

# POLYMER PRINCIPLES IN PROTEIN STRUCTURE AND STABILITY

*Hue Sun Chan and Ken A. Dill*

Department of Pharmaceutical Chemistry, University of California,  
San Francisco, California 94143-1204

KEY WORDS: protein conformations, excluded volume, local and nonlocal interactions, heteropolymer collapse

---

## CONTENTS

<i>Perspectives and Overview</i> .....	447
<i>The Configurational Partition Function</i> .....	449
<i>Local Interactions</i> .....	451
<i>Nonlocal Interactions</i> .....	455
<i>Exhaustive Lattice Simulations</i> .....	459
<i>The Path-Integral Method</i> .....	460
<i>Homopolymer Collapse Theories</i> .....	464
<i>Heteropolymer Collapse</i> .....	468
<i>Spin-Glass Models</i> .....	473
<i>Internal Structure in Compact Polymers</i> .....	478
<i>Conclusions</i> .....	485

## *Perspectives and Overview*

Proteins are polymers. This review is intended as an introduction to methods and principles of polymer statistical mechanics, old and new, that bear on the large conformational changes and internal organization in protein molecules. We focus on the relative roles played by local vs non-local forces. *Local* defines interactions among near neighbors in the chain sequence; these have often been referred to as short-ranged or secondary forces (see Figure 1). *Nonlocal* describes interactions among monomers distant in the chain sequence, often referred to as long-ranged or tertiary forces.

447

0883-9182/91/0610-0447\$02.00

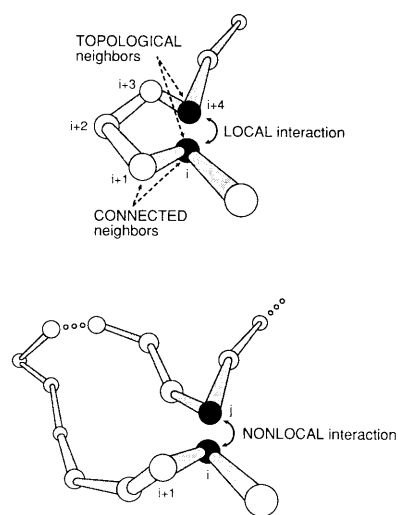


Figure 1 Spatially neighboring monomers ( $i, j$ ) are defined as *connected* neighbors if they share a backbone bond,  $j = i + 1$ ; otherwise, they are *topological* neighbors. Interactions are local or non-local depending on the separation along the chain of the interacting monomers.

First we discuss the concept of separability, the idea that interactions in polymers can be separated into two types of force, local and nonlocal. Then we describe local forces. Under conditions in which nonlocal forces are small (the Flory “theta” state), chain conformations are determined by local forces. In these cases, conformational changes depend on an elastic entropy. Effects of monomer sequence are well-modeled using the rotational isomeric state (RIS) model, implemented in one-dimensional Ising model matrix methods. This approach is the basis of helix-coil transition theories, widely applied to polypeptides. Next we describe nonlocal forces. Monomers that are far apart in the sequence interact through their mutual excluded volume and through solvent-mediated interactions. We review the Flory lattice model of nonlocal interactions, then computer lattice simulation results, and path integral methods. We then discuss solvent-driven compactization of homopolymers, heteropolymers, and proteins. Finally, we consider how the compactness of a chain molecule can lead to internal structure.

Among the several conclusions summarized here are two recent and unexpected results. First, whereas helices and sheets in globular proteins have been generally assumed to arise from local forces, evidence now indicates that excluded volume instead may be significantly responsible for protein structure. Second, whereas homopolymers in poor solvents will collapse to a relatively large ensemble of compact conformations, heteropolymers are predicted to collapse to an extremely small ensemble:

one or two or a few conformations. Thus, the ways in which proteins apparently differ from simpler polymers—(a) that they have unique native structures dictated by the monomer sequence and (b) that native conformations are comprised of helices, sheets, and irregular conformations—appear to arise as natural consequences of heteropolymer collapse processes.

### *The Configurational Partition Function*

In this review, we focus on how the configurations of polymers, particularly proteins, derive from two types of interaction: those among monomers that are near neighbors in the sequence, and those among monomers that are far apart in the sequence. First, we describe the rationale for using the terms “local” and “nonlocal” to describe these two types of interaction. For some years (47), standard usage has been to refer to interactions in chain molecules as “short-ranged” or “long-ranged,” depending on whether the interacting monomers are near neighbors in the chain sequence or not. However, this conflicts with terminology, considerably older, wherein the “range” of interaction refers to its spatial distance dependence: ion-ion interactions are long-ranged, and van der Waals interactions are short-ranged, for example. Ambiguity arises because the context may not clarify whether “range” refers to a particular distance dependence of a given type of energy or to the separation of the monomers along the chain sequence. Two monomers of a chain molecule that are far apart in the sequence may interact via either van der Waals or electrostatic interactions; conversely, two monomers that are close in sequence may also interact via either van der Waals or electrostatic interactions. The potential confusion is greatest for proteins, for which the significant interactions can involve both types of distance dependence (van der Waals and electrostatic) and both types of chain separation (together and apart in the sequence). To avoid these ambiguities, we use long-ranged and short-ranged to refer to the distance dependence of the type of energy because that usage is older and more widespread, and we use local and nonlocal to refer to interactions among monomers close or distant in the sequence of a chain molecule. We do not favor “secondary” and “tertiary” in reference to interactions because these terms are vague and refer neither to the range of interaction nor to the positions of interacting monomers in the chain: for example secondary interaction can refer to either helix (local) or sheet (nonlocal) contacts.

