EUROPHYSICS LETTERS 20 October 1994

Europhys. Lett., 28 (3), pp. 169-174 (1994)

Collapse of Randomly Self-Interacting Polymers.

Y. KANTOR (*) (**) (***) and M. KARDAR (*)

(*) *Department of Physics, Massachusetts Institute of Technology Cambridge, MA 02139, USA*
(**) *Physics Department, Harvard University - Cambridge, MA 02138, USA*
(***) *School of Physics and Astronomy, Tel Aviv University Tel Aviv 69978, Israel* (§)

(received 15 June 1994; accepted in final form 16 September 1994)

PACS. 36.20 – Macromolecules and polymer molecules.
PACS. 35.20B – General molecular conformation and symmetry; stereochemistry.
PACS. 64.60 – General studies of phase transitions.

Abstract. – We use complete enumeration and Monte Carlo techniques to study self-avoiding walks with random nearest-neighbor interactions described by $v_0 q_i q_j$, where $q_i = \pm 1$ is a quenched sequence of «charges» on the chain. For equal numbers of positive and negative charges ($N_+ = N_-$), the polymer with $v_0 > 0$ undergoes a transition from self-avoiding behavior to a compact state at a temperature $\theta \approx 1.2 v_0$. The collapse temperature $\theta(x)$ decreases with the asymmetry $x = |N_+ - N_-|/(N_+ + N_-)$ and vanishes at $x \approx 0.6$. For $v_0 < 0$, a θ-point is present at all x.

Polymers in a solvent are subject to both the hard-core repulsions between the monomers and somewhat longer-range (*e.g.* van der Waals) attractive interactions. At high temperatures T, the repulsive interactions are dominant, and the radius of gyration (root-mean-squared size) of the polymer scales with the number of monomers N as $R_g \sim N^\nu$, where $\nu \approx 0.588$ in three dimensions. On a discrete lattice, the hard-core interactions are usually modeled by self-avoiding walks (SAWs). The attractive interactions are then included by introducing a negative energy for each pair of monomers residing on neighboring lattice sites. As temperature is lowered, a point is reached ($T = \theta$) where the repulsive and attractive interactions effectively cancel and the polymer behaves essentially as an ideal random walk [1,2] with $\nu = 1/2$. For $T < \theta$, the polymer collapses into a compact object with $\nu = 1/3$. Numerous Monte Carlo (MC) and exact-enumeration studies of the θ-point have been performed (see ref. [3] and references therein).

The collapse transition for *heteropolymers* is particularly interesting in view of its possible relation to the problem of *protein folding* [4,5]. While models based on random heteropolymers significantly oversimplify the specificity and complexity of real proteins, they do bring in fresh perspectives from the statistical mechanics of random systems and spin

(§) Permanent address.

glasses [6-11]. We shall consider polymers formed from two types of monomers, labeled by $q_i = \pm 1$, and subject to a short-range interaction

$$\mathscr{H}_{\mathrm{I}} = \frac{1}{2} \sum_{i,j} q_i q_j \mathscr{V}(\boldsymbol{r}_i - \boldsymbol{r}_j), \tag{1}$$

where $\boldsymbol{r}_i$ are the coordinates of the monomers. Do these interactions modify the collapse transition of heteropolymers described above? A perturbative answer is provided by noting that [12] the relevance of such interactions in d space dimensions is controlled by the scaling exponent $y_{\mathrm{I}} = 1 - d\nu$. The inhomogeneities are marginal in the compact state ($\nu = 1/d$), and may cause the polymer to choose a particular configuration dependent on the specifics of the sequence. The statistical properties of such states are the subject of considerable current interest [13, 14]. Here we address the simpler question of how $\mathscr{H}_{\mathrm{I}}$ affects the non-compact states. For ideal (non-interacting) chains with $\nu = 1/2$, $\mathscr{H}_{\mathrm{I}}$ is relevant in $d < 2$: in $d = 1$ the polymer swells if like charges repel, and collapses to a point, if they attract [12] (see also ref. [15, 16]). For all other non-compact states ($\nu > 1/d$) *weak* interactions are never relevant. However, we shall demonstrate that strong interactions described by $\mathscr{H}_{\mathrm{I}}$ do lead to the collapse of a self-avoiding chain.

