Scaling properties of maximally compact chains

K. W. Foreman*

*The Department of Pharmaceutical Chemistry, University of California, San Francisco, 3333 California Avenue,
San Francisco, California 94118*

Scaling of the exact function for the number of intramolecular nonbonded contacts in a single maximally compact linear homopolymer on hypercubic lattices is determined as a function of number N of monomers and dimension d. A representative maximally compact structure is designed and an exact recursive expression for the maximum number m_{max} of contacts is derived from that design. The equivalent nonrecursive expression yields the asymptotic scaling of m_{max} as $(d-1)N - dN^{\Delta} + 1$, with $\Delta = (d-1)/d$. Implications in polymer and protein studies are discussed. [S1063-651X(99)12405-X]

PACS number(s): 05.50.+q

I. INTRODUCTION

Understanding the fundamental properties of macromolecules has led to advances in catalysis, drug design, and even manufacturing. Advances in these areas have been aided in part by lattice models, whose simple microscopic descriptions facilitate connections between molecular structure and fundamental physical properties [1,2]. For example, structural information about the equilibrium conformations of single lattice chains (representing proteins) can be extracted, at worst, by exhaustive enumeration. The lattice representation of chains reduces the continuum of conformations down to an enumerable number of conformations, with the lowest-energy conformations readily identifiable. These lowest energy conformations are typically assumed to represent the equilibrium state conformations under chain collapse conditions. Properties of pathways to these lowest energy conformations, including the location of kinetic traps, can be discerned [3] and the thermodynamics of the system can be calculated exactly [4]. Despite the multiple uses of lattice models, the physics of single lattice chains is still incompletely understood [2,5–11].

One of the unsolved problems involving single lattice chains is how many contacts are present in structures with the maximum number m_{max} of contacts and how this number scales with both the number N of monomers and the dimension d of the lattice. The solution to this problem has multiple applications. For example, the exact form for m_{max} may help determine the scaling of the free energy for compact lattice chains [12] and serves as an upper bound on the maximum number of contacts possible in random copolymers, although tight bounds have been independently developed for this case [13]. Further, knowing the number of contacts in the most compact state helps determine the speed and frequency with which equilibrium conformations of polymers in poor solvents are sampled by Monte Carlo techniques.

The exact formulas and asymptotic scaling properties of m_{max}, however, are incompletely determined. Chan and Dill [14,15] postulate formulas for m_{max} when $d=1-3$ on cubic

lattices, but the scaling properties are not readily extractable. On the other hand, Douglas and Ishinabe [5] propose an asymptotic scaling of m_{max} with N and d. In developing this scaling, they define an optimized walk on hypercubic lattices as a representative self-avoiding walk (SAW) (i.e., a representative lattice chain) which makes m_{max} nonbonded nearest-neighbor contacts for a linear homopolymer chain containing N monomers in d dimensions. Proving the existence of these optimized walks is tantamount to determining the exact form for m_{max}, and hence its scaling properties with N and d. Douglas and Ishinabe derive an exact expression for m_{max} in $d=2$, and a tight upper bound in $d=3$. Nevertheless, the differences between the exact expression in $d=2$ and their general asymptotic expression makes confirmation of their postulated scaling difficult. Therefore, developing an asymptotic scaling function for m_{max} requires systematically constructing an optimized walk.

This work constructs an optimized walk in arbitrary dimensions and determines the asymptotic scaling behavior of m_{max}. Section II describes the construction of an optimized walk. Section II A establishes the fact that hypercubic conformations and certain hyperrectangular conformations must be obtained by an optimized walk. In establishing these conformations, tools are developed which severely restrict the number of alternative conformations to an optimized walk. Section II B details a putative construction of the optimized walk (the cross-back spiral) and Sec. II C proves that the construction in Sec. II B is valid by determining that the few possible alternative structures have at best an equal number of contacts. Section III provides recursive (Sec. III A) and nonrecursive (Sec. III B) forms for m_{max} based on the construction in Sec. II. These results confirm the previously suggested [5] scaling of m_{max} with N and d, as well as the formulas for m_{max} in $d=1-3$ dimensions [14,15].

II. THE ''OPTIMIZED'' HAMILTON WALK

Many polymer simulations and models represent polymer chains as self-avoiding walks. The chain is constructed in a stepwise fashion. One end is placed, without loss of generality, at the origin. A step, usually of fixed length, is taken in an arbitrary direction, with the end point of the step demark-

*FAX: (415)502-4222. Electronic address: kw@maxwell.ucsf.edu

ing the location of the next monomer in the polymer (provided the end point does not coincide with a previously placed monomer). This step and place routine is repeated until all monomers are assigned locations. For the lattice model considered here, each step length is exactly the distance between nearest-neighbor lattice sites and each step can fall only along one of z lattice directions. Thus, every monomer occupies a lattice site, with consecutive monomers in the polymer lying on adjacent lattice sites. To make the walk self-avoiding, each lattice site contains at most one monomer. This paper considers only hypercubic lattices, in which the number z of nearest neighbors is twice the dimension d. Two nonbonded monomers form a contact when they are nearest neighbors. Although bonded monomers are adjacent, they are excluded when counting contacts since any chain configuration will always incur these ''contacts.'' Thus, the terminal and interior monomers in a chain on a hypercubic lattice make at most $z-1$ and $z-2$ contacts, respectively (a contact in every direction except along bonded ''contacts'').

The optimized walk is defined [5] as a representative SAW which, for any number N of monomers in d dimensions, makes the maximum number $m_{max}(N,d)$ of contacts for a (linear) lattice polymer chain. This definition implies that for all optimized walks in dimension d with $N \geq k > 0$ monomers, the first k monomers are always placed in *exactly* the same configuration. For example, in two dimensions, a spiral is an optimized walk [see Fig. 1(a)]. The construction below systematically creates an optimized walk and derives an exact formula for $m_{max}(N,d)$.