The conformational properties of a polymer molecule are described by its partition function. Consider a polymer chain of  $n$  bonds ( $n + 1$  monomers) in one specific configuration in  $d$ -dimensional space, such that monomer 0 is at position  $\mathbf{r}_0$ , monomer 1 is at position  $\mathbf{r}_1$ , and monomer  $n$  is at

position  $\mathbf{r}_n$ . The conformation of the molecule is completely specified by the set of vectors  $\{\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n\}$ . Alternately, the spatial vectors of the monomers may be labeled by a continuous contour measure  $\tau$  from one end of the chain, such that  $\mathbf{r}(\tau_i) \equiv \mathbf{r}_i$ , where  $\tau_i$  is the position of monomer  $i$  along the contour of the chain; see Figure 2. The latter notation is more general; it is applicable to chains with either discrete or continuous segments (see the path integral section below). The conformation is specified by the values of  $\{\mathbf{r}(\tau_0), \mathbf{r}(\tau_1), \mathbf{r}(\tau_2), \dots, \mathbf{r}(\tau_n)\}$ .

The partition function,  $Q$ , is the total number of conformations accessible to the chain,

$$Q = \mathcal{N} \int d\mathbf{r}(\tau_0) \int d\mathbf{r}(\tau_1) \dots \int d\mathbf{r}(\tau_n) P[\mathbf{r}(\tau_0), \mathbf{r}(\tau_1), \dots, \mathbf{r}(\tau_n)], \quad 1.$$

where  $P(\{\mathbf{r}(\tau)\}) \equiv P[\mathbf{r}(\tau_0), \mathbf{r}(\tau_1), \dots, \mathbf{r}(\tau_n)]$  is the relative probability density of a conformation in the  $d\mathbf{r}(\tau_0)d\mathbf{r}(\tau_1)\dots d\mathbf{r}(\tau_n)$  neighborhood specified by  $\{\mathbf{r}(\tau_0), \mathbf{r}(\tau_1), \dots, \mathbf{r}(\tau_n)\}$ . The quantity  $\mathcal{N}$  is an overall normalization, and  $\{\mathbf{r}(\tau)\}$  is shorthand for the set of  $n+1$  spatial vector variables  $\mathbf{r}(\tau_i)$ . The quantitative value of the normalization  $\mathcal{N}$  is unimportant because it cancels out in most practical applications. The summation over conformations is carried out by integrating  $P$  over all possible positions  $\mathbf{r}(\tau_i)$  for each chain segment  $0 \leq i \leq n$ . The range of each  $\int d\mathbf{r}(\tau_i)$  integration is the entire  $d$ -dimensional space accessible to the chain.

The probability density  $P(\{\mathbf{r}(\tau)\})$  is given by the Boltzmann distribution

$$P(\{\mathbf{r}(\tau)\}) = \exp[-W(\{\mathbf{r}(\tau)\})]. \quad 2.$$

If the chain molecule is immersed in a solvent,  $W(\{\mathbf{r}(\tau)\})$  is a potential of mean force (54, 64, 68) in units of  $kT$ . That is, if the degrees of freedom of the solvent are represented as a vector  $\zeta$ , then  $W(\{\mathbf{r}(\tau)\})$  is the reversible work required to put the polymer into configuration  $\{\mathbf{r}(\tau)\}$  averaged over

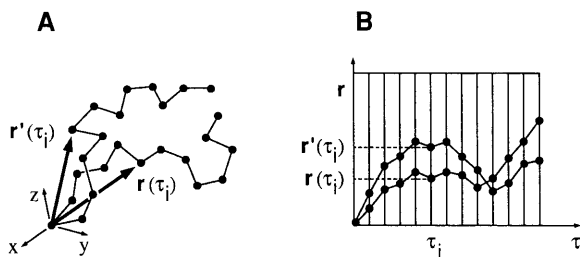


Figure 2 Chain conformations (A) as paths in the  $\mathbf{r}$ - $\tau$  space (B). Summation over conformations is equivalent to summation over  $\mathbf{r}$ - $\tau$  paths with appropriate statistical weights.

the ensemble of all the solvent degrees of freedom  $\xi$ , i.e. the force acting on the  $i$ th monomer is

$$-kT\nabla_{\mathbf{r}(\tau)}W(\{\mathbf{r}(\tau)\}) \equiv \frac{-\int d\xi \nabla_{\mathbf{r}(\tau)}V(\{\mathbf{r}(\tau)\}, \xi) \exp[-V(\{\mathbf{r}(\tau)\}, \xi)/kT]}{\int d\xi \exp[-V(\{\mathbf{r}(\tau)\}, \xi)/kT]}, \quad 3.$$

where  $V(\{\mathbf{r}(\tau)\}, \xi)$  is the potential energy for the chain and the solvent configuration specified by  $\{\mathbf{r}(\tau)\}$  and  $\xi$ .