We investigate a discretized model in which the only homogeneous interaction is the repulsion caused by the constraints of self-avoidance. In the absence of randomness the polymer explores all SAWs with equal probabilities. A short-range random interaction is incorporated by assigning an energy $v_0 q_i q_j$ to every pair of monomers, i and j, on *neighboring* lattice sites. We primarily focussed on $v_0 > 0$, for which like charges repel and opposite ones attract. This choice corresponds to strongly screened Coulomb interactions. The asymmetry in the amount of positive and negative charges is measured by $x \equiv |N_+ - N_-|/N$. We shall show that for moderate values of x the polymer undergoes a θ-transition and explore its properties. *Negative* v_0 describes a situation in which the like monomers attract, and unlike ones repel. A mixture of hydrophobic and hydrophilic monomers would exhibit such a tendency. In this case the θ-transition is present for all x.

We used exact enumeration to study the properties of chains of up to $L = 12$ steps ($N = 13$ monomers). We examined all spatial conformations of SAWs and all possible quenched sequences of charges. Taking advantage of rotation and reflection symmetries, as well as the degeneracies related to inverting the order of the sequence or the signs of all charges, we reduced the number of independent configurations to 4 162 866 for 2080 sequences, *i.e.* a total of roughly $8 \cdot 10^9$ possibilities. Since the number of cases grows by an order of magnitude when L is increased by unity, we could not go beyond $L \leqslant 12$. By comparison, the maximal chain lengths considered in enumeration studies of *compact* walks in $d = 3$ [14] and $d = 2$ [17] are about *twice longer*. In the two-dimensional case the number of spatial conformations is smaller by 2 to 3 orders of magnitude, while in the former case only a random sample of quenches was used. Our results were supplemented by MC simulations for $L \leqslant 95$, using the pivoting method [18]. (The method provides excellent equilibration at high and intermediate temperatures, but does not permit equilibration of almost compact structures [19].)

Exact enumeration provides a quite detailed picture of the energy landscape. Figure 1 depicts contours of the density of states in the variables R_g^2 and E for *a*) a homogeneous chain with $q^i = 1$ for all i, and *b*) a particular quench that is approximately neutral. All energy levels are proportional to v_0 and fig. 1 corresponds to choosing $v_0 = +1$. The corresponding densities for $v_0 < 0$ are obtained by simply reflecting the figures around $E = 0$. From such histograms, the behavior as a function of temperature is deduced as follows: at high T the system explores the region of highest density and moves to the lowest-energy states at $T \to 0$. Both chains are SAWs at $T \to \infty$ with $R_g^2 \approx 2.4$. As T is reduced the uniformly charged

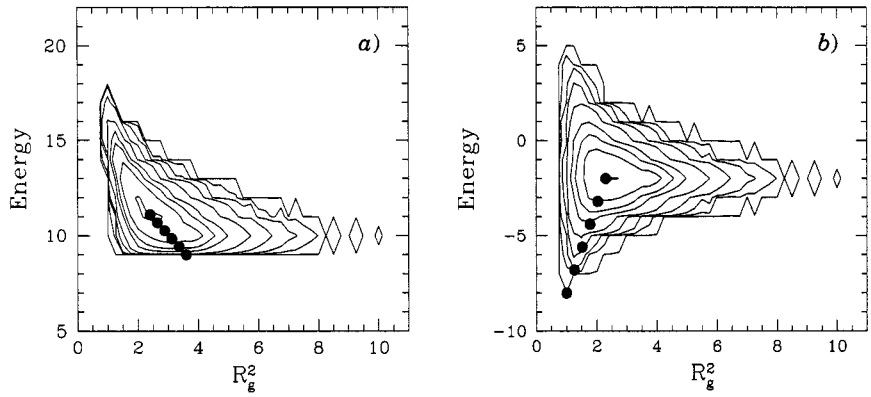

Fig. 1. – Contour plots of the number of states as a function of R_g^2 (in lattice units) and energy (in units of v_0) for a 10-step polymer, in *a*) uniformly charged case ($N_+ = 11$), and *b*) for one quench in an almost neutral case ($N_+ = 6$, $N_- = 5$). The data bins are of size 1 on the energy scale, and size 1/4 on the R_g^2 scale. The smooth contour lines were created by interpolation at levels 0.5, 33, 129, 513, 2049, 8000, 16 000, 32 000, 64 000. The full circles indicate the location of average R_g^2 as the temperature changes between 0 and ∞.

