

A simple theory and Monte Carlo simulations for recognition between random heteropolymers and disordered surfaces

Arup K. Chakraborty

Department of Chemical Engineering and Department of Chemistry, University of California, Berkeley, California 94720

D. Bratko

Department of Chemical Engineering, University of California, Berkeley, California 94720

(Received 14 July 1997; accepted 16 October 1997)

We study the adsorption of random heteropolymers (RHPs) on disordered multifunctional surfaces. Recent replica calculations suggest that when the statistics describing the sequence distribution and the surface site distributions are related in a special way, a first-order adsorption transition occurs in such systems. This phenomenon of recognition between RHPs and disordered surfaces due to statistical pattern matching is studied via Monte Carlo simulations and a simple nonreplica theory. The theory serves to elucidate the simulation results, and suggests that the origin of the first-order adsorption transition is the suppression of loop fluctuations due to competing interactions and the quenched disorder (i.e., frustration). © 1998 American Institute of Physics.
[S0021-9606(98)51504-8]

I. INTRODUCTION

Polymer adsorption has been studied extensively because of its importance in applications, and interest in understanding the fundamental physics of long polymer chains near surfaces. Most studies, however, have focused on the adsorption of homopolymers and diblock copolymers (for a review, see Ref. 1); in other words, polymers with ordered sequence distributions. In nature, protein adsorption is often associated with specificity or pattern matching. This is to say that particular proteins adsorb on a given substrate and not on others. A question one may ask is: What generic physical features does nature exploit in order to mediate such pattern recognition between polymers and surfaces? There seem to be two main features, multifunctionality and disorder. Proteins are made up of more than one type of monomer, and all coarse-grained models show that the sequence distribution is disordered.^{2,3} Furthermore, the surfaces on which adsorption occurs also usually contain sites of more than one type arranged in an aperiodic manner.

The above observation motivates the following question. Can one take a lesson from nature and exploit multifunctionality and disorder to achieve pattern recognition between polymers and solid surfaces in abiotic applications? In order to address this question, the simplest system that one can study is the adsorption of random heteropolymers (RHPs) on disordered multifunctional surfaces. RHPs are multicomponent polymers with a disordered sequence distribution. Thus, they embody multifunctionality and disorder. The disordered sequence distribution is characterized by a statistics that is ultimately related to the chemical identity of the segments that constitute the chain.⁴ Surfaces that contain multiple types of sites arranged in a disordered manner occur in nature and can also be synthesized (e.g., Refs. 5–7). The disordered arrangement of sites can also be characterized by a statistics. Studying the adsorption of RHPs on such surfaces can therefore address the following intriguing issue. While

evolution has led nature to specific pattern matching, can a sharp adsorption transition (i.e., recognition) occur when the statistics that describe the RHP sequence and the surface sites obey a particular relationship? In other words, is matching statistics enough for recognition?

The physics of RHPs has excited much interest in recent years (e.g., Refs. 8–12). One reason for this is that they constitute a special class of soft matter wherein a quenched disorder (the sequence distribution) is carried by the fluid whose statistical properties are of interest. This feature lends them many interesting properties, including the existence of a transition wherein below a certain temperature the properties are determined by a few dominant conformations.^{8,9} This transition is akin to the protein folding transition. The interfacial behavior of RHPs has also been studied in a few contexts: their behavior near homogeneous surfaces and those with ordered patterns (e.g., Refs. 13–18) as well as their role in strengthening polymer–polymer interfaces^{19,20} have been investigated.

The behavior of homopolymers in disordered media has also been studied extensively following the pioneering work of Muthukumar, Edwards, and Baumgartner (e.g., Refs. 21–23). Muthukumar²⁴ has also considered issues related to pattern matching between polyelectrolytes and surfaces bearing ordered patterns. RHPs in three-dimensional disordered media have also been considered.^{25–27}

Recently, a replica field-theoretic analysis was carried out to examine the simplest possible scenario for RHPs interacting with disordered surfaces.²⁸ RHPs made up of two kinds of units (A and B) were considered to interact with a surface functionalized with two kinds of sites; one type of site preferred to interact with segments of type A, and the other exhibited the opposite preference. The bare RHPs were taken to be Gaussian, and in order to examine the essential physics, the sequence and surface site fluctuations were considered to be Gaussian with short-ranged correlations. A

mean-field theory with a one-step replica symmetry breaking (RSB) scheme showed that indeed a sharp adsorption transition occurred when the product of the widths of the two statistical distributions acquired a threshold value. In other words, recognition occurs when the statistics of the two distributions are related in a special way.

