Randomly Charged Polymers

Mehran Kardar
Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Yacov Kantor
School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69 978, Israel

Polyampholytes (PAs) are polymers with a random sequence of positive and negative charges along their backbone. We have studied systematically the dependence of internal energy and shape of the PA on its excess charge by combining analytic arguments, Monte Carlo simulations, and exact enumeration of all configurations of short chains. The results indicate that the overall excess charge, Q, is the main determinant of the size of the PA. A polymer composed of a mixture of N positive and negative charges $\pm q_0$, is compact for $Q < Q_c \approx q_0\sqrt{N}$, and expanded otherwise. The transition between the two states at low temperatures is reminiscent of the Rayleigh shape instability of a charged drop. A uniform excess charge causes the breakup of a fluid drop; we show that a uniformly charged polymer stretches out to a *necklace shape*. The inhomogeneities in charge distort the shape away from an ordered necklace. The freezing transition of a PA, and its relevance to proteins is also discussed.

I. INTRODUCTION

Randomly charged polymers, or *polyampholytes* (PAs), are of interest to statistical mechanics for several reasons: (1) While the physics of homogeneous polymers is fairly well understood [1], considerably less is known about heteropolymers. The seemingly random sequence of monomers in the latter, provides an interesting challenge to the physics of inhomogeneous and glassy systems [2]. Simple heteropolymers have been studied in connection with the folding of proteins. [3] Since real proteins incorporate charged amino acids, they are examples of PAs [4]. (3) The long–range nature of the Coulomb interactions profoundly effects the behavior of polymers. Experiments on charged heteropolymers [5] and gels [6] reveal interesting and puzzling aspects in need of explanation.

Our model PAs are polymers of length N, where each monomer (pictured as a sphere of size a) has a charge $q_i = \pm q_0$ $(i = 1, \cdots, N)$. Configurations of a polymer are completely specified by listing the position vectors $\{\mathbf{r}_i\}$ of its monomers. The energy of each configuration is then given by

$$\mathcal{H} = \sum_{i=1}^{N-1} \mathcal{V}\left(\mid r_{i+1} - r_i \mid\right) + \sum_{i<j}^{N} \left[\mathcal{U}\left(\mid r_i - r_j \mid\right) + \frac{q_i q_j}{\mid r_i - r_j \mid^{d-2}}\right]. \tag{1}$$

The first term describes the covalent bonding of neighboring monomers. The potential $\mathcal{V}$ diverges at short and long distances, to forbid monomer overlap, and bond breaking, respectively. The excluded volume between non-neighboring pairs of monomers is described by the potential $\mathcal{U}$. Finally, the last term in Eq.(1) describes the (unscreened) Coulomb interaction between the charges q_i and q_j. Note that, when generalized to d–dimensional space, the Coulomb interaction falls off with separation as $1/r^{d-2}$.

We have performed several numerical and analytical studies of the above Hamiltonian [7–10]. Our analytical studies included a high temperature expansion on a continuum version of the problem, *without excluded volume effects*, i.e. starting with

$$\frac{\mathcal{H}}{T} = \frac{K}{2}\int_0^N dx \left(\frac{d\vec{r}}{dx}\right)^2 + \frac{1}{2T}\int dx dx' \frac{q(x)q(x')}{|\vec{r}(x) - \vec{r}(x')|^{d-2}}. \tag{2}$$

We also employed an analogy to charged droplets which will be described in later sections.

For purposes of *Monte Carlo* (MC) simulations [9], we employed a "tether potential" $\mathcal{V}$, which allows the bonded monomers to explore configurations in which their separation is within an allowed range (without any additional energy cost). In these simulations the relaxation time at high temperatures scales as N^4, and we typically looked at chains with $N \leq 128$. However, as the PA falls out of equilibrium at low temperatures, we only examined temperatures $T \geq 0.01 q_0^2/a$ (the basic energy scale of the problem). To perform quench averages, we typically examined 10 different realizations of random sequences.

For exact enumeration studies [10], we looked at *all* self-avoiding random walk configurations on cubic lattices. We also examined all sequences. Since the numbers of configurations (and sequences) grows exponentially with N, we

were limited to lengths $N \leq 13$. The exact enumeration allows us to look at all temperatures, and to do exact quench averages, at the expense of severe finite size limitations.