chain expands; its scaling at $T = 0$ is still that of a SAW, as the only effect of the potential is to exclude configurations with chain segments on neighboring sites. This effectively increases the range of the excluded-volume interaction and swells R_g by a finite factor. The random chain in fig. 1*b*) behaves quite differently, collapsing to a dense low-energy configuration as $T \to 0$.

If the collapse of random chains is through a θ-transition similar to homogeneous polymers, we expect that at the transition $R_g^2 \sim L$ as in ideal random walks. This expectation is confirmed in fig. 2 which depicts the T-dependence of (quench-averaged) R_g^2/L for T's ranging from 3 to 95. (The quench average is exact for enumeration data, and obtained from 20 quenches in the MC simulations.) The curves intersect in the vicinity of the same point. Since for $T < \theta$ the polymers are compact ($R_g^2/L \sim L^{-1/3}$), while for $T > \theta$ they are expanded SAWs ($R_g^2/L \sim L^{0.17}$), the graphs representing larger L's show steeper crossover. Despite a slight L-dependence of the intersection points we can locate the θ-temperature with reasonable accuracy at $T \approx 1.2v_0$, but because of small chain lengths the crossover exponents cannot be determined reliably. We also examined the heat capacity per monomer which exhibits a sharp peak whose height increases with L. The position of the peak is close to $T = 0.5v_0$ for short chains and moves towards higher T's for longer chains.

A θ-transition in a *nonrandom* heteropolymer was observed by Victor and Imbert in ref. [20]. These authors consider an *alternating* sequence (hence at $x = 0$) of charges, subject to long-range Coulomb interactions. Due to self-screening, the interactions are effectively short-ranged, and the resulting behavior is not dissimilar to the θ-transition of homopolymers. Randomizing the sequence in a chain with Coulomb interactions, however, leads to a quite different behavior [21].

What happens to the collapse transition if the charges on the chain are not exactly balanced ($x > 0$)? Due to the excess of repulsive interactions, the chains initially swell on lowering temperature. Nevertheless, at sufficiently low temperatures, they may find some compact configurations of low energy. The θ-temperature is expected to decrease due to charge imbalance. Figure 3 depicts partial plots of R_g^2/L *vs.* T at $x = 0$, 1/2, and 5/6. The curves for $x = 1/2$ are qualitatively similar to $x = 0$, but with a lower θ-temperature. For

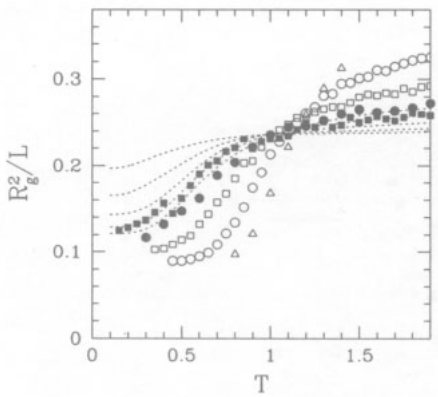

Fig. 2.

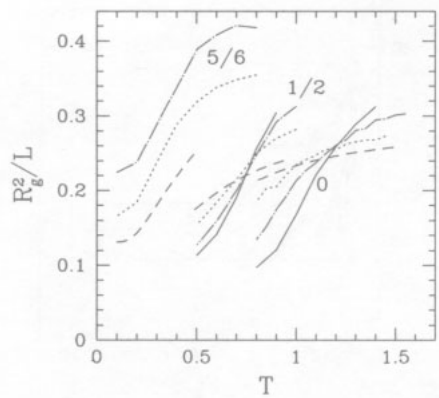

Fig. 3.