In order to investigate this phenomenon further, we have carried out extensive Monte Carlo (MC) simulation studies.^{29,30} The simulation studies have been carried out using the nondynamic ensemble growth method. The results show that the basic physics revealed by the replica mean-field theory is correct. When the statistics of the RHP sequences and the distribution of sites are related in a special way, there is a sharp adsorption transition (the hallmark of recognition). The MC simulation results have been carried out for situations far more complex and interesting than the simplest scenario studied via replica field theory. Correlated fluctuations in sequence and surface site distribution statistics have been studied, and RHPs have been shown to be able to discriminate between surfaces with different statistical patterns (i.e., recognition).³⁰

In this paper, however, we focus on some details of the physics revealed by the simulation studies. In order to do so, we will restrict attention to a few simulation results for the case wherein both sequence and surface site fluctuations exhibit short-ranged correlations, and the composition is symmetric. These results suggest that while the replica field theory²⁸ revealed the phenomenon of recognition due to statistical pattern matching correctly, it may be lacking in detail. Here we develop a simple nonreplica model that seems consistent with some of the details of the simulation results. This simple theory is physically more transparent than the replica theory, and points to the essential origin of why the adsorption transition is sharp in this case (recall that adsorption transitions are continuous for polymers with ordered sequence distributions). The theory is inspired by a similar model developed by Pande *et al.*³¹ to describe protein folding. We hope that this simple theory can be employed to guide experiments.

The rest of this paper is organized as follows. In Sec. II, we briefly outline the simulation results that prompt the development of the simple model. In Sec. III the model is developed, and this is followed by a discussion in Sec. IV. Brief concluding remarks are offered in Sec. V.

II. OUTLINE OF SIMULATION RESULTS

The simulations have been performed on a cubic lattice using the ensemble growth MC method.^{26,32–34} A particular sequence is first drawn from the statistical distribution under consideration. M monomers are then placed randomly with Boltzmann statistics dictating the positional probabilities. We then attempt to add a segment of type A or B (as specified by the particular sequence that has been chosen) at the end of each monomer. Of these $6M$ trials, M dimers are chosen with Boltzmann probabilities. The potential energy of each conformation is determined by intersegment interactions and interactions with sites that comprise the disordered

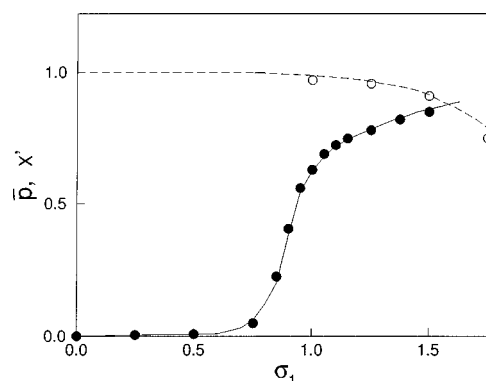


FIG. 1. MC simulation results for adsorbed fraction (\bar{p}), filled circles, and x (defined in the text), open circles as a function of loading, σ_1 .

surface. For the results that we show here, in the absence of the surface, the intersegment interactions are merely excluded volume constraints. The process of growing chains is continued until chains with desired length (N) are obtained. For $M \gg N$, a canonical distribution results, and properties are computed as nonweighted averages over the ensemble. If necessary for accuracy, many populations of the same sequence are simulated. The same procedure is then repeated for several realizations of the quenched sequence distribution. Averages over different realizations of the disordered sites are computed in two ways; by employing analytic preaveraging over the surface disorder or by explicit quenched averages over disorder realizations. Both approaches lead to statistically equivalent results when sufficiently large surfaces are simulated. Simulations have been carried out for $N = 32, 64$ and 128 ; most detailed results exist for $N = 32$. Typically, M equals $64\,000$ in our simulations.

Consider RHPs with two types of segments (say, A and B) interacting with a surface bearing two types of sites with A preferring one type of site and B the other. Let the RHPs have fixed sequence fluctuation statistics. For sequence fluctuations with short-ranged correlations this is tantamount to fixing the width of the distribution (or average composition). The statistics of the surface sites is also measured by the width when the distribution exhibits fluctuations with short-ranged correlations, with the variance σ_1 being directly proportional to the loading of surface sites. By loading, we mean the fraction of the surface area covered by interacting sites. Let us examine the adsorbed fraction of RHPs with fixed sequence statistics as a function of the loading (measure of the surface site distribution statistics). Note that the loading can be experimentally manipulated (e.g., Refs. 5 and 6).