It is useful to separate the total potential of mean force into a sum of two terms, one owing to local and the other to nonlocal interactions. Local interactions are defined as those that occur among monomers that are separated by up to  $p$  units; nonlocal interactions involve monomers separated in the sequence by more than  $p$  units,

$$W(\{\mathbf{r}(\tau)\}) = W_l(\{\mathbf{r}(\tau)\}) + W_{nl}(\{\mathbf{r}(\tau)\}),$$

$$W_l(\{\mathbf{r}(\tau)\}) = \sum_{j=0}^{n-p} w_j[\mathbf{r}(\tau_j), \mathbf{r}(\tau_{j+1}), \dots, \mathbf{r}(\tau_{j+p})], \quad 4.$$

where  $w_j$ s account for the local potential of mean force.  $W_{nl}$  is the nonlocal potential of mean force that accounts for all other interactions not expressible in the functional form of  $W_l$ . This separation of terms is useful because two qualitatively different types of behavior are observed in polymers depending on whether local interactions are dominant or not (see below). Since by definition  $W_l$  only contains local interactions,  $p$  is a small number of order 1. Typically  $p$  is chosen to be in the range of 1 to 4. For  $p = 1$ ,

$$W(\{\mathbf{r}(\tau)\}) = \sum_{j=0}^{n-1} w_j[\mathbf{r}(\tau_j), \mathbf{r}(\tau_{j+1})] + W_{nl}(\{\mathbf{r}(\tau)\}). \quad 5.$$

In this case, the local potential of mean force,  $W_l$ , is defined as interactions among connected neighbors, whereas the nonlocal potential of mean force,  $W_{nl}$ , is among topological neighbors (9–11) (see Figure 1); the latter involves excluded volume and solvent interactions (see below).

### Local Interactions

Under special solvent conditions, referred to as the Flory theta state (47, 48) (see below), the nonlocal interactions are negligible,  $W_{nl} \simeq 0$ , and the configurational states of polymers are governed primarily by local interactions. Local interactions are important for protein structure and stability. They account for stabilities of helices in solution and they contribute an elastic entropy to the folding of globular proteins. In this section,

we review the following principles for conformations of chains driven by local interactions. 1. In the most elementary model of local interactions, for which properties of bonds adjacent along the chain are uncorrelated, the chain conformations may be suitably modeled as random flights. For random flights, the configurational distribution function of the end-to-end distance is Gaussian, which leads to an elastic contribution to the conformational entropy. 2. Even if the bond directions among neighboring links are correlated, provided only that the correlations are local, then the statistics remain those of a random flight. Only an effective bond length is changed; the dependence of end-to-end distance on square root of the chain length remains unaffected, except for a constant. 3. More accurate than the random-flight model is the rotational isomeric state model for capturing the chemical bond details, the effects of monomer sequence, and the finite lengths of short chains.

The simplest classical model for long chains subject only to local interactions assumes that bonds have fixed length and that adjacent bond angles are uncorrelated. This model can be expressed equivalently in terms of a Gaussian bond distribution function (54, 148),

$$\exp\{-w_j[\mathbf{r}(\tau_j), \mathbf{r}(\tau_{j+1})]\} = \left(\frac{d}{2\pi l^2}\right)^{d/2} \exp\left[-\frac{d}{2l^2}|\mathbf{r}(\tau_j) - \mathbf{r}(\tau_{j+1})|^2\right], \quad 6.$$

where  $l$  is the bond length. The distribution of end-to-end lengths  $R_n \equiv |\mathbf{r}(\tau_n) - \mathbf{r}(\tau_0)|$  is obtained by integrating Equation 6 over all spatial coordinates  $\mathbf{r}(\tau_1), \dots, \mathbf{r}(\tau_{n-1})$  except the ends; the result is a Gaussian distribution,

$$P(R_n) = \left(\frac{d}{2\pi R_0^2}\right)^{d/2} \exp\left[-\frac{d}{2}\left(\frac{R_n}{R_0}\right)^2\right], \quad 7.$$

where  $R_0 = \sqrt{\langle R_n^2 \rangle} = ln^{1/2}$  is the root-mean-square, end-to-end length of the random-flight chain.