A. Required chain configurations

Before constructing an optimized walk, it is helpful to consider which configurations that walk must pass through. Douglas and Ishinabe [5] find that an optimized walk with $N = n^d$ (n, an integer) monomers fills a d-dimensional hypercube with side of length n monomers. This section establishes their finding as a special case of the proof that an optimized walk with $N = n^{d-i}(n+1)^i$ ($0 \leq i < d$) monomers must fill a hyperrectangle with $d-i$ sides of length n monomers and with i sides of length $n+1$ monomers.

This proof is simplified by a lemma that establishes the relationship between the maximum number of contacts possible for a configuration and the minimal bounding hyperrectangle for that configuration. A minimal bounding hyperrectangle is the smallest hyperrectangle which completely contains a configuration. These hyperrectangles must exist for configurations with a finite number of monomers since the configuration has finite extent in each of the d dimensions. Intuitively, minimal bounding hyperrectangles with larger volumes contain configurations with fewer possible contacts, provided N remains constant. To make this measure quantitative, consider the general properties of any configuration with N monomers. No more than $zN/2 = dN$ contacts are possible since each monomer has at most z nearest neighbors. The factor of 2 in this upper bound prevents double counting of each contact between pairs of monomers. Correcting this upper bound for the $N-1$ bonds between monomers yields a better upper bound on the number of contacts as
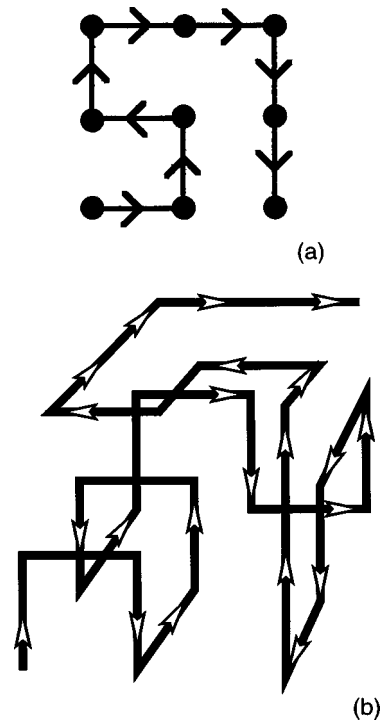

FIG. 1. A cross-back spiral with (a) $N=9$ monomers in $d=2$ dimensions and (b) $N=27$ monomers in $d=3$ dimensions. Both spirals begin in the lower left corner and the arrows indicate the order of monomer placement. The spiral in (a) [(b)] proceeds from one square [cubic] configuration to another (at $N=1$, 4, and 9 [1, 8, and 27] monomers) while passing through (nearly square) rectangular configurations (at $N=2$ and 6 [2, 4, 12, and 18] monomers). To generate a cross-back spiral with fewer monomers, simply follow the arrows until the desired number of monomers has been reached. Note that the final nine monomers of the cross-back spiral in (b) are arranged exactly like the monomers in (a).

$$(d-1)N+1, \qquad (2.1)$$

regardless of the configuration. The maximum number of contacts for any configuration with N monomers is less than in Eq. (2.1) since some monomers are incompletely surrounded by other monomers, i.e., they lie on an exposed surface.

Consider an N monomer configuration with the smallest surface area for a minimal bounding hyperrectangle with sides of length $l_1, l_2, \ldots, l_d$. Construct a second configuration from the first with a minimal bounding hyperrectangle of dimensions $l_1, l_2, \ldots, l_i + 1, \ldots, l_d$ by moving only one monomer in the first configuration. The moved monomer must lie on a $(d-1)$-dimensional surface on which no other monomers are present, thus exposing all faces in the $d-1$ dimensions. If the moved monomer makes more contacts in the second structure, the moved monomer cannot be bound to any other in the original ''configuration,'' an impossibility. Thus, increasing the volume of the minimum bounding hyperrectangle while keeping N constant does not permit an increase in the maximum possible number of contacts, provided that the structure in the starting configuration has the smallest surface area possible for the first bounding hyperrectangle. Therefore, minimization of the surface area of

configurations which fit into the smallest minimum bounding hyperrectangle yields the maximum number of contacts.

It is proven here that when $N=(n+1)^{d-i}n^i$, $0 \leq i \leq d$, the smallest minimum bounding hyperrectangle has sides in $d-i$ dimensions of length $n+1$ units and sides in the remaining i dimensions of length n units. Assume that an alternative configuration also has N monomers, but that two of the dimensions for the bounding hyperrectangle have been changed. If a better bounding hyperrectangle exists, more than two dimensions may change, but this pairwise process may be repeated until no further benefit is obtained. Now the two side lengths that changed may be A) both in the directions initially with either $n+1$ monomers or n monomers or B) one along a side initially with n monomers and the other along a side initially with $n+1$ monomers.

If the two edges to be changed are initially of the same length t ($=n$ or $n+1$), then the new bounding edges have lengths $t+a$ and $t-b$, a, $b \geq 0$, $b < t$. Since a smaller total surface area for the bounding hyperrectangle is desired, the area of the original square is required to be greater than the corresponding area of the new rectangle

$$t^2 \geq (t+a)(t-b). \tag{2.2}$$

The total volume of the new hyperrectangle cannot be less than that of the original, since not all of the monomers would fit in the new hyperrectangle otherwise. This restriction forces the equality to hold in Eq. (2.2). Algebraic rearrangement of Eq. (2.2) gives

$$\frac{ab}{t} = (a-b) \geq 0, \tag{2.3}$$

where the inequality in Eq. (2.3) is permitted since ab/t is non-negative. For the new hyperrectangle to have a smaller surface area, it must have a net loss in total length of the two changed sides. This implies that $a-b<0$, in direct contradiction to Eq. (2.3). Thus, if smaller minimum bounding hyperrectangles exist, they must be created by modifying a pair of edges with disparate lengths.