Figure 1 depicts the pertinent simulation results. We find that when the loading exceeds a threshold value, a sharp adsorption transition occurs. In MC simulations, we can measure a quantity $x = 1 - \sum_i P_i^2$, where P_i is the probability associated with conformation i . When a multitude of surface conformations are sampled, x is asymptotically equal to unity. However, if only a few dominant conformations prevail, x acquires values less than unity. Thus, when $x < 1$, the thermodynamics is determined by a few dominant conforma-

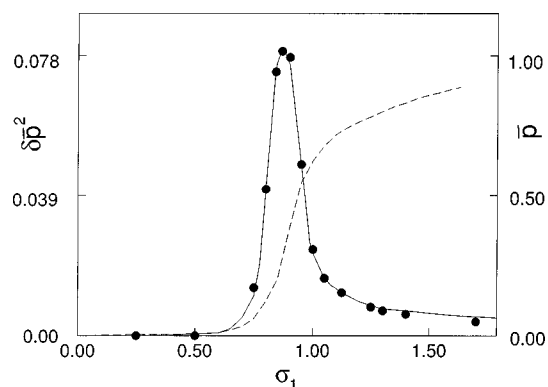


FIG. 2. MC simulation results for the fluctuations in adsorbed fraction, $\langle(\delta\bar{p})^2\rangle$, as a function of surface loading, σ_1 . Also included is the adsorbed fraction (\bar{p}), dashed line.

tions. Figure 1 shows that the sharp adsorption transition is accompanied by x acquiring values less than unity. These simulation results are consistent with a previous replica field theory that predicted a first-order adsorption transition accompanied by replica symmetry breaking (i.e., freezing into a few dominant conformations).

The physical reason underlying this adsorption transition appears to be the following. When the statistics of the RHP sequence fluctuations and that of the surface site distribution are related in a particular way (*vide infra*), a few pattern-matched conformations become energetically much more favorable than all others. By pattern matched, we mean a registry between segments and the type of sites that they prefer. Thus, at this point the RHPs give up entropy in order to take advantage of the substantial energy gain associated with sampling a few pattern matched adsorbed conformations. This leads to the sharp adsorption transition, which mimics recognition when the statistics are matched. The replica theory²⁸ and the simulations, however, do not provide clear physical insights regarding the phenomenon (especially, on the issue of why the transition is sharp).

While in broad terms the exciting physics revealed by the replica theory²⁸ is borne out by our MC simulation results, a detailed examination indicates that some features are not captured by the replica theory. Consider, for example, the simulation results shown in Figs. 2 and 3. Here the fluctuations in adsorbed fraction and in the energy are shown superimposed on the profile for p . We find that $\langle(\delta p)^2\rangle$ exhibits one peak coinciding with the sharp adsorption transition. This peak grows in size and narrows as N is increased from 32 to 64 and 128. This is consistent with a first-order transition in finite size simulations (we have deemed three chain lengths inadequate for estimating and reporting a finite size scaling exponent). Interestingly, however, the energy fluctuations exhibit two peaks; one that coincides with the sharp adsorption transition, and the other that occurs at higher values of σ_1 . The replica theory does not predict the second peak for the energy fluctuations. Closer examination of the simulation results also shows that x acquires values less than unity *after* the sharp adsorption transition. Further, the simulation results also show that fluctuations become

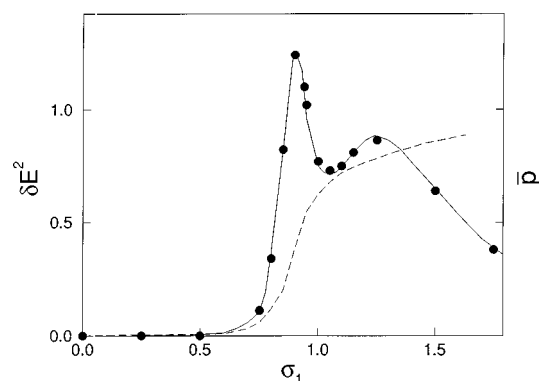


FIG. 3. MC simulation results for the fluctuations in energy, $\langle(\delta E)^2\rangle$, as a function of surface loading, σ_1 . Also included is the adsorbed fraction (\bar{p}), dashed line.

very small (i.e., very few dominant conformations are adopted) only after the adsorption transition occurs.

The aforementioned details that do not seem to be explained by the replica theory motivate the development of a complementary theoretical model that is described in Sec. III. The model is conceptually very simple, and sheds light on the essential physics. It predicts the phenomenon of statistical pattern matching first revealed by the replica theory, and also seems to explain some of the features noted above.

III. MODEL DEVELOPMENT AND THEORY

Again, we consider the interaction of a RHP chain made up of A and B segments interacting with a surface functionalized with two type of sites which are such that the A segments prefer to interact with one type of site more than the other, while the opposite is true for the segments of type B. Therefore, if a segment adsorbs, there are two possible types of contacts; good contacts (those that find a preferred type of site) and bad contacts (those that do not adsorb on a preferred type of site). In order to formulate a simple theory, we now develop an expression for the free energy density as a function of the coarse-grained order parameters that measure the total number of sorbed segments and the number of good contacts. Minimization of this free energy density will yield a mean-field theory.