Fig. 2. – R_g^2 (in lattice constants) divided by the length of the walk L as a function of temperature, for $N_+ = N_-$. Dashed lines represent the results of exact enumeration for $L = 3, 5, 7, 9$ and 11, in order of increasing slope. Full squares and circles, open squares and circles, and open triangles represent MC results for $L = 9$, 13, 23, 47 and 95, respectively.

Fig. 3. – R_g^2 (in lattice constants) divided by the length of the walk L, for $L = 95$ (solid lines), $L = 47$ (dot-dashed lines), $L = 23$ (dotted lines), and $L = 1$ (dashed lines). The groups of lines represent (from right to left) the values of $x = 0$, 1/2, and 5/6.

$x = 5/6$, however, the curves do not intersect and there is no collapse transition; the chain is a SAW at both high and low temperatures. An approximate phase diagram in the (x, T) plane is constructed in fig. 4: the θ-temperature decreases monotonically with x, finally disappearing at $x \simeq 0.6$. In the $T \to 0$ and $x \to 1$ limit, the action of minority charges is to form loops in the SAW of majority charges. This is somewhat reminiscent of a model of SAWs with periodically distributed «stickers» which attach in pairs [22]. Unlike our model, the minority stickers are not randomly situated and do not stick to the majority monomers. With such stickers, the collapse transition apparently persists all the way to $x = 1$.

The phase diagram is different for $v_0 < 0$, *i.e.* when like charges attract. The limit of $x = 1$ describes a homogeneous polymer with nearest-neighbor attractions. This is precisely the model used to investigate the collapse of uniform polymers, and is known to undergo a θ-transition at $T \approx 3.7\,|v_0|$ [23]. At the other extreme, when $x = 0$, the situation somewhat

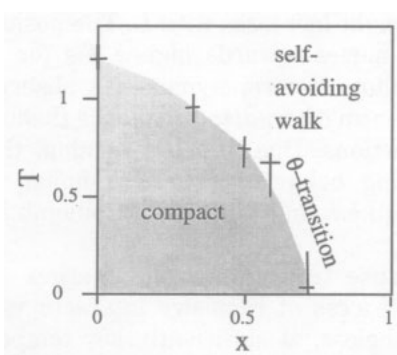

Fig. 4. – Phase diagram of a random polymer with $v_0 > 0$ in the plane of temperature (in units of v_0), and asymmetry x. Vertical bars indicate estimated uncertainties in the extrapolated values of the θ-points.

resembles the case of $v_0 > 0$. Although the model is *not symmetric* under changing of the sign of v_0 which corresponds to changing the sign of E in fig. 1, the energy landscape of fig. 1*b*) appears approximately unchanged by such a reflection. Therefore, it is not surprising that there is again a collapse for $x = 0$ at $\theta \approx 1.2\,|v_0|$. We verified that the model with $v_0 < 0$ does indeed have a collapse temperature for all values of x which interpolates between the above two limits.

Most models of protein folding assume that a *homogeneous* short-range attraction leads to formation of a compact phase, whose details are then determined by the particulars of the sequence. In this work we have shown that even random sequences with no «built-in» overall attraction can create a compact state. Related behavior is observed in models where instead of considering a random sequence, a random interaction is assigned to any pair of monomers. A replica analysis of such polymers indicates a similar tendency towards collapse, even in the absence of a bias in the average interaction [24, 25]. Our polymers are not long enough to determine the critical behavior at the collapse transition. (Even for uniform polymers, chains with $L > 200$ are needed to determine exponents with any degree of accuracy.) However, scaling arguments suggest that randomness is irrelevant for θ-chains, and that the collapse transition should be in the same universality as that of uniform polymers [26].