II. EXPANSION VERSUS COLLAPSE

The spatial extent of a polymer is characterized by the critical exponent ν, which relates its radius of gyration (r.m.s. size) R_g to the number of monomers N, through the power law $R_g \propto N^{\nu}$. A polymer is "compact" if $\nu = 1/d$, where d is the dimension of the embedding space, and fully "stretched" if $\nu = 1$. For self-avoiding polymers $\nu \approx 0.6$ in $d = 3$ [1]. Perhaps the simplest question regarding the PA is whether the introduction of Coulomb interactions causes the polymer to be more or less expanded. There was initially some controversy regarding this question, which is now resolved. If the charges within a PA behave as in a plasma, they will rearrange themselves so as to neutralize excess charge in any region of space. The energy gain from such a screening will provide an inward pressure which compactifies the PA [11]. On the other hand, for a sequence in which each monomer takes a charge $q_i = \pm q_0$ independent of the others; there is a "typical" excess charge of $Q_c \equiv q_0 N^{1/2}$. Independent of its sign, the excess charge produces a repulsive energy, which should swell the PA [7].

The apparent contradiction between the above results was resolved in Ref. [8]: If the monomers comprising the chain are assembled together in an organic solvent, the strong Coulomb forces ensure the neutrality of the resulting polymer. The neutral polymer in any environment then behaves as a plasma and compactifies to screen the charges. There is no such bias towards neutrality if the polymer is prepared in a solution where the Coulomb forces are screened. The typical polyampholyte will then have a net charge due to the random imbalance of positive and negative components. Such net charge is actually sufficient to stretch the chain in an environment with no Coulomb screening. This is nicely confirmed in MC simulations: the size of a neutral PA monotonically decreases with temperature, while typical unconstrained sequences expand at low temperatures due to their excess charge.

In a high temperature expansion, the Coulomb interaction is treated as a perturbation, and various quantities are obtained as a series in inverse temperature. For example, the radius of gyration is given by

$$\overline{\langle R^2 \rangle} = \langle R^2 \rangle_0 \left[1 + A \left(\frac{T_0}{T} \right) + B \left(\frac{T_0}{T} \right)^2 + \cdots \right], \tag{3}$$

where $\overline{\cdots}$ and $\langle \ldots \rangle$ denote averaging over sequences and conformations respectively. The characteristic temperature is obtained by dimensional analysis, and given by

$$T_0 \equiv \frac{q_0^2 N}{\langle R^2 \rangle_0^{\frac{d-2}{2}}} \propto N^{1-(d-2)\nu}. \tag{4}$$

In $d \leq 4$, T_0 grows with N signalling the breakdown of perturbation theory, and the relevance of interactions. In Ref. [8], the coefficients A and B were calculated, starting from an ideal (non self–avoiding) chain at infinite temperature: for a neutral PA, $A = -\sqrt{192/4050\pi}$, while for a random sequence $A = 0$ and B is negative.

Quite generally, the coefficient A is proportional to the correlator of charges, i.e. $A \propto \overline{q_i q_j}$. While $\overline{q_i q_j} = 0$ for independently selected charges, constraining the net charge to Q leads to $\overline{q_i q_j} = (Q^2 - q_0^2 N)/N^2$. The first term in the high temperature expansion thus changes sign at $Q = Q_c = q_0\sqrt{N}$. We confirmed this behavior by MC simulations of PAs with fixed charge. For $N = 64$, chains with $Q > 8q_0$ expand upon lowering temperature, while those with $Q \leq 8q_0$ collapse. Rather surprising, the radii of gyration appear to vary monotonically with temperature, and the above trends continue to low temperatures. Histograms constructed at $T = 0.05 q_0^2/a$ suggested that Q is a good determinant of the size of the PA, with polymers of different Q falling into distinct bands [9].

III. EXCESS CHARGE AND THE NECKLACE MODEL

The high temperature expansion *does not* explain why PAs are compact for $Q \approx Q_c$ at low temperatures. Instead, we use the observation that PAs with $Q = 0$ form dense globules, as a starting point for the investigation of the dependence of low–T shapes on Q in $d = 3$. The energy (or rather the quench–averaged free energy) of the PA is then phenomenologically related to its shape by

$$E = -\epsilon_c N + \gamma S + \frac{bQ^2}{R}. \tag{5}$$

The first term is a condensation energy proportional to the volume (assumed compact) with $\epsilon_c \sim q_0^2/a$; the second term is proportional to the surface area S (with a surface tension $\gamma \sim q_0^2/a^3$), while the third term represents the long–range part of the electrostatic energy due to an excess charge Q. (For a uniformly charged spherical drop $b = 0.6$, while the migration of charges to the surface reduces b to 0.5 in a conducting drop.)