Consider the energy first. In the spirit of examining the simplest model, we consider the chain to be noninteracting in the absence of the surface. This allows us to examine the basic physics of RHPs interacting with disordered multifunctional surfaces. Also, it can be shown that excluded volume interactions do not change the qualitative physics;³⁵ our MC simulations enforce excluded volume constraints, and the same physics is observed. If distinct types of intersegment interactions are included,³⁵ then the competition between intersegment interactions and segment-surface interactions leads to richer behavior (including the sharp adsorption transition, however). For the bare chains being Gaussian, the energy for a given number of total contacts with the surface (p) and the number of good contacts (q) is

$$\frac{E}{kT} = [q\delta E + pE_1], \quad (1)$$

where E is the energy, δE is the energy difference between good contacts and others, and E_1 is the energy associated with making bad contacts.

In order to obtain the free energy we now need to compute the entropy for a chain of length N with p adsorbed segments, of which q are good contacts. This entropy is a sum of three contributions that stem from different physical effects. First, there is a ‘‘mixing’’ entropy (S_{mix}) associated with alternative ways of choosing p adsorbed segments out of a total of N . Second, there is the entropy loss (S_{ads}) associated with the adsorption of a chain segment. Finally, there is the crucial part of the entropy associated with the formation of loops (S_{loop}). Formally, we may write

$$S(p, q, N) = S_{\text{mix}} + S_{\text{ads}} + S_{\text{loop}}. \quad (2)$$

For S_{mix} , we make the simplest possible estimate, i.e.,

$$\frac{S_{\text{mix}}}{N} = -\bar{p} \ln \bar{p} - (1 - \bar{p}) \ln (1 - \bar{p}), \quad (3)$$

where \bar{p} is the fraction of adsorbed segments. Equation (3) does not account properly for chain connectivity, and hence overestimates the mixing entropy that we seek. However, as in all polymer problems, the important physics is contained in the loop entropy, and so we take this simple form for S_{mix} .

The entropy loss associated with segmental adsorption is well-known,³⁶ and we write it in the usual way:

$$\frac{S_{\text{ads}}}{N} = -\omega \bar{p}, \quad (4)$$

where ω is a constant related to chain flexibility and solution conditions.

We now consider the most important part of the entropy calculation, viz. the computation of loop entropy. In our problem, there are two types of loops. There are those loops that are associated with forming good contacts at both ends, and those that are associated with forming the other contacts. These two types of loops are fundamentally different in character. For loops associated with good contacts *only certain loop lengths and distances (on the two-dimensional surface) between the ends of the loops are allowed*. These lengths and distances are determined by the statistics of the chain sequence and surface site distributions. *Thus, because of competing interactions and quenched disorder, loop fluctuations that are characteristic of unfrustrated polymer adsorption are suppressed*. Loops associated with contacts that are not good are not restricted in this manner, and the usual fluctuations in loop length and distance between loop ends on the surface occur. For ease of reference, we shall label these two types of loops as quenched and annealed loops, respectively.

We need to compute the entropy associated with both these types of loops when they coexist. In order to do so, the first step is to be able to compute the entropy associated with each type of loop when it is the only type of loop that exists.

Once these formulas are available, as we shall see, it is relatively straightforward to combine them properly to reflect the physics under consideration.

Let us begin with the quenched loops. The loop factor associated with a loop of length n returning to the plane with the two ends separated by a distance d is

$$P(n, d) = \frac{C}{n^{3/2}} \exp\left[-\frac{d^2}{2nb^2}\right], \quad (5)$$

where C is a constant related to chain flexibility, and b is the statistical segment length. We have written Eq. (5) for the bare chains being Gaussian since we are considering such a situation in order to describe the essential physics. The loop factor for quenched loops depends upon the statistics of the sequence and surface site distributions. For example, for the simplest case of surface and site fluctuations with short-ranged correlations, the way in which d depends upon the statistics of the disordered surface is: $d^2 \sim 1/\sigma_1^2$, where σ_1^2 is the width of the distribution (simply related to the average loading). If there are q quenched loops, then the average length of the quenched loops is N/q . Later, we shall see that the number of quenched loops, and hence n , will depend on the statistics of the RHP sequence distribution. For now, for q quenched loops, using Eq. (5) and the considerations noted above, we can write down the entropy to be

$$\frac{S_{\text{quen loops}}}{N} = \bar{q} \ln(C\bar{q}^{-3/2}) - \left[\frac{a\bar{q}^{-2}}{2b^2\sigma_1^2}\right], \quad (6)$$

where $\bar{q} = q/N$, and a is a constant.

Now consider the entropy associated with forming annealed loops only. We note again that we shall consider the problem of simultaneously forming both quenched and annealed loops after we derive formulas for the entropic factors for each type of loop. In order to compute the entropy associated with forming annealed loops only, consider the well-known problem of a homopolymer interacting with an attractive surface with an energy bonus associated with segmental adsorption. Thus, we consider a potential $\Phi(z)$ that is zero everywhere except at the surface; z is the coordinate normal to the attractive surface. The effective energy bonus associated with segmental adsorption is then calculated in the usual way as

$$\beta = \int_0^\infty dz [e^{-\phi(z)/kT} - 1], \quad (7)$$

where $\ln \beta$ is the effective energy bonus.