An alternative hyperrectangle which changes the lengths of unequally long edges has new lengths $n+1+a$ and $n+b$, with a and b of arbitrary sign and magnitude (negative values of a and b are less in magnitude than $n+1$ and n, respectively). As above, an equality between the two areas can be established, which, after rearrangement, gives

$$(b+a)n = -a(1+b). \tag{2.4}$$

If the new surface area is to be smaller than the original, $b+a$ must be negative. Thus, the condition on a and b from Eq. (2.4) becomes

$$a(1+b) > 0, \tag{2.5}$$

since $n>0$. If $a>0$, $b>-1$ from Eq. (2.5), and $b<0$ if $b+a<0$. But b must take an integer value and hence no so-

lution is possible if $a>0$. If $a<0$, $b<-1$ from Eq. (2.5). Solving for a in Eq. (2.4) gives

$$a = \frac{-bn}{n+1+b} < 0. \tag{2.6}$$

Now $n+1+b>0$ by definition, so Eq. (2.6) implies that $b>0$. Thus, no solution exists when a is less than zero. Therefore, when $N=(n+1)^{d-i}n^i$, the hyperrectangle with $d-i$ sides of length $n+1$ and i sides with length n is the smallest minimal bounding hyperrectangle. Since the monomers completely fill this hyperrectangle, any arrangement of monomers within the hyperrectangle produces the same surface area. Thus, chains with $N=(n+1)^{d-i}n^i$ monomers make the maximum number of contacts when they fill these smallest minimal bounding hyperrectangles.

B. Cross-back spirals

Section II A requires that an optimized walk in d dimensions grow from one hypercubic configuration with n^d monomers to the next larger one with $(n+1)^d$ monomers. In order to create a hypercube with $n+1$ monomers per side from one with n monomers per side, one must completely cover d faces whose normals are mutually orthogonal. In addition, the walk should also pass through hyperrectangle-filling configurations when $N=(n+1)^{d-i}n^i$. The growth process, therefore, completely coats one face at a time. Thus, the optimized walk grows by covering sequentially chosen adjacent orthogonal faces, passing through the requisite intermediate hyperrectangular configurations between a hypercube and the next larger hypercube.

One putative candidate for an optimized walk is the cross-back spiral. This spiral passes through hyperrectangular configurations by placing new monomers in a $(d-1)$-dimensional version of itself along the surface of the previously completed hyperrectangle until the face is completely coated. For example, a three-dimensional cross-back spiral covers its faces with two-dimensional cross-back spirals. The process of growing a cross-back spiral is illustrated in Fig. 1. To make the pattern of placement explicit, consider the direction of the bonds in Fig. 1(a), starting from the first monomer (a one-by-one ''square''). Consecutive monomers are added right, up, and then left to generate the two-by-two square. The next layer is then constructed by adding monomers up, right, right, down, and down to complete the $n=3$ square. Note that the dimer completes a 2×1 rectangle and the hexamer completes a 2×3 rectangle.

C. Cross-back spirals are optimized walks

Now consider if, for any N, there exists a configuration which has more contacts than the cross-back spiral. By definition, the cross-back spiral obtains a hyperrectangular configuration whenever the number of monomers N is $(n+1)^{d-i}n^i$ $(0 \leq i \leq d)$. As shown in Sec. II A, these hyperrectangular configurations are representative of the only configurations with maximal contacts when $N=(n+1)^{d-i}n^i$. Thus the proof that the cross-back spiral is an optimized

walk is only required for chains with length intermediate between those yielding ''consecutive'' hyperrectangular configurations [i.e. $(n+1)^{d-i}n^i<N<(n+1)^{d-i+1}n^{i-1}$].

For these intermediate configurations, the cross-back spiral is building on the $(d-1)$-dimensional face of the last completed hyperrectangle. The configuration must be contained within the hyperrectangle with $d-i+1$ sides of length $n+1$ and $i-1$ sides of length n. Since the dimensions of this bounding hyperrectangle are minimal, any alternative configuration must fit within the same bounding hyperrectangle. Thus the problem simplifies to proving that no rearrangement of the surface monomers in the cross-back spiral (within the same minimal bounding hyperrectangle) produces a smaller surface area. This proof proceeds inductively, with concurrent demonstration that cross-back spirals make the maximum number of contacts in two and in higher dimensions.

Consider cross-back spirals with $(n+1)^{d-i}n^i<N<(n+1)^{d-i+1}n^{i-1}$. The possible rearrangements of the surface monomers (ignoring connectivity constraints) may involve some combination of monomers on the incomplete face and monomers from previously completed faces. Rearrangements of only monomers from the incomplete face cannot create more contacts since the arrangement of monomers in the cross-back spiral provides the maximum number of contacts already: in two dimensions, the monomers are arranged consecutively in a straight line (the ''one-dimensional cross-back spiral''). The only possible rearrangement of these monomers that is different from the one-dimensional cross-back spiral (in terms of surface area) is achieved by placing spaces between individual monomers. Clearly, each space introduced between monomers increases the surface area. In higher than two dimensions, the arrangement of the monomers on the surface is a cross-back spiral in $d-1$ dimensions, which, by assumption, makes the maximum number of contacts. Likewise, rearrangements of only the monomers from the previously completed hyperrectangle do not change the surface area of the cross-back spiral at all, thus leaving the total number of contacts the same.

Rearranging combinations of monomers from incomplete and complete faces does not decrease the surface area relative to that of the cross-back spiral, as is now shown. Consider rearrangements where monomers are taken from only one of the completed faces. More general cases are possible, but the arguments below hold equally well for those cases. In two dimensions, the monomers along both incomplete faces should be arranged consecutively in a straight line (since leaving spaces between monomers increases the surface area). Both before and after rearrangement, the corner monomers on each face have two surfaces exposed and all other monomers on each face have one surface exposed. Therefore, the surface area does not decrease (and in rare cases may increase) on rearrangement.