A random-flight molecule subject to small deformation has a retractive *elastic* entropy, so-called because of its importance in the theory of rubber elasticity (48). For a uniform isotropic expansion or contraction of the chain by a factor  $\alpha$  to a root-mean-square, end-to-end distance  $R = \alpha R_0$ , the change in (entropic) free energy  $\Delta F_{cl}$  owing to elastic deformation is deduced from simple scaling (48) of the distribution equation, Equation 6:

$$\frac{\Delta F_{cl}}{kT} = \frac{d}{2}(\alpha^2 - 1) - d \ln \alpha. \quad 8.$$

To establish from experiments whether local or nonlocal interactions are dominant, one must measure  $\langle R^2 \rangle$  vs  $n$ . If nonlocal interactions are small, the mean square end-to-end distance  $\langle R_n^2 \rangle$  takes the general form

$$\langle R_n^2 \rangle = Cnl^2, \quad 9.$$

where  $C = 1$  for chains that obey the random-flight statistics (i.e. Equations 6 and 7). A most remarkable feature of chain conformations is that this dependence,  $\langle R_n^2 \rangle \sim n$  for large  $n$ , is unaffected even if nearest neighbor bonds are correlated, and also if second neighbor bonds are correlated, etc, provided that correlations extend out to only some finite number of neighbors. In those cases, the effect of local bond correlations is only to change the constant  $C$  but not the scaling behavior of  $\langle R_n^2 \rangle$  with  $n$  (48, 54, 148). Thus the distinction between local and nonlocal is not sharply defined in terms of some particular number of neighbors distant in the sequence over which interactions occur (i.e. the exact choice of  $p$  in Equation 4 is not critical as long as  $p \ll n$ ). Rather, the critical evidence that indicates when chain behavior is dominated by local interactions is experimental observation that the mean square end-to-end length scales with the first power of the chain length. The scaling behavior is determined by the nature of the solvent and the temperature (see below).

A more accurate molecular treatment of polymer conformations that are subject only to local interactions  $W_1$  is given by the rotational isomeric state theory (5, 49, 139). In this model, the continuum of conformational states of sequential bonds is approximated as a set of discrete rotational isomers. Now a pair of sequential bonds need no longer be represented as uncorrelated; local details of bond chemistry can be more suitably represented. The main advantages of this approach are: (a) the constant  $C$  and other moments of a chain configurational distribution function can be calculated from molecular details of the chemical bonds and are found to be generally in good agreement with experiments, (b) short chains of any length can be treated, and (c) the effects of the monomer sequence can be treated explicitly.

The degrees of freedom considered by rotational isomeric state models are the bond rotation angles  $\phi_i$  between pairs of planes defined by sequential pairs of bonds,  $\phi_i \equiv \cos^{-1}[(\mathbf{b}_{i-2} \times \mathbf{b}_{i-1}) \cdot (\mathbf{b}_{i-1} \times \mathbf{b}_i)] / [|\mathbf{b}_{i-2} \times \mathbf{b}_{i-1}| |\mathbf{b}_{i-1} \times \mathbf{b}_i|]$ ,  $2 \leq i \leq n$ , where  $\mathbf{b}_i \equiv \mathbf{r}(\tau_{i+1}) - \mathbf{r}(\tau_i)$  is the bond vector from monomer  $i$  to monomer  $i+1$ . Bond lengths  $|\mathbf{b}_i|$  and bond angles  $\theta_i \equiv \cos^{-1}(\mathbf{b}_{i-1} \cdot \mathbf{b}_i / |\mathbf{b}_{i-1}| |\mathbf{b}_i|)$  are fixed. The model makes the simplifying assumption that  $\phi_i$  may adopt any one of a finite set of discrete values,  $\phi_i = \phi_i^{(k)}$ ,  $k_i = 1, 2, \dots, s$ , hence the integrations over  $\mathbf{r}(\tau_i)$ s in the expression (Equation 1) for the partition function  $Q$  can be approximated by summations over discrete values of  $\phi_i$ s. Positions of four sequential

monomers are required to determine a single bond rotation angle; therefore when pairs of sequential bond rotation angles  $\phi_i$  and  $\phi_{i+1}$  are correlated, the local interaction terms  $w_j$  in Equation 4 have  $p = 4$ :

$$\begin{aligned} W_1(\{\mathbf{r}(\tau)\}) &= \sum_{j=0}^{n-4} w_j[\mathbf{r}(\tau_j), \mathbf{r}(\tau_{j+1}), \dots, \mathbf{r}(\tau_{j+4})] \\ &= w_2^{(1)}(\phi_2) + \sum_{j=2}^{n-1} w_j^{(2)}(\phi_j, \phi_{j+1}). \end{aligned} \quad 10.$$