In this problem, the loop lengths and distances between loop ends can fluctuate with only the constraint of a fixed total chain length. Thus, it is most convenient to work in the grand canonical ensemble. In order to obtain the entropy corresponding to a given number of annealed loops, we follow a method adapted from Ref. 37. In this reference the entropy for chain adsorption on to a three-dimensional point potential has been computed. Adapting this method, and exploiting the symmetries in our problem, we find that the Schrödinger-type equation corresponding to the problem under consideration is

$$[1 + \beta \delta(z)] \hat{g} \psi(z) = \lambda \psi(z), \quad (8)$$

where \hat{g} is the standard connectivity operator, and the eigenfunction (ψ) and the eigenvalue (λ) have their usual meaning. Upon Fourier transformation, very simple manipulations (see Ref. 37) of Eq. (8) lead us to the following relationship between β and λ :

$$\frac{1}{\beta} = \frac{1}{2\pi} \int dk \frac{g(k)}{[\lambda - g(k)]}, \quad (9)$$

where k is the Fourier variable conjugate to z .

Our strategy now is the following. We will first compute the equilibrium value of the number of annealed loops (P) for a given energy bonus per segment, $\ln \beta$. Then, we will compute the corresponding entropy, thereby obtaining the entropy as a function of the number of annealed loops. This is the relationship that we seek.

Let $Z(N, P)$ be the partition function for a chain of length N with P annealed loops. Then, we can define a generating function, $z(N, \beta)$ in the usual way

$$z(N, \beta) = \sum_P \beta^P Z(N, P). \quad (10)$$

For the situation at hand, ground state dominance is appropriate, and so using a saddle point approximation, substituting Eq. (9) for β , and noting that $\lambda^N = z(N, \beta)$, we obtain the following relationship for the equilibrium value of the adsorbed fraction P/N for the annealed situation we are considering:

$$\frac{P}{N} = \frac{1}{\lambda} \frac{\int dk g(k)/[\lambda - g(k)]}{\int dk g(k)/[\lambda - g(k)]^2} \quad (11)$$

Now we are ready to compute the entropy. Noting that the free energy, F , equals $-N \ln \lambda$, and that $\ln \beta$ is the energy bonus for adsorbing a segment, we obtain the entropy (S) to be

$$\frac{S}{N} = -\frac{F}{NT} + \frac{E}{NT} = \ln \lambda - \frac{P}{N} \ln \left[\frac{1}{2\pi} \int dk \frac{g(k)}{[\lambda - g(k)]} \right]. \quad (12)$$

Equations (9), (11), and (12) can be solved simultaneously to establish the relationship between the entropy and the number of contacts P for annealed loops.

Having computed the entropy of quenched and annealed loops separately, we are now in a position to compute the loop entropy when both types of contacts between RHP chain segments and surface sites occur.

On physical grounds, it is clear that the total number of contacts with the surface must be greater than or equal to the number of good contacts. This, in turn, implies that the annealed loops live within quenched ones. Since the redistribution of annealed loops among quenched ones must be unconstrained, the chemical potential of annealed loops must be the same in each quenched loop. Of course, the chemical potential in question is easily computed from Eqs. (9) to (12) as it equals $-\partial S/\partial P$. One finds that the chemical potential is a monotonic function of P . This fact combined with our

observation that the chemical potential for annealed loops must be the same in each quenched loop implies that the concentration of annealed loops, $\bar{p}_a = (\text{No. of annealed loops}/\text{length of the quenched loop})$ must be the same in all quenched loops. Note that if \bar{p} is the total adsorbed fraction and \bar{q} the fraction of good contacts, then $\bar{p}_a = \bar{p} - \bar{q}$.

The above remarks allow us to write down the loop factor for quenched loops, with each quenched loop containing the appropriate concentration of annealed loops. Thus, we can compute the entropy for quenched and annealed loops as a function of \bar{p} and \bar{q} . We find

$$\frac{S_{\text{loops}}}{N} = \bar{q} \left[\ln P(1/\bar{q}, d) + \frac{1}{\bar{q}} s(\bar{p}_a) \right], \quad (13)$$

where P is the probability written down in Eqs. (5) and (6), and $s(\bar{p}_a)$ is the entropy of annealed loops with concentration \bar{p}_a divided by n . $s(\bar{p}_a)$ is obtained from Eqs. (9) to (12) with one minor modification. These equations were derived assuming that the surface was uniform. In the physical situation at hand, there is an entropy loss associated with the fact that adsorption can only occur on the sites on the surface. Thus, we account for an entropy reduction that is related to the total surface loading (proportional to $\ln \zeta_1^2$).