In higher dimensions, both incomplete faces must be rearranged as $(d-1)$-dimensional cross-back spirals, since these spirals, by assumption, are optimal. Assume that k_1 monomers lie in the first face and k_2 monomers in the second face. By definition, the optimal arrangement for $k_1+k_2=K$ bonded monomers in $d-1$ dimensions is a $(d-1)$-dimensional cross-back spiral. If this $(d-1)$-dimensional cross-back spiral is split into multiple

independent (nontouching) groups of monomers, each group can be rearranged to form smaller (optimal) cross-back spirals, but not as many monomers in the smaller clusters can be buried as deeply as in the original K-mer structure. Thus, the optimal solution for the two incomplete faces has $k_1=K$ monomers on one face and $k_2=0$ monomers on the other. This arrangement is unattainable since all the monomers cannot fit on one face. With this restriction on the maximum size of one cluster, the optimal solution fills as many faces as possible (from largest to smallest) within the bounding hyperrectangle and makes a cross-back spiral which incompletely fills the final face with any remaining monomers. But this solution is exactly the configuration obtained by a cross-back spiral in d dimensions with N monomers. Therefore, no rearrangement of the surface elements can reduce the surface area of a cross-back spiral, proving that the cross-back spiral must be an optimized walk in any dimension.

III. CALCULATION OF THE MAXIMUM NUMBER OF CONTACTS

A. Recursive form

The maximum number of contacts for an optimized walk in d dimensions can be determined recursively. The optimized walk not only proceeds from one hypercube to the next size larger hypercube, but passes through $d-1$ hyperrectangles along the way. The equation for the maximum number of contacts for hyperrectangles (including hypercubes) requires the calculation of the surface area correction to the upper bound in Eq. (2.1). Assume that i edges have length $n+1$ monomers and the remaining $d-i$ edges have length n monomers. The surface area for a given face is the volume of the hyperrectangle [$=n^{d-i}(n+1)^i$] divided by the length of the edge normal to the face (either n or $n+1$). Thus, the maximum number of contacts for a walk which exactly fills a hyperrectangle is

$$m_{\max}[N=n^{d-i}(n+1)^i,d]$$
$$=(d-1)n^{d-i}(n+1)^i-(d-i)n^{d-i-1}(n+1)^i$$
$$-in^{d-i}(n+1)^{i-1}+1, \qquad (3.1)$$

where $n=\text{Int}(\sqrt[d]{N})$, Int($x$) is the integer portion of x, and $0\leq i<d$, i an integer. The first and last terms on the right-hand side of Eq. (3.1) are identical to Eq. (2.1) and the first and second of the remaining terms on the right-hand side arise from surface corrections from faces with an orthogonal edge of length n and $n+1$, respectively.

Generally, walks do not exactly fill hyperrectangles, but have an extra $k>0$ monomers which partially complete a face. All but the first monomer added to the incomplete face contribute the maximum of one contact with the face of the completed hyperrectangle. In addition, contacts among the monomers of the incomplete face are given by the maximum number $m_{\max}(k,d-1)$ of contacts for a k monomer walk in $d-1$ dimensions. Combining these contributions with Eq. (3.1) implies that the maximum number $m_{\max}(N,d)$ of contacts for $d\geq 2$ is given by

$$m_{max}[N=n^{d-i}(n+1)^i+k,d]=[(d-1)n^{d-i}(n+1)^i-(d-i)n^{d-i-1}(n+1)^i-in^{d-i}(n+1)^{i-1}+1]+[k-1+\delta(k,0)]$$
$$+[m_{max}(k,d-1)], \quad 0\leq k\leq n^{(d-i-1)}(n+1)^i$$
$$=(d-1)n^{d-i}(n+1)^i-(d-i)n^{d-i-1}(n+1)^i-in^{d-i}(n+1)^{i-1}$$
$$+k+\delta(k,0)+m_{max}(k,d-1),$$
$$=(d-1)n^{d-i}(n+1)^i-dn^{d-i-1}(n+1)^i\left[1-\frac{i}{d(n+1)}\right]+k+\delta(k,0)+m_{max}(k,d-1),$$

$$(3.2)$$

where the Kronecker delta $\delta(x,y)$ is unity when $x=y$ and zero otherwise. The Kronecker delta in Eq. (3.2) must be included since the minimum number of contacts between the incomplete new layer and the completed old layer is zero, but $k-1$ is negative when $k=0$. For $d=1$, m_{max} is always zero.

Computational algorithms can easily apply Eq. (3.2) to tabulate the exact maximum: calculation of m_{max} for $1\leq N\leq 2^{15}$ and $2\leq d\leq 15$ takes less than half a minute on a 200 MHz Pentium PC. Nevertheless, an exact, but nonrecursive, formula is developed below in order to elucidate the asymptotic scaling properties of m_{max} with N and d.

B. Nonrecursive form and scaling properties of m_{max}

Douglas and Ishinabe [5] propose the asymptotic ($N\rightarrow\infty$) scaling relation of m_{max} in arbitrary dimensions to be

$$m_{max}\sim a_c N-b_c N^\Delta+c_c, \quad (3.3a)$$

where $\Delta=(d-1)/d$, b_c and c_c are constants, and $a_c=d-1$ for hypercubic lattices. In addition, they propose that for arbitrary N on hypercubic lattices,

$$m_{max}\approx Int[(d-1)N-dN^\Delta+1], \quad (3.3b)$$

where $Int[x]$ is the integer portion of x and $\Delta=(d-1)/d$. Equation (3.3b) is exact for all N in $d\leq 2$ and for $N=n^d$, n an integer, in all dimensions. In $d=3$, the expression on the right-hand side of Eq. (3.3b) provides a fairly tight upper bound, never exceeding m_{max} in Eq. (3.2) by more than three for $N<250$. Yet, comparison between $m_{max}(250=3^5+7,5)$ from Eq. (3.2) and the right-hand side of Eq. (3.3b) yields a difference of 9 between the upper bound and the exact solution. Making b_c a function of N such that the right-hand side of Eq. (3.3b) is exact for configurations which completely fill hyperrectangles does not significantly tighten the upper bound except for those particular configurations.

