, p. 228101.
12. Einert, T.R., Nager, P., Orland, H., and Netz, R.R., *Phys. Rev. Lett.*, 2008, vol. 101, p. 048103.
13. Manosas, M. and Ritort, F., *Biophys. J.*, 2005, vol. 88, p. 3224.
14. Morita, T., *J. Math. Phys.*, 1964, vol. 5, p. 1401.
15. Kuhn, R., *Markov Processes Relat. Fields*, 2004, vol. 10, p. 523.
16. Serva, M. and Paladin, G., *Phys. Rev. Lett.*, 1993, vol. 70, p. 105.
17. Liu, T. and Bundschuh, R., *Phys. Rev. E*, 2005, vol. 72, p. 061905.
18. Mezard, M., Parisi, G., and Virasoro, M., *Spin Glass Theory and Beyond*, Singapore: World Scientific, 1987.