Combining Eqs. (1)–(4), and (13) we obtain the free energy density, f , as a function of p and q ,

$$f = \bar{q} \delta E + (\omega + E_1) \bar{p} + \bar{p} \ln \bar{p} + (1 - \bar{p}) \ln(1 - \bar{p}) - \bar{q} \ln(C \bar{q}^{3/2}) + \frac{a \bar{q}^2}{2 \sigma_1^2 b^2} - s(\bar{p} - \bar{q}) - (\bar{p} - \bar{q}) \ln \left[\sigma_1^2 \frac{b^2}{a} \right]. \quad (14)$$

However, \bar{p} and \bar{q} are related by the fact that the number of good contacts equal the total number of contacts multiplied by the probability of finding a good contact. This is to say that

$$\bar{q} = \bar{p} P_g, \quad (15)$$

where P_g is the probability of finding a good contact. At infinite temperature, P_g is simply related to the statistics of the sequence and surface site distributions. For example, in the case of uncorrelated fluctuations in sequence and site distributions, it is simply proportional to the width of the two Gaussian distributions. Simple combinatorics can be used to determine similar relationships for correlated fluctuations. (A relation for the probability of finding good contacts at infinite temperature that seems to describe all qualitative aspects of the simulation results is: $\sum_m P_c(m) P_s(m)$, where P_c and P_s are the probabilities of finding a block of like segments of length m and a patch of size m on the surface, respectively.) At finite temperatures, such relationships must be weighted by the free energy associated with making good contacts. This is to say that

$$P_g = P'_g \frac{e^{-\beta F_q}}{P'_g e^{-\beta F_q} + (1 - P'_g) e^{-\beta F_a}}, \quad (16)$$

where P'_g is the probability of making good contacts at infinite temperature, F_a and F_q are the free energies associated

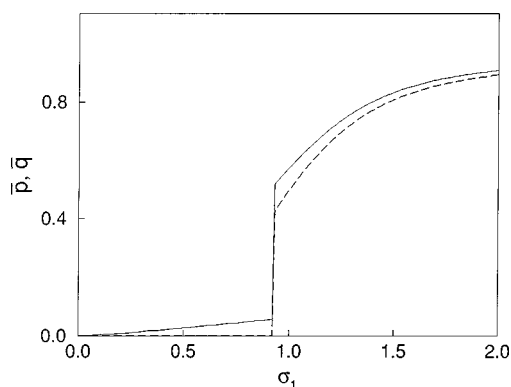


FIG. 4. Theoretical predictions for total adsorbed fraction (\bar{p}), solid line, and the fraction of good contacts (\bar{q}), dashed line, as a function of surface loading, σ_1 . For illustrative purposes, the other parameters are chosen to be: $\delta E = -2$; $\omega + E_1 = 1$; $C = 2$; $a/b^2 = 4$; $P'_g = 1/2$.

with annealed and quenched contacts, respectively. The above discussion makes clear that the length of the quenched loops ($1/\bar{q}$) depends upon the statistics of both the chain sequence and the surface site distribution.

In order to obtain \bar{p} and \bar{q} for a given set of conditions, we minimize the free energy density [Eq. (14)] with respect to \bar{p} [note that \bar{p} and \bar{q} are related via Eq. (15)], with the minimum finite number of annealed loops per quenched loop being one. The procedure has to be carried out iteratively since P_g is determined by Eq. (16). This simple numerical solution of our equations can be carried out using a rudimentary computing machine. In Fig. 4, we present the results for fixed RHP sequence statistics, and varying surface statistics. For illustration of the qualitative aspects, we have chosen to depict the situation corresponding to uncorrelated fluctuations in sequence and surface site distributions. Thus, we plot \bar{p} and \bar{q} as a function of σ_1 (for fixed RHP sequence statistics). Note, however, that the theory we have developed is general, and is not restricted to the special case of short-range correlations in sequence and surface site fluctuations.

IV. DISCUSSION

Figure 4 shows that for small values of σ_1 both \bar{p} and \bar{q} are zero. As the loading increases, we find a continuous adsorption transition (finite \bar{p}), but the number of good contacts remains negligible. Then, at a threshold value of σ_1 , the theory predicts a first-order adsorption transition in both \bar{p} and \bar{q} . For the case of short-ranged sequence and surface site distribution fluctuations, the average composition f and σ_1 are measures of the statistical patterns carried by the chains and the surface. Thus, Fig. 4 shows that for fixed sequence statistics (f), a sharp adsorption transition occurs when the surface site statistics are of a particular type (σ_1 acquires threshold value), i.e., we have recognition due to statistical pattern matching. As noted in Sec. I, simulations show³⁰ this phenomenon for more elaborate statistical patterns than the one considered here to illustrate the essential physics. In the context of the theory developed in this paper, the sharp adsorption transition occurs when $P'_g = \sum_m P_c(m)P_s(m)$ acquires a threshold value, i.e., the RHP sequence and surface