Formulating the nonrecursive exact expression for m_{max} elucidates potential corrections to Eq. (3.3). Consider the contributions from the recursive relation Eq. (3.2). The number N of monomers is essentially broken down into two parts, the number $N_d[=n^{d-i}(n+1)^i]$ of monomers which complete a hyperrectangle in d dimensions plus the remaining $k=N-N_d$ monomers which only partially complete a face in d dimensions. This decomposition then recurs for k in $d-1$ dimensions: $N_{(d-1)}$ monomers complete a hyperrectangle in $d-1$ dimensions and $(N-N_d-N_{(d-1)})$ are decomposed in the next lower dimension. This process repeats until

either all monomers have been accounted for in higher than one dimension (in $j_{min}>1$ dimensions) or else the number of remaining monomers is assigned to equal N_1 (and $j_{min}=1$), giving the relation for N,

$$N=N_d+N_{(d-1)}+\cdots+N_{j_{min}}=\sum_{j=j_{min}}^{d} N_j. \quad (3.4)$$

Substituting $k_d=N_{d-1}+k_{d-1}$ into Eq. (3.2) gives the starting point for the nonrecursive form

$$m_{max}(N=N_d+k_d,d)=(d-1)N_d-\frac{dN_d}{n_d}\left[1-\frac{i_d}{d(n_d+1)}\right]$$
$$+k_d+\delta(N,N_d)+m_{max}(N_{d-1}+k_{d-1},d-1),$$

$$(3.5)$$

where $n_d=Int(\sqrt[d]{N_d})$ and

$$i_d=Int\left\{\frac{\ln(N/n_d^d)}{\ln[(n_d+1)/n_d]}\right\}.$$

The last term on the right-hand side of Eq. (3.5) is of the same form as the left-hand side of Eq. (3.5), so the collection of similar terms from lower dimensional evaluations of m_{max} yields the preliminary nonrecursive form for $m_{max}(N,d)$,

$$m_{max}(N,d)=\sum_{j=j_{min}}^{d}\left\{(j-1)N_j-\frac{jN_j}{n_j}\left[1-\frac{i_j}{j(n_j+1)}\right]+k_j\right\}$$
$$+1, \quad (3.6)$$

where

$$n_j=Int(\sqrt[j]{N_j}),$$

$$i_j=Int\left\{\ln\left(\frac{N_j}{n_j^j}\right)\Big/\ln[(n_j+1)/n_j]\right\},$$

and

$$k_j=N-\sum_{t=j}^{d} N_t.$$

A factor of unity in Eq. (3.6) replaces the Kronecker deltas from Eq. (3.2). In dimension j_{min}, there are no extra monomers, thus satisfying the nonzero condition of the Kronecker delta, while in higher dimensions, the Kronecker deltas give

a zero, since a nonzero number of monomers remain after filling the largest hyperrectangle. Evaluation of the sum over k_j in Eq. (3.6) is facilitated by counting how many times a particular N_x appears, x a dimension between j_{min} and d. The sum over k_j is thus replaced as

$$\sum_{j=j_{min}}^{d} k_j = \sum_{j=j_{min}}^{d} \left\{ N - \sum_{1=j}^{d} N_1 \right\}$$

$$= (d-j_{min}+1)N - \sum_{j=j_{min}}^{d} (j-j_{min}+1)N_j. \tag{3.7}$$

Substitution of Eq. (3.7) into Eq. (3.6) followed by algebraic rearrangement yields the simplified expression,

$$m_{max}(N,d) = (d-j_{min}+1)N + \sum_{j=j_{min}}^{d} \left\{ (j_{min}-2)N_j - \frac{jN_j}{n_j} \right.$$

$$\left. \times \left[1 - \frac{i_j}{j(n_j+1)} \right] \right\} + 1. \tag{3.8}$$

A trivial substitution of Eq. (3.4) for the sum over N_j in Eq. (3.8) gives the final expression for m_{max},

$$m_{max}(N,d) = (d-1)N + 1 - \sum_{j=j_{min}}^{d} \frac{jN_j}{n_j} \left[1 - \frac{i_j}{j(n_j+1)} \right]. \tag{3.9}$$

Thus, $a_c=(d-1)$ and $c_c=1$ for hypercubic lattices and the error in the estimate provided by Eq. (3.3b) arises exclusively from the approximate dN^{Δ} surface correction term.

The exact surface area corrections to Eq. (2.1) take a polynomial form. Typically, the optimal walk has N monomers with the restriction

$$N^{d-i}(n+1)^i \leq N \leq n^{d-i-1}(n+1)^{i+1}. \tag{3.10}$$

By definition, $N_d=n^{d-i}(n+1)^i$ and $N_{(d-1)} \leq N - N_d$. Applying these definitions to Eq. (3.10) provides a restriction on $N_{(d-1)}$,

$$N_{(d-1)} \leq n^{d-i-1}(n+1)^i = \frac{N_d}{n}. \tag{3.11}$$

Each correction in every lower dimension (though j_{min}) has a similar restriction as in Eq. (3.11), i.e., $N_j \leq N_{(j+1)}/n_{(j+1)}$. Now n_d is of order $\sqrt[d]{N}$. Thus, the order of N in Eq. (3.9) can vary as

$$O(N^{\Delta}) \sim \sum_{i=0}^{d-1} O(N^{(d-1-i)/d}) \sim O(N^{(d-1)/d}) \text{ for large } N. \tag{3.12}$$