Here  $w_2^{(1)}(\phi_2)$  is the potential of mean force for a single rotation angle  $\phi_2$ , and  $w_j^{(2)}(\phi_j, \phi_{j+1})$  is the potential of mean force for a pair of sequential rotation angles. The term  $w_2^{(1)}$  for  $\phi_2$  is necessary in addition to the  $w_j^{(2)}$  terms because  $\phi_2$  has no predecessor; the rotation angle of the first internal bond,  $\phi_1$ , is undefined. In this representation, the partition function is given by the summation of the local interaction Boltzmann factor  $\exp(-W_1)$  over all possible combinations of bond rotation angles,

$$Q = \sum_{k_2=1}^s \sum_{k_3=1}^s \cdots \sum_{k_n=1}^s \left\{ \exp[-w_2^{(1)}(\phi_2^{(k_2)})] \prod_{j=2}^{n-1} \exp[-w_j^{(2)}(\phi_j^{(k_j)}, \phi_{j+1}^{(k_{j+1})})] \right\}. \quad 11.$$

Using one-dimensional Ising model (71) matrix methods (49, 75), Equation 11 can be expressed more conveniently as

$$Q = (1 \ 0 \ 0 \ \dots \ 0) \left( \prod_{j=1}^{n-1} \mathbf{U}_j \right) \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = (1 \ 0 \ 0 \ \dots \ 0) \mathbf{U}^{n-2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \quad 12.$$

In this expression, both the row and column matrices consist of  $s$  elements, and  $\mathbf{U}_j$  is an  $s \times s$  matrix of statistical weights, representing relative populations of all possible rotational isomers for monomer  $j$  in the sequence. The matrix elements of  $\mathbf{U}_j$  are given by  $(\mathbf{U}_2)_{1,k} = \exp[-w_2^{(1)}(\phi_2^{(k)})]$ ,  $(\mathbf{U}_2)_{k,k'} = 0$  for  $k > 1$ , and  $(\mathbf{U}_j)_{k,k'} = \exp[-w_j^{(2)}(\phi_j^{(k)}, \phi_{j+1}^{(k')})]$  for  $j > 2$ . The last equality in Equation 12 holds if all monomer units are identical (i.e. in a homopolymer) and the numbering of the rotational isomeric states is such that  $w_2^{(1)}(\phi_2^{(k)}) = w_j^{(2)}(\phi_j^{(1)}, \phi_{j+1}^{(k)})$ . For a symmetric chain with three rotational isomers, such as  $n$ -alkanes,  $\mathbf{U}$  takes the form

$$\mathbf{U} = \begin{matrix} & t & g^+ & g^- \\ \begin{matrix} t \\ g^+ \\ g^- \end{matrix} & \begin{pmatrix} 1 & \sigma & \sigma \\ 1 & \sigma\psi & \sigma\omega \\ 1 & \sigma\omega & \sigma\psi \end{pmatrix} \end{matrix}, \quad 13.$$



where  $t$  denotes the *trans* state and  $g^\pm$  denotes the two *gauche* states;  $\sigma$ ,  $\sigma\psi$ , and  $\sigma\omega$  are statistical weights of combinations of these rotational states, the details of which can be found elsewhere (49).

The methods described above provide the basis for the helix-coil transition theories widely applied to polypeptides in solution (49, 81, 108, 110, 150). For example, in the Zimm-Bragg model (150), the two possible conformations of a bond pair are taken to be helix (h) or coil (c). Then  $\mathbf{U}$  becomes

$$\mathbf{U} = \begin{matrix} & \begin{matrix} c & h \end{matrix} \\ \begin{matrix} c \\ h \end{matrix} & \begin{pmatrix} 1 & \sigma s \\ 1 & s \end{pmatrix} \end{matrix}, \quad 14.$$

where  $\sigma$  is a helix initiation equilibrium constant, representing the bond-pair conformation  $\dots ch\dots$ , and  $s$  is a helix-coil equilibrium constant for an h unit preceded by either h or c. Thus, the helix-coil partition function is given by substituting the matrix in Equation 14 into Equation 12.

From the partition function, various configurational averages and their temperature dependences are obtained using standard methods (49, 150). If the helical configuration of a bond pair is energetically favored (for example by hydrogen bonding between monomers  $i$  and  $i+3$ , as in the  $\alpha$ -helix), and if the coil configuration of the bond pair is entropically favored by having more accessible states than the helix, then the helix will be stable at low temperatures and will undergo a transition to the coil state at high temperatures. The transition between helix and coil will be sharp if the chain is long. This type of theory has been highly successful in predicting solvent and temperature dependences of helix stabilities of polymers and peptides (110).

Thus the helix-coil behavior of a chain is governed by factors local in the sequence because the chain partition function is a product of the elementary partition functions of bonded neighbors taken pairwise, and because the use of this product of bond-pair partition functions is based on the assumption that the configuration of any bond pair is independent of any other.

### *Nonlocal Interactions*

Flory (47, 48) was the first to appreciate the importance of nonlocal interactions in governing polymer conformations. He made a simple argument to show that a homopolymer molecule should be more expanded in an inert solvent (e.g. a solvent of chain monomers) than would be predicted by the random-flight model. Consider a random-flight conformation. Any

two particular monomers  $i$  and  $j$  distant in the sequence are not likely to be close in space. But a long chain contains many monomer pairs. On balance, a random flight is likely to have at least one steric violation, whereby some pair of monomers would occupy the same volume element in space. Because this is physically impossible for a real chain, the random-flight model errs and requires correction. Such physical violations will be more common in compact conformations. Therefore, steric exclusions will eliminate more of the compact conformations than open ones, leading to a net expansion of the average molecular radius relative to the random flight prediction. The distribution function,  $P$  in Equation 7 above, accounts only for local interactions of connected neighboring residues along the chain and neglects these consequences of nonlocal interactions among monomers distant in the sequence.