site statistics are related in a special way. The quantities, $P_c(m)$ and $P_s(m)$, which have been defined earlier, can be easily computed from the sequence statistics (related to RHP chemistry and synthetic route) and the surface site distribution statistics. After the first-order adsorption transition, \bar{q} still remains smaller than \bar{p} ; however, at this stage a preponderance of the contacts are good ones. The theory predicts that this first-order adsorption transition is followed by a continuous transition wherein \bar{q} and \bar{p} approach each other.

Let us consider these results in the context of the simulation results, the replica field theory, and a related problem (protein folding). The prediction of a sharp adsorption transition when the loading acquires a threshold value is in harmony with both the replica field theory and our simulation results. This first-order adsorption transition seems analogous to the “coil-globule” transition in proteins. The coil to globule transition in proteins (and heteropolymeric models of the phenomenon) is one wherein most of the contacts become native ones, but the protein is not “folded” into its final low entropy native state. The transition is first order for RHPs. For a very different system, we find related physics, i.e., when the statistics are matched, a sharp transition occurs wherein a large number of good contacts are established. Our theory predicts a second continuous transition where the fraction of good contacts approaches the total adsorbed fraction. This second transition may be viewed to be the freezing transition in the replica theory and analogous to the folding transition of proteins (that follows the coil-globule transition). In the replica theory, the first-order adsorption transition and freezing into a few dominant adsorbed conformations occurs simultaneously. However, the MC simulations indicate that the freezing transition occurs after the adsorption transition (*vide supra*). Thus, the simple theory we have presented seems to be consistent with the simulation results in this regard. It is not possible at this time to determine the order of the second transition from simulation data, and hence we cannot make a meaningful comparison with the prediction that the second transition is continuous. This is to say that a detailed finite size scaling study of the peaks for $(\delta\bar{p})^2$ and $(\delta E)^2$ has not been carried out. As noted earlier, the results do show, however, that the peaks narrow as the system size increases. The existence of two transitions, however, does explain why the simulation results show two peaks for the energy fluctuations and only one for $(\delta\bar{p})^2$. As we see from Fig. 4, there is only one adsorption transition, so there should be only one peak in the fluctuations associated with \bar{p} . During the second transition, the major effect is the change in \bar{q} , i.e., the change to essentially all good contacts and freezing into very few dominant conformations. The energy fluctuations would be sensitive to such a transition as well as the adsorption transition.

The above arguments and the simulation results suggest that the replica field theory may not be correct in predicting that the adsorption and freezing transitions occur simultaneously. While we are unable to say anything precise at this time, we speculate that the origin of this prediction may be related to the one-step replica symmetry breaking scheme that was employed in Ref. 28.

Our (Refs. 28 and 30, and this paper) most significant result is that when the statistics of the sequence and surface site distributions are related in a special way (*vide supra*), a sharp adsorption transition (mimicking recognition) occurs. The simple model developed in this paper provides some physical insight into why this transition is predicted to be first order. The physical reason is the suppression of loop fluctuations due to the fact that for the problem at hand only certain loop lengths are allowed when good contacts (which are favored) are made. This is a direct consequence of the fact that we are dealing with a disordered system with competing interactions. In our theory, this frustration manifests itself as the quenched loop entropy. We find that if the quenched nature of the loops resulting from good contacts is "turned off" in the theory, the adsorption transition becomes continuous. This offers support for the physical reason noted above. Rigorous proof of this is expected to be provided by ongoing renormalization group calculations.

V. CONCLUDING REMARKS

MC simulations show that when RHPs interact with disordered multifunctional surfaces, a sharp adsorption transition occurs when the statistics of the RHP sequence and the surface site distribution are related in a particular way. This sharp transition is the hallmark of recognition, and suggests that statistical pattern matching is sufficient to mimic recognition. MC simulations not discussed in this paper also show that this phenomenon can be exploited to discriminate between different types of statistical patterns.³⁰ In this paper, we have developed a simple physically transparent model for RHPs near disordered surfaces that does not employ replicas. The model suggests that the physical origin of the first order adsorption transition is the suppression of loop fluctuations due to disorder and competing interactions.

ACKNOWLEDGMENTS

The nonreplica model described in this paper was inspired by a talk by Dr. V. S. Pande where a similar model was developed for describing protein folding. Subsequent discussions with Dr. Pande, Professor Eugene Shakhnovich, and Simcha Srebnik are also gratefully acknowledged. Financial support for this work was provided by NSF and the U.S.D.O.E., Basic Energy Sciences.