The optimal shape is obtained by minimizing the overall energy in Eq.(5). The first term is the same for all compact shapes, while the competition between the surface and electrostatic energies is controlled by the dimensionless parameter

$$\alpha \equiv \frac{Q^2}{16\pi R^3 \gamma} = \frac{Q^2}{12 V \gamma} \equiv \frac{Q^2}{Q_R^2} \ . \tag{6}$$

Here, R and V are the radius and volume of a spherical drop of N particles, and we have defined the *Rayleigh charge* Q_R. We note that in the case of PAs [9,10] $Q_R \approx q_0 N^{1/2} = Q_c$. The behavior of a system described by Eq.(5) was first analyzed more than a century ago in the context of charged liquid drops by Lord Rayleigh [12,13]. He showed that a conducting drop is locally unstable to an infinitesimal sphroidal distortion for $\alpha > 1$. In fact, if the possible shapes of a globule are restricted to spheroids, this instability is preempted by a discontinuous transition [9] to a prolate shape for $\alpha \gtrsim 0.899$. Both MC simulations [9] and experimental results [6] indicate that for $Q < Q_R$, the R_g of PAs at low temperature is almost independent of Q, while it increases extremely fast as a function of Q for larger charges. The question of the stable shape of the PA for $Q > Q_R$ is taken up next.

Although the ellipsoidal globule may have a lower energy than the spherical one, the former is *not* the minimum energy configuration of a charged drop. This was first noted in the context of liquid drop models of atomic nuclei [14] A uniformly charged nucleus minimizes its energy by undergoing *fission*, and splitting into two equal parts [15–17]. If the resulting parts (droplets) can be infinitely separated, this instability occurs for $\alpha > 0.3$. Obviously, the PA must maintain its connectivity and cannot undergo fission. We proposed in Ref. [9] that as a compromise a charged polymer splits into two droplets that are connected by a narrow tube of length L and diameter a (a dumbbell shape). As long as $La^2 \ll R^3$, most of the charge remains in the spherical globules. The total electrostatic energy is proportional to Q^2/L, while the energy cost of creating the tube (from surface or condensation energy) grows as $\gamma a L$; equating the two gives $L \propto Q$. For larger Q, a liquid drop will split into a larger number of droplets. The corresponding picture for the PA is a *necklace* of globules connected by narrow strands. This picture has been independently confirmed by MC simulations of uniformly charged polymers [18].

The astute reader will note that the picture developed in the previous paragraph in fact applies to a uniformly charged polymer, such as a polyelectrolyte in a poor solvent [18]. Inhomogeneities in charge are expected to modify this scenario. If a random PA with $\alpha > 0.3$ splits into globules of charges Q_1 and Q_2, it can be shown that for each globule $\overline{\alpha_{\rm subchain}} = (1+\alpha)/2$, while the mean product of the charges is $\overline{Q_1 Q_2} = q_0^2 N(\alpha - 1)/4$. For $\alpha = 1$, each globule has, on average, values of α close to unity, while the average value of the product of charges vanishes. We thus have the paradoxical situation in which most spherical shapes are unstable, while there is on average no energetic gain in splitting the sphere into two equal parts.

While charge inhomogeneities drastically modify the necklace picture, the resulting PA is probably still composed of rather compact globules connected by a (not necessarily linear) network of tubes. The globules are selected preferentially from segments of the chain that are approximately neutral (or at least below the instability threshold), while the tubes are from subsequences with larger than average excess charge. A few globules can indeed be identified in recent MC simulations of PAs [19]. It can be shown [20] that the probability of finding a very large neutral segment in a random sequence of charges is large, i.e. it is possible to build a configuration consisting of one very large neutral globule with highly charged "tails" sticking out of it.

IV. FREEZING TRANSITION

Compact heteropolymers with short range interactions are known to undergo a freezing transition in which the number of thermodynamically relevant states goes from an exponentially large value ($\mathcal{O}(e^N)$) in the random globule state, to only a few ($\mathcal{O}(1)$) conformations in the frozen state [3]. This transition is believed to be well described by the Random Energy Model (REM) [21]. The principal assumption of REM is that, for a given sequence, the energies of different configurations are unrelated, i.e. $\overline{E_\alpha E_\beta} \approx 0$. This is certainly not true, as configurations related by simple moves will have correlated energies. The approximate validity of REM resides in such configurations making a negligible contribution to the overall probability distribution for overlaps $\mathcal{P}(E_\alpha E_\beta)$, which is expected to approach a delta function as N becomes large.