The first and simplest calculation for the magnitude of this effect was given by Flory based on the Flory-Huggins lattice model (46–48, 69, 70). If the local and nonlocal interactions are independent, i.e. have additive free energies, then the effects of excluded volume can be described by a factor,  $\omega_{\text{ex}}$ , that multiplies the random-flight distribution to give the number of accessible conformations of the chain. The required excluded-volume factor is defined as the ratio of the number of conformations,  $\Omega$ , that are accessible when one considers excluded volume, divided by the number of conformations,  $\Omega_0$ , accessible when excluded volume is ignored, i.e.  $\omega_{\text{ex}} = \Omega/\Omega_0$ . Assuming that the monomers can be distributed uniformly throughout the available ( $d$ -dimensional) volume  $R^d$ , this ratio is calculated by parceling the volume occupied by the chain into  $m = R^d/v$  sites, each having a volume  $v$  equal to that of one of the  $n$  chain segments. The excluded volume contribution to the distribution function can be calculated through use of a fictitious process whereby each segment is considered to be inserted, one at a time, into a lattice of sites. The first chain segment can be inserted into any one of  $m$  accessible sites. The second chain segment can only be inserted into any of the remaining  $m - 1$  sites. The third segment can be inserted into any of the remaining  $m - 2$  sites, and so on, until the final insertion of chain segment  $n$  occurs into any of the remaining  $m - n + 1$  sites. Thus the total number of configurations resulting from this insertion process is

$$\Omega = m(m-1)(m-2)\cdots(m-n+1) = \frac{m!}{(m-n)!}. \quad 15.$$

This quantity must be divided by the number of configurations accessible to the chain if excluded volume were ignored. In that case, the first segment could occupy any of the  $m$  sites. If excluded volume is neglected, then the

second and all subsequent segments will also have access to any of the  $m$  sites, hence  $\Omega_0 = m^n$ . Thus the ratio required to correct for the excluded volume contribution to the number of accessible conformations is

$$\omega_{\text{ex}} = \frac{\Omega}{\Omega_0} = \frac{m!}{m^n(m-n)!} = \frac{(n/\rho)!}{(n/\rho)^n(n/\rho-n)!}, \quad 16.$$

where the latter equality follows from the definition  $\rho = n/m$ . If we use Stirling's approximation  $x! = \sqrt{2\pi x}(x/e)^x$  and the Boltzmann-Planck equation  $S = k \ln \Omega$ , then this contribution to the conformational entropy is

$$\frac{\Delta S_{\text{ex}}(\rho)}{nk} = - \left[ \left( \frac{1-\rho}{\rho} \right) \ln(1-\rho) + 1 \right], \quad 17.$$

where terms of order  $n^{-1}$  are neglected. This formula estimates the effect of nonlocal steric repulsions on chain conformational entropy.

There is a second nonlocal contribution if the solvent is not inert. This contribution is described in terms of the relative interactions of monomer-monomer, monomer-solvent, and solvent-solvent, expressed by the Flory  $\chi$  parameter (48). The quantity  $\chi$  is the free energy (divided by  $2kT$ ) for the following thermodynamic exchange process: one monomer from a pure medium of monomers is transferred into a pure solvent and one solvent molecule is similarly transferred from the pure solvent into a medium of pure monomers. Combined with the nonlocal excluded volume, the total nonlocal free energy in the Flory model is

$$\frac{\Delta F_{\text{nonlocal}}}{nkT} = \left( \frac{1-\rho}{\rho} \right) \ln(1-\rho) + 1 - \chi\rho. \quad 18.$$

Expansion of the nonlocal free energy in a Taylor series around the low-density  $\rho = 0$  conformations yields

$$\frac{\Delta F_{\text{nonlocal}}}{kT} = n\rho \left[ \frac{1}{2} - \chi \right] + \sum_{s=2}^{\infty} \frac{n\rho^s}{s(s+1)}. \quad 19.$$

If only the term linear in  $\rho$  is kept ( $\rho = nv/R^d$  corresponds to the probability of forming a monomer-monomer contact), the sum of local (Equation 8) and nonlocal contributions (Equation 19) gives the total free energy

$$\frac{\Delta F_{\text{total}}}{kT} = \frac{\Delta F_{\text{el}}}{kT} + \frac{\Delta F_{\text{nonlocal}}}{kT} = \frac{d}{2}(\alpha^2 - 1) - d \ln \alpha + \frac{n^2v}{\alpha^d R_0^d} \left( \frac{1}{2} - \chi \right), \quad 20.$$

in terms of  $R = \alpha R_0$ . By setting the derivative  $\partial \Delta F_{\text{total}} / \partial \alpha = 0$ , we obtain the equilibrium expansion,  $\alpha$ , of the molecule. The Flory result generalized

to  $d$  dimensions is

$$\alpha^{d+2} - \alpha^d = \frac{n^2 v}{R_0^d} \left( \frac{1}{2} - \chi \right) \propto n^{2-d/2} \left( \frac{1}{2} - \chi \right), \quad 21.$$

where the second step follows from the random-flight relation  $R_0 \sim n^{1/2}$ .