Some aspects of the model deserve further study. It is plausible that a random polymer initially collapses to a dense-liquid («molten globule») state, followed by a glass transition upon further cooling. (Studies that only enumerate compact states [14], as well as models with random pairwise interactions [25], appear to find such a transition.) We could not investigate such transitions due to either limitations on polymer size in exact enumeration, or difficulties in equilibrating MC runs in the appropriate (low) temperature range. It would also be interesting to investigate the behavior of polymers close to the point at which the θ-transition disappears. Detailed investigation of this point, however, requires use of much longer chains ($L \approx 1000$) and methods which permit reliable equilibration of dense structures at low temperatures.

* * *

This work was supported by the US-Israel BSF grant No. 92-00026, by the NSF through No. DMR-87-19217 (at MIT's CMSE), DMR-91-15491 (at Harvard), and the PYI program (MK). YK would like to thank D. ERTAŞ for numerous discussions of the subject.

REFERENCES

[1] DE GENNES P. G., *J. Phys. (Paris) Lett.*, **36** (1975) L-55.
[2] DUPLANTIER B., *J. Phys. (Paris) Lett.*, **41** (1990) L-409.
[3] KREMER K., *J. Phys. A*, **15** (1981) 2879.
[4] See, *e.g.*, CREIGHTON T. E., *Proteins: Their Structure and Molecular Properties* (Freeman, San Francisco) 1984.
[5] CHAN H. S. and DILL K. A., *Phys. Today*, February issue (1993), p. 24.
[6] STEIN D. L., *Proc. Natl. Acad. Sci. USA*, **82** (1985) 3670; BRYNGELSON J. D. and WOLYNES P. G., *Proc. Natl. Acad. Sci. USA*, **84** (1987) 7524.
[7] GAREL T. and ORLAND H., *Europhys. Lett.*, **6** (1988) 307.
[8] SHAKHNOVICH E. I. and GUTIN A. M., *Europhys. Lett.*, **8** (1989) 327.
[9] KARPLUS M. and SHAKHNOVICH E., in *Protein Folding*, edited by T. E. CREIGHTON (Freeman & Co., New York, N.Y.) 1992, chapt. 4, p. 127.
[10] IORI G., MARINARI E., PARISI G. and STRUGLIA M. V., *Physica A*, **185** (1992) 98.
[11] SFATOS C. D., GUTIN A. M. and SHAKHNOVICH E. I., *Phys. Rev. E*, **48** (1993) 465.

[12] KANTOR Y. and KARDAR M., *Europhys. Lett.*, **14** (1991) 421.
[13] SHAKHNOVICH E. I. and GUTIN A. M., *Nature*, **346** (1990) 773.
[14] SHAKHNOVICH E. I. and GUTIN A. M., *J. Chem. Phys.*, **93** (1990) 5967.
[15] FOSTER D. P., VANDERZANDE C. and YEOMANS J., *J. Stat. Phys.*, **69** (1992) 857.
[16] STEPANOW S., SCHULTZ M. and SOMMER J.-U., *Europhys. Lett.*, **19** (1992) 273.
[17] CAMANCHO C. J. and THIRUMALAI D., *Phys. Rev. Lett.*, **71** (1993) 2505.
[18] LAL M., *Mol. Phys.*, **17** (1969) 57; MADRAS N. and SOKAL A. D., *J. Stat. Phys.*, **50** (1988) 109.
[19] KANTOR Y. and KARDAR M., in preparation.
[20] VICTOR J. M. and IMBERT J. B., *Europhys. Lett.*, **24** (1993) 189.
[21] KANTOR Y. and KARDAR M., *Europhys. Lett.*, **27** (1994) 643.
[22] CATES M. E. and WITTEN T. A., *Macromolecules*, **19** (1986) 732; BALJON A. R. C., *Macromolecules*, **26** (1993) 4339.
[23] JANSSENS M. and BELLEMANS A., *Macromolecules*, **9** (1976) 303.
[24] SHAKHNOVICH E. I. and GUTIN A. M., *Biophys. Chem.*, **34** (1989) 187.
[25] DINNER A., ŠALI A., KARPLUS M. and SHAKHNOVICH E. I., *J. Chem. Phys.*, **101** (1994) 1444.
[26] OBUKHOV S. P., *J. Phys. A*, **19** (1986) 3655.