In REM the single energy density of states is assumed to be a Gaussian of width $u\sqrt{N}$. Typically the number of conformations of a polymer scales as $\mathcal{M} \sim e^{\omega N}$, with ω of the order of unity, resulting in an entropy

$$S_{\rm REM}(E) = \ln\left[\exp\left(\omega N - \frac{E^2}{2Nu^2}\right)\right]. \tag{7}$$

On lowering temperature, the number of available states decreases, leading to an entropy crisis, and hence a freezing transition at $T_f = u/\sqrt{2\omega}$.

The exact enumeration studies of short chains [10] also provided some preliminary evidence of a freezing transition in PAs. However, it is unclear how Coulomb interactions modify freezing. Indeed the basic assumption of REM is not valid as the energies in a PA are highly correlated. In fact dimensional analysis, supported by numerical studies [22], indicates that

$$\overline{E_\alpha E_\beta} \approx \frac{N^2}{R^{2(d-2)}} \propto u^2 N^{4/d}. \tag{8}$$

Enumeration of energy levels for chains of up to $N = 36$ indicates that while the mean energy fluctuates strongly with sequence, each sequence has an approximately Gaussian density of states. If we naively apply the REM freezing criterion, we thus obtain a $T_f \propto N^{(4-d)/2d}$, i.e. the freezing temperatures goes up with N, and all chains are frozen in the thermodynamic limit for $d \leq 4$.

It is tempting to compare this conclusion with an earlier result [8] that PA interactions also remove the collapse (θ) transition of chains in $d < 4$, replacing it by a smooth crossover. This is probably just a coincidence as the result for freezing is obtained from energy *correlations* ($N^2/[N^{1/d}]^{d-2}$) for *compact* chains of the size $\sim N^{1/d}$, while the earlier result is obtained by considering the interaction of two ideal chains, with charge imbalance $N^{1/2}$, over the Gaussian separation $N^{1/2}$, i.e. $[N^{1/2}]^2/[N^{1/2}]^{d-2}$.

An important element left out of the previous analysis is screening. It is likely that at temperatures lower than T_0 in Eq.(4), the charges are screened over a Debye length $r_D \sim (TR^d/Ne^2)^{1/2}$. Assuming that in this case the entropy is proportional to the number of independent Debye volumes $\mathcal{N} \approx (R/r_D)^d$, we obtain

$$S_{\rm PA}(T) = \ln\left[\exp\left(\omega N - c\left(\frac{e^2 N}{TR^{d-2}}\right)^{d/2}\right)\right]. \tag{9}$$

In this case, freezing (signalled by $S(T_f) = 0$) occurs for an N independent value of $T_f = e^2/a^{d-2} \cdot (c/\omega)^{2/d}$. However, the results of lattice simulations [22] appear to indicate that chains of up to $N = 36$ are still in a short chain regime where screening effects are not important.

V. CONCLUSIONS

We have demonstrated that while neutral PAs form compact globules, a small net charge of $Q_c \propto \sqrt{N}$ is sufficient to cause instabilities in the globular shape. Beyond this threshold, a uniformly charged polymer breaks up into a regular sequence of globules arranged like pearls on a stretched necklace. The inhomogeneities in charge of a random PA are expected to destroy the regularity of the necklace. However, MC simulations indicate that the picture of globules connected by tubes continues to remain applicable.

The experimental signature of the necklace shape is the appearance of multiple length scales in scattering probes of PAs. There is some indication of this happening in the experiments of Ref. [5]. Another intriguing observation is that gels composed of PAs exhibit multiple phases with different volumes [6,24]. A possible explanation of this observation is that the different phases are composed of globules of different sizes at the mesoscopic level, the transitions between them induced by surface tension instabilities similar to those of individual PAs. Again a possible experimental test is to search for inhomogeneities in density at short scales.

Finally, we note that *proteins* are weakly charged PAs: Of the 20 natural amino-acids, three are positively charged (Lys, Arg, His), two are negatively charged (Asp, Glu), and the rest are neutral [4]. It is often assumed that Coulomb interactions are not essential to proteins, as the screening length in biological solvents is often quite small. It is less clear that screening is also effective in compact globular configurations with little or no solvent in their interiors. Secondary structural elements such as α-helices effectively reduce the conformational flexibility of proteins. Thus, while the total charge on a given protein may be small, in solvents with few counter ions, this may be sufficient to lead to a REM-violating correlated energy landscape, making the results of this study relevant.