The important conclusion from this result is that it defines three types of solvent, depending on their effects on the chain expansion. First, the so-called Flory theta solvent is defined as that solvent for which  $\chi = 1/2$ , or alternatively the Flory theta temperature is that temperature, for a given solvent, for which  $\chi = 1/2$ . For the theta solvent, Equation 21 shows that  $\alpha = 1$ , and the chain adopts its ideal unperturbed distribution. This solvent is somewhat unfavorable, such that chain contraction resulting from monomer-monomer attractions (relative to monomer-solvent attractions) exactly cancels the expansion resulting from nonlocal steric repulsions; thus in the theta solvent, only the local interactions remain.<sup>1</sup> The second is referred to as a “good” solvent,  $\chi < 1/2$  ( $\chi = 0$  for an inert solvent), because then, for sufficiently long chains and  $d < 4$ , the  $(d+2)$ th power term clearly dominates the left side of Equation 21,  $\alpha^{d+2} \sim n^{2-d/2}$ , resulting in an expanded radius relative to the ideal chain,

$$R = \alpha R_0 \sim [n^{(2-d/2)/(d+2)}] [n^{1/2}] = n^{3/(d+2)}. \quad 22.$$

This expression reduces to  $R \sim n$ , in  $d = 1$  dimension,  $R \sim n^{3/4}$  in  $d = 2$  dimensions, and the well-known result that  $R \sim n^{3/5}$  in  $d = 3$  dimensions. Equation 21 shows that for dimensions  $d \geq 4$ , effects of excluded volume are negligible; for all  $d > 4$ ,  $\alpha \rightarrow 1$  with increasing  $n$ , and for  $d = 4$ ,  $\alpha$  is independent of  $n$ , indicating that the chain is ideal in four dimensions, i.e.  $R = \alpha R_0 \sim n^{1/2}$ . Thus, in a good solvent (for  $d < 4$ ), the balance of nonlocal interactions is such that the excluded volume expansion exceeds the solvent-driven contraction. The third case is referred to as a “poor” solvent,  $\chi > 1/2$ , because the monomer-monomer attractions are favored so strongly relative to monomer-solvent attractions that  $\alpha < 1$  and the chain contracts. Thus in a poor solvent, the solvent-driven contraction exceeds the excluded volume expansion, and, as  $n \rightarrow \infty$ , the radius becomes zero. Clearly, the model is unphysical in this latter case because the chain radius cannot diminish beyond the steric size dictated by the sum of the monomer sizes. Moreover, as the polymer collapses to its minimum size,  $\rho$  approaches 1, which violates the small- $\rho$  requirement of Equation 20. This

<sup>1</sup> The conclusion that solvent interactions and excluded volume cancel exactly is only valid to the second order, i.e. to the level of the two-body interactions. The second virial coefficient is zero, but the third virial term is not (54, 119, 146). Hence, exact cancellation does not occur, so the distribution will not be exactly Gaussian in any solvent condition.

situation is repaired by including the three-body repulsive interaction, which is the basis for the simplest homopolymer collapse theories (see below).

### *Exhaustive Lattice Simulations*

The Flory approximation, Equation 16, estimates the effect of excluded volume by supposing that the chain monomers are distributed uniformly in the available volume; in this way, it neglects the severity of steric constraints among monomers close together in the sequence that is greater than among those far apart in the sequence. Despite its simplicity, it has led to remarkably accurate predictions of scaling behavior for chain expansion; for example, the exponent 0.6 for  $d = 3$  in Equation 22 is now known in better approximation to be 0.588 (78). One of the most important tools for refinement of the theory of polymer excluded volume has been the use of exhaustive simulations of self-avoiding chains on lattices. Such methods permit the prediction of long-wavelength properties of chains that depend little on microscopic details such as bond angles and local steric effects.