- ¹A. K. Chakraborty and M. Tirrell, *MRS Bull.* **21**, 28 (1996).
- ²V. S. Pande, A. Yu. Grosberg, and T. Tanaka, *Proc. Natl. Acad. Sci. USA* **91**, 12972 (1994).
- ³A. Irback, C. Peterson, and F. Potthast, *Proc. Natl. Acad. Sci. USA* **93**, 9533 (1996).
- ⁴G. G. Odian, *Principles of Polymerization* (Wiley, New York, 1991).
- ⁵R. J. Todd, R. D. Johnson, and F. H. Arnold, *J. Chromatography A* **662**, 13 (1994); R. D. Johnson, Z. G. Wang, and F. H. Arnold, *J. Phys. Chem.* **100**, 5134 (1996).
- ⁶P. Berndt, G. B. Fields, and M. Tirrell, *J. Am. Chem. Soc.* **117**, 9515 (1995).
- ⁷M. Mammen, G. Dahmann, and G. M. Whitesides, *J. Med. Chem.* **38**, 4179 (1995).
- ⁸E. I. Shakhnovich and A. M. Gutin, *Nature (London)* **346**, 773 (1990).
- ⁹P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, *Science* **267**, 1619 (1995).
- ¹⁰A. Sali, E. I. Shakhnovich, and M. Karplus, *Nature (London)* **369**, 248 (1994).
- ¹¹H. S. Chan and K. A. Dill, *Phys. Today* **46**, 24 (1993).
- ¹²H. Frauenfelder and P. G. Wolynes, *Phys. Today* **47**, 57 (1994).
- ¹³L. Gutman and A. K. Chakraborty, *J. Chem. Phys.* **101**, 10074 (1994).
- ¹⁴L. Gutman and A. K. Chakraborty, *J. Chem. Phys.* **104**, 7306 (1996).
- ¹⁵J. F. Joanny, *J. Phys. II* **4**, 1281 (1994).
- ¹⁶L. Gutman and A. K. Chakraborty, *J. Chem. Phys.* **103**, 10738 (1995).
- ¹⁷A. C. Balazs, M. C. Gempe, and Z. Zhou, *Macromolecules* **24**, 4918 (1991).
- ¹⁸D. Gersappe and A. C. Balazs, *Phys. Rev. E* **52**, 5061 (1995).
- ¹⁹S. T. Milner and G. H. Fredrickson, *Macromolecules* **28**, 7953 (1995).
- ²⁰C. A. Dai, B. J. Dair, K. H. Dai, C. K. Ober, and E. J. Kramer, *Phys. Rev. Lett.* **73**, 2472 (1994).
- ²¹A. Baumgartner and M. Muthukumar, *J. Chem. Phys.* **87**, 3082 (1987).
- ²²S. F. Edwards and M. Muthukumar, *J. Chem. Phys.* **89**, 2435 (1988).
- ²³K. Sumithra and K. L. Sebastian, *J. Phys. Chem.* **98**, 9312 (1994).
- ²⁴M. Muthukumar, *J. Chem. Phys.* **103**, 4723 (1995).
- ²⁵A. K. Chakraborty and E. I. Shakhnovich, *J. Chem. Phys.* **103**, 10751 (1995).
- ²⁶D. Bratko, A. K. Chakraborty, and E. I. Shakhnovich, *Phys. Rev. Lett.* **76**, 1844 (1996).
- ²⁷V. S. Pande, A. Yu. Grosberg, and T. Tanaka, *J. Phys. A* **28**, 3657 (1995).
- ²⁸S. Srebnik, A. K. Chakraborty, and E. I. Shakhnovich, *Phys. Rev. Lett.* **77**, 3157 (1996).
- ²⁹D. Bratko, A. K. Chakraborty, and E. I. Shakhnovich, *Comp. Theor. Polym. Sci.* (submitted).
- ³⁰D. Bratko, A. K. Chakraborty, and E. I. Shakhnovich, *Chem. Phys. Lett.* (in press).
- ³¹V. S. Pande, A. Yu. Grosberg, and T. Tanaka, *Phys. Rev. Lett.* (submitted).
- ³²P. G. Higgs, H. Orland, *J. Chem. Phys.* **95**, 4506 (1991).
- ³³D. Bratko, A. K. Chakraborty, and E. I. Shakhnovich, *J. Chem. Phys.* **106**, 2064 (1997).
- ³⁴T. Garel and H. Orland, *J. Phys. A* **23**, L621 (1990).
- ³⁵S. Srebnik and A. K. Chakraborty (unpublished).
- ³⁶C. A. J. Hoeve, E. A. diMarzio, and P. Peyser, *J. Chem. Phys.* **42**, 2558 (1965).
- ³⁷A. Yu. Grosberg and A. R. Khoklov, *Statistical Physics of Macromolecules* (AIP Press, New York, 1994).