Equation (3.12) suggests that $\Delta=(d-1)/d$ asymptotically, and confirmation of this exact asymptotic value is straightforward. The exponent Δ is well estimated by assuming that the surface area term $A(N,d) \equiv m_{max}(N,d) - a_c N - c_c$ scales as N^{Δ} and then taking the average of the ratio $\ln[A(N,d)]/\ln(N)$ for $N \gg 1$. For $d=2$, $\langle \Delta \rangle$ is 0.500002 and for $d=3$, $\langle \Delta \rangle$ is

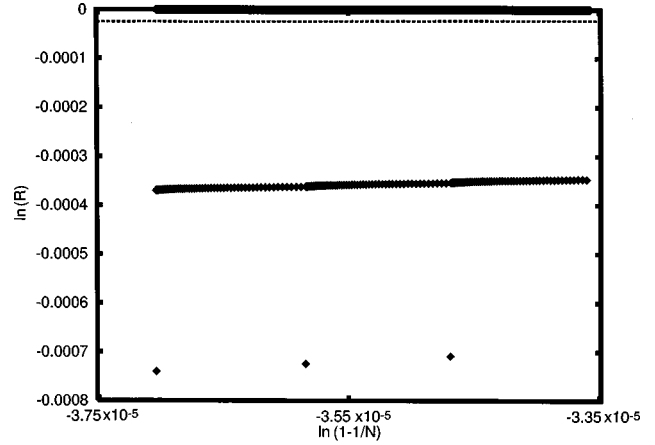FIG. 2. Log-log plot of $R=A(N-1,3)/A(N,3)$ as a function of $1-1/N$. The surface area correction $A(x,y)$ is the difference between Eq. (2.1) and $m_{max}(x,y)$ and N is the number of monomers. The best fitting line (dashed) has slope of almost exactly $\frac{2}{3}[=(d-1)/d]$ and a y intercept of 0. The tight clustering of points and zero y intercept for plots in this and other dimensions suggests that $\Delta=(d-1)/d$ asymptotically.

0.666667 which is close to the values $\frac{1}{2}$ and $\frac{2}{3}$ from $(d-1)/d$. Further support for Δ asymptotically equaling $(d-1)/d$ comes from an analysis using the ratio method [12]. Since $A(N,d) \sim N^{\Delta}$, the ratio $A(N-1,d)/A(N,d)$ goes as $(1-1/N)^{\Delta}$. The best fit line through the log-log plot of the ratio of surface area terms as a function of $(1-1/N)$ gives Δ as the slope. The validity of the method hinges on having a y intercept of zero. Figure 2 displays the results of the ratio method for $d=3$ and $30^3 < N < 31^3$. Although scatter is present in the plot, the variation is numerically minuscule and the y intercept is zero to within parts per million. The best fitting line has the equation $\ln(A[N-1,d]/A[N,d]) = 0.666665 \ln(1-1/N)$, suggesting that $\Delta=(d-1)/d$ for $d=3$. Similar results are observed for other dimensions, implying that $\Delta=(d-1)/d$ asymptotically in all dimensions.

A similar analysis provides the asymptotic value for b_c. The ratio $A(N,d)/N^{\Delta}$ defines b_c for a given N-mer in d dimensions. Taking the average over large N $(N \gg 1)$ of this ratio provides an estimate for b_c. For example, in $d=2$ dimensions and for $100^2 < N < 101^2$, values of b_c range from 2 to 2.1, with the average 2.00494 ± 0.00029. As N grows larger, the upper bound of the range grows smaller while the lower bound cannot change. Thus, the asymptotic limit must be 2 in $d=2$. Similar trends are observed in other dimensions suggesting that $b_c=d$ asymptotically. To illustrate this trend further, consider the ratio between $A(N,d)$ and $A(N_0,d)$ as a function of N, where $N_0 \equiv n^{d-i}(n+1)^i$ is the number of monomers which complete the largest possible ''optimized'' hyperrectangle in d dimensions with no more than N monomers. Figure 3 depicts the typical scaling of the ratio as a function of N for $d>3$: the log of the ratio varies linearly with $\ln(N)$ for $N_0 \leq N \leq N_0(n+1)/n-1$ and appears as a series of nearly parallel lines with slope approximately $(d-1)/d$. When $i=0$, the slope is greater than $(d-1)/d$ and when $i=d-1$, the slope is less than $(d-1)/d$, with a monotonic decrease in between. The range of slope values decreases with increasing n, but always brackets $(d-1)/d$.
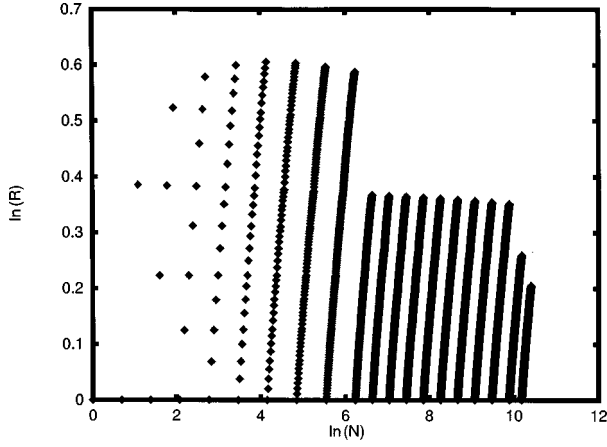
FIG. 3. Log-log plot of
$$R \equiv A(N,3)/(dn^{d-1-i}(n+1)^i\{1-i/[d(n+1)]\})$$
as a function of N in $d=9$. The points for which $n^{d-i}(n+1)^i \leq N < n^{d-1-i}(n+1)^{i+1}$ appear to form straight lines with slope approximately $(d-1)/d$. The slope, however, is greater than $(d-1)/d$ for $i=0$ and monotonically decreases to less than $(d-1)/d$ for $i=d-1$. The range of slope values decreases as n grows.