We can also (indirectly) examine the relevance of Coulomb interactions in protein structure, by searching for correlations of charged amino acids along the tabulated sequences. Firstly, the typical net charge of proteins is less than that expected on the basis of statistical fluctuations. Secondly, protein sequences are anti-correlated with respect to their charge [23]. This indicates that perhaps protein evolution was not dictated solely by the degree of hydrophobicity of monomers (which depends on the degree of charge, not its sign), and that Coulomb interactions also played an important role.

ACKNOWLEDGMENTS

This work was supported by the US–Israel BSF grant No. 92–00026, by the NSF through grant No. DMR–94–00334. Discussions with V.J. Pande and A.Yu. Grosberg are gratefully acknowledged. Many thanks are due to Prof. Miguel Rubi for organizing an interesting and productive conference in the beautiful setting of Sitges.

---

[1] P.G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell Univ. Press, Ithaca (1979).
[2] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond*, World Scientific, Singapore (1987).
[3] D.L. Stein, Proc. Natl. Acad. Sci. USA **82**, 3670 (1985); J.D. Bryngelson and P.G. Wolynes, Proc. Natl. Acad. Sci. USA **84**, 7524 (1987); H.S. Chan and K.A. Dill, Physics Today **46**(2), 24 (1993); T. Garel and H. Orland, Europhys. Lett. **6**, 307 (1988); E.I. Shakhnovich and A.M. Gutin, Europhys. Lett. **8**, 327 (1989); M. Karplus and E.I. Shakhnovich, in *Protein Folding*, ed. by T.E. Creighton, ch.4, p. 127, (Freeman & Co., New York, 1992).
[4] D. Dressler and H. Potter, *Discovering Enzymes*, (Scientific American Library, NY, 1990).
[5] J. Copart and F. Candau, *Macromolecules* **26**, 1333 (1993); M. Scouri, J.P. Munch, S.F. Candau, S. Neyret, and F. Candau, Macromol. **27**, 69 (1994).
[6] X.-H. Yu, A. Tanaka, K. Tanaka, and T. Tanaka, J. Chem. Phys. **97**, 7805 (1992); Yu X.-H., Ph. D. thesis, MIT (1993); A. E. English, S. Mafe, J. A. Manzanares, X.-H. Yu, A. Yu. Grosberg, and T. Tanaka, *J. Chem. Phys* **104**, 8713 (1996).
[7] Y. Kantor and M. Kardar, Europhys. Lett. **14**, 421 (1991).
[8] Y. Kantor, H. Li, and M. Kardar, Phys. Rev. Lett. **69**, 61 (1992); Y. Kantor, M. Kardar, and H. Li, Phys. Rev. **E49**, 1383 (1994).
[9] Y. Kantor and M. Kardar, Europhys. Lett. **27**, 643 (1994); Y. Kantor and M. Kardar, Phys. Rev. **E51**, 1299 (1995).
[10] Y. Kantor and M. Kardar, Phys. Rev. **E52**, 835 (1995).
[11] P.G. Higgs and J.-F. Joanny, J. Chem. Phys. **94**, 1543 (1991); J. Wittmer, A. Johner and J.F. Joanny, Europhys. Lett. **24**, 263 (1993).
[12] Lord Rayleigh, Phil. Mag. **14**, 184 (1882).
[13] G. Taylor, Proc. R. Soc. London **A280** 383 (1964).
[14] J.M. Blatt and V.F. Weisskopf, *Theoretical Nuclear Physics*, ch. 7, p. 303, Willey, New York (1952); R.D. Evans, *The Atomic Nucleus*, ch. 11, p. 387, McGraw–Hill, New York (1955).
[15] N. Bohr and J. A. Wheeler, Phys. Rev. **56**, 426 (1939).
[16] E. Feenberg, Phys. Rev. **55**, 504 (1939).
[17] F. Weizsäcker, Naturwiss. **27**, 133 (1939).
[18] A.V. Dobrynin, S.P. Obukhov, and M. Rubinstein, preprint (1995).
[19] N. Lee and S.P. Obukhov, work in progress (1996); see http://www.phys.ufl.edu/∼nlee/research.html.
[20] Y. Kantor and D. Ertaş, J. Phys. **A27**, L907 (1994).
[21] B. Derrida, Phys. Rev. Lett. **45**, 79 (1980).
[22] V.S. Pande, A.Yu. Grosberg, C. Joerg, M. Kardar, and T. Tanaka, preprint (1996).
[23] V.S. Pande, A.Yu. Grosberg, and T. Tanaka, Proc. Nat. Acad. Sci. USA **91**, 12976 (1994).
[24] M. Annaka and T. Tanaka, Nature **355**, 430 (1992).