Thus, asymptotically, the slope is $(d-1)/d$ as observed previously and the asymptotic ratio obeys the relation

$$\frac{A(N,d)}{A(N_0,d)} = \left(\frac{N}{N_0}\right)^{(d-1)/d}, \quad N \to \infty, \quad (3.13a)$$

implying that

$$A(N,d) = d\left(\frac{n+1}{n}\right)^{i/d}\left[1 - \frac{i}{d(n+1)}\right]N^{(d-1)/d}, \quad N \to \infty. \quad (3.13b)$$

The formulation of $A(N,d)$ in Eq. (3.13b) corresponds to the form of the surface area term in Eq. (3.3a), suggesting that b_c must be d asymptotically:

$$b_c = \lim_{n\to\infty} d\left(\frac{n+1}{n}\right)^{i/d}\left(1 - \frac{i}{d(n+1)}\right) = d. \quad (3.14)$$

Hence, the asymptotic scaling of m_{max} for linear chains on hypercubic lattices is

$$m_{max} \sim (d-1)N - dN^{(d-1)/d} + 1, \quad (3.15)$$

which is essentially the right-hand side of Eq. (3.3b). The right-hand side of Eq. (3.3b) is used as a numerical upper bound for m_{max} in MC simulations. But for small enough N [i.e., $n \sim O(1)$], the lower order corrections in Eq. (3.12) become more important (see Fig. 4), suggesting that the asymptotic formula Eq. (3.15) only provides correct order of magnitude information, especially in higher dimensions.

IV. DISCUSSION

An exact formulation is developed for the maximum number m_{max} of intramolecular contacts in a single linear homopolymer of any length on hypercubic lattices of any dimension. The formulas for m_{max} follow from a proof that a
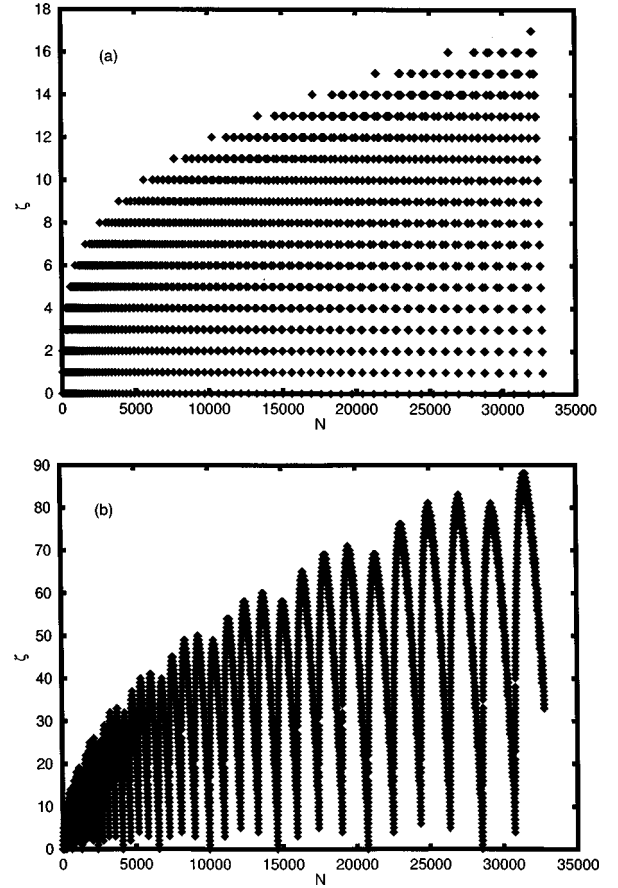

FIG. 4. Difference ζ between the exact m_{max} of Eq. (3.2) and the approximate Eq. (3.3b) in (a) $d=3$ and (b) $d=4$ as a function of N. The asymptotic limit of Eq. (3.3a) is approached even though the largest difference between the exact and approximate equations for m_{max} continues to grow, since the difference only grows as $O(N^{(d-2)/d})$. These differences make the approximation Eq. (3.3b) only good for order of magnitude estimates of m_{max}, especially in higher dimensions.

single self-avoiding walk, the cross-back spiral (defined in Sec. II B), always obtains a configuration with the maximum number of contacts. The proof begins by establishing that any representative walk must pass through configurations which exactly fill a hyperrectangle with i sides of length n and $d-i$ sides of length $n+1$ whenever the number N of monomers equals $n^i(n+1)^{d-i}$, with n a positive integer and $0 \leq i \leq d$, i an integer. Establishing these required configurations places fairly strong restrictions both on the potential representative walks and on alternative configurations with a competitive number of contacts.

The cross-back spiral is then inductively shown to be a walk with maximal contacts for any length linear chain in any dimension. Alternative configurations to the cross-back spiral must be bound by the same hyperrectangle as the cross-back spiral (see Sec. II A). Further, the rearrangements of the cross-back spiral which have the potential to improve the number of contacts only involve monomers on its surface. These relatively few rearrangements are shown to yield at best the same number of contacts as the cross-back spiral: in two dimensions, the maximum number of contacts occurs when no spaces exist between any pairs of monomers along any of the faces. In higher dimensions, the

$(d-1)$-dimensional cross-back spiral is assumed to be representative, from which it follows that the arrangement of monomers on the surface of the d-dimensional cross-back spiral produces the maximum number of contacts.

The construction of the cross-back spiral provides exact recursive [Eq. (3.2)] and nonrecursive [Eq. (3.9)] formulas for the maximum number $m_{max}(N,d)$ of contacts for a linear lattice polymer with N monomers in d dimensions. The exact scaling of m_{max} with N, derived in Eq. (3.12) from Eq. (3.9), reveals that m_{max} scales with N in a nonuniform way, depending on the roughness of the surface of the cross-back spiral. Nevertheless, m_{max} asymptotically takes the form

$$m_{max}(N,d) \sim (d-1)N - dN^{(d-1)/d} + 1, \qquad (4.1)$$

in accordance with the predictions of Douglas and Ishinabe [5]. Yet, m_{max} frequently has significant lower order corrections which render Eq. (4.1) an order of magnitude upper bound, especially in higher dimensions. Thus, the exact Eq. (3.2) is preferable to Eq. (4.1) in calculating m_{max} for comparison with the number of contacts in collapsed chain configurations as sampled from MC simulations.

Although the cross-back spiral is representative of configurations with the maximum number of contacts, they are not necessarily representative of the thermodynamically favored state. The increase in contacts lowers the internal energy of the system, but may incur too large an entropic penalty, producing a configuration which does not correspond to a free energy global minimum. Nevertheless, the minimum energy structure is usually assumed to represent the equilibrium conformation in protein folding studies. For a single homopolymer under chain collapse conditions, this assumption is probably acceptable, since the permutations of geometries which permit the maximum number of contacts produce a combinatorially large number of arrangements.

The assumption about the equivalence between the global minimum of the internal and free energies does not necessarily hold for systems as simple as linear random copolymers under chain collapse conditions. For these copolymers, the connectivity and excluded volume constraints remain the same, but the ''chemical'' or energetic interactions of adjacent monomers may be completely different. For example, in the HP model [16], monomers along the chain are labeled as either noninteracting polar ''P'' monomers or as hydrophobic ''H'' monomers, which are attracted only to other H's. It is predicted that finding the maximum number of contacts for a given HP sequence on hypercubic lattices is an NP-complete problem [17]. This prediction is quite plausible, since replacement of H monomers in the all H homopolymer (used in this paper) with P's can place the P's in the interior of the cross-back spiral, a highly unfavorable energetic event. Reduction of this energy penalty requires large scale rearrangement of monomers in the chain such that as much as possible the P's lie on the surface and the H's are placed in the interior. But for different sequences, the requirements for placing these buried P's on the surface vary strongly with the ordering of the P's in the chain. That is, the sequence order plus connectivity constraints requires considering more than the minimum bounding volume for the given number of H's. In principle, the HP model may have a single lowest energy configuration, but many other configurations with a slightly higher internal energy. For this circumstance, the representative minimum free energy conformation may not correspond to the polymer in the single lowest energy conformation.

Likewise, determining the minimum internal energy structure is frequently used to predict the native states of proteins. Generally, the algorithms place hydrophobic residues in the interior of the protein and hydrophilic residues on the surface, all while keeping connectivity constraints satisfied. More fine tuning is added by differentiating among the types of hydrophobes and of hydrophiles and by approximating what proclivities groups of residues have to form α helices and β sheets. Consistent success using these techniques to predict the correct three-dimensional structures of proteins remains elusive. Probably, a major component missing in the analysis of protein structure is proper description of side-chain packing [18]. As in lattice studies [19–23], structural detail may also lead to increased specificity and altered equilibrium properties, even in simple energy functions such as Geocore [24,25], but, most likely, a more sophisticated method which includes entropic effects must be developed for predicting representative native conformations.

[1] K. W. Foreman and K. F. Freed, Adv. Chem. Phys. **103**, 335 (1998), and references therein.

[2] K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, and H. S. Chan, Protein Sci. **4**, 561 (1995), and references therein.

[3] H. S. Chan and K. A. Dill, Proteins: Struct., Funct., Genet. **30**, 2 (1998).

[4] D. Baker, H. S. Chan, and K. A. Dill, J. Chem. Phys. **98**, 9951 (1993).

[5] J. F. Douglas and T. Ishinabe, Phys. Rev. E **51**, 1791 (1995).

[6] E. N. Gilbert, Bell Syst. Tech. J. **37**, 815 (1957).

[7] U. Schnell, Monatsh. Math. **116**, 331 (1993).

[8] P. D. Lax, J. Funct. Anal. **46**, 280 (1982).

[9] J. E. Brothers and F. Morgan, Michigan Math. J. **41**, 419 (1994).

[10] M. Silver, Pac. J. Math. **40**, 451 (1972).

[11] E. A. Bender, Am. Math. Monthly **69**, 742 (1962).

[12] J. Douglas, C. M. Guttman, A. Mah, and T. Ishinabe, Phys. Rev. E **55**, 738 (1997).

[13] K. Yue and K. A. Dill, Proc. Natl. Acad. Sci. USA **92**, 146 (1995).

[14] H. S. Chan and K. A. Dill, Macromolecules **22**, 4559 (1989).

[15] H. S. Chan and K. A. Dill, J. Chem. Phys. **92**, 3118 (1990).

[16] K. F. Lau and K. A. Dill, Macromolecules **22**, 3986 (1989).

[17] W. E. Hart and S. Istrail, J. Comput. Biol. **4**, 1 (1997).

[18] M. Levitt, M. Gerstein, E. Huang, S. Subbiah, and J. Tsai, Annu. Rev. Biochem. **66**, 549 (1997).

[19] J. Dudowicz and K. F. Freed, Macromolecules **24**, 5112 (1991).

[20] J. Dudowicz and K. F. Freed, Macromolecules **29**, 8960 (1996).

[21] K. F. Foreman, K. F. Freed, and I. M. Ngola, J. Chem. Phys. **107**, 4688 (1997).

[22] K. F. Foreman and K. F. Freed, Macromolecules **30**, 7279 (1997).

[23] K. F. Freed, J. Dudowicz, and K. F. Foreman, J. Chem. Phys. **108**, 7881 (1998).

[24] K. Yue and K. A. Dill, Protein Sci. **5**, 254 (1996).

[25] B. H. Park, E. S. Huang, and M. Levitt, J. Mol. Biol. **266**, 831 (1997).