In 1947, W. J. C. Orr (101) first exhaustively enumerated all the conformations of short chains on simple lattices subject to excluded volume. He considered chains of lengths  $n \leq 9$  on two-dimensional square lattices and  $n \leq 7$  on three-dimensional simple cubic lattices. With advances in computer technology in the 1950s, Domb and his colleagues (34) extended the exhaustive simulations; when combined with sophisticated counting theorems of Sykes (133), this work led to improved enumeration methods for longer chains. The main emphasis of these systematic efforts has been the determination of scaling exponents, connective constants, and the distributions of various chain properties in the presence of excluded volume. The most thoroughly studied of these properties include the exponent  $\nu$  in the asymptotic scaling of the mean square end-to-end distance  $\langle R_n^2 \rangle \sim n^{2\nu}$  (33, 44), the exponent  $\gamma$  and the connective constant  $\mu$  in the asymptotic form of the total number of conformations  $Q_0(n) \sim n^\gamma \mu^n$  (44, 134), the exponent  $g$  in the total number of end-to-end cycles or rings  $Q(n; 1, n) \sim n^g \mu^n$  [which is equivalent to  $Q(N; 1, N+1)$  (11), where  $N = n - 1$  is the number of bonds while  $n$  is the number of monomers] (67, 83, 118), and the distribution of end-to-end vectors  $\mathbf{R}_n$  (36). The study of lattice self-avoiding walks (3) has also been extended to incorporate solvent interactions (44) and nearest-neighbor attractive energies (86) to account for monomer-monomer association. One noteworthy finding is that even at theta point when  $\langle R_n^2 \rangle$  scales as  $n$ , the distribution of  $\mathbf{R}_n$  is not Gaussian (36, 43, 44, 87, 140, 141). Thus no attraction between monomers can fully compensate for the excluded volume effects (see also Footnote 1). These

exhaustive simulation studies contributed significantly to the development of modern polymer theory, including scaling (24) and renormalization group methods (20, 22, 23, 26, 27, 54, 55, 95–99). Recently, exhaustive simulations have also been applied to chains with several constraints as a function of compactness (9–12), as well as to compact chains confined to specific shapes (10, 12, 19). Exhaustive simulations are of special importance to the study of compact chains because the efficiency and ergodicity of Monte Carlo samplings of compact conformations are problematic (58, 103, 104). Compact conformations are of interest because of their relevance to globular native states of proteins.

### *The Path-Integral Method*

More satisfactory than the Flory approximation for the nonlocal interactions is the self-consistent field path-integral method, introduced by S. F. Edwards (38). In addition to self-consistent field considerations, the path integral is often analyzed using perturbation theory, which is applicable to chains with few intrachain contacts (53, 54). This approach is important for proteins for at least two reasons. First, near the theta point, it currently provides the most accurate description of excluded volume effects and may provide the best approach for refining polymer collapse theories, or at least the initial stages of collapse from the open conformations. Second, it provides an off-lattice alternative for the study of internal architectures in polymers and proteins (see below).

In general, the nonlocal potential of mean force  $W_{nl}$  in Equation 5 is expressed in terms of two-monomer, three-monomer, and higher many-monomer interaction terms as follows:

$$W_{nl}(\{\mathbf{r}(\tau)\}) = \sum_{i < j} u^{(2)}[\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j)] + \sum_{i < j < k} u^{(3)}[\mathbf{r}(\tau_i), \mathbf{r}(\tau_j), \mathbf{r}(\tau_k)] + \dots \quad 23.$$

In the path integral approach, interactions among connected neighbors along the chain sequence are usually assumed to obey random-flight statistics (Equation 6). In many applications, the series for  $W_{nl}$  in Equation 23 is truncated to retain only the binary term,

$$W_{nl}(\{\mathbf{r}(\tau)\}) = \sum_{i < j} u[\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j)]. \quad 24.$$

This restriction is equivalent to the superposition approximation in the theory of simple fluids (68).

When the monomer-monomer interactions  $u(\mathbf{r})$  are spatially short-ranged, then  $u$  is sharply peaked at  $\mathbf{r} = \mathbf{0}$ . Thus the pair interaction is usually approximated by the Dirac  $\delta$ -function, which may be expressed either in terms of the Mayer function  $f_{ij}$ ,

$$\exp[-u(\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j))] = 1 - \beta_0 \delta[\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j)] = 1 + f_{ij}, \quad 25.$$

as in the Ursell-Mayer theory of nonideal gas (85, 107), first introduced by Zimm (149), or in the form

$$u[\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j)] = \beta_0 \delta[\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j)]. \quad 26.$$

In both cases,  $\beta_0$  is a constant determined by the strength of interaction of the chain and its environment. Provided  $\beta_0$  is small and the  $\delta$ -functions are handled appropriately to avoid unphysical singularities, Equations 25 and 26 are equivalent in perturbation theory, as noted by Fixman (45).

Fixman studied chains with excluded volume ( $\beta_0 > 0$ ) (45) by using a method resembling the Mayer cluster expansion technique (85, 107), which consists of expanding the interaction term as

$$\begin{aligned} \exp[-W_{nl}(\{\mathbf{r}(\tau)\})] &= \prod_{i < j} (1 + f_{ij}) = 1 + \sum_{i < j} f_{ij} + \underbrace{\sum_{i < j} \sum_{l < m} f_{ij} f_{lm}}_{i < l} + \cdots \\ &+ \underbrace{\sum_{i < j} \sum_{l < m} \cdots \sum_{r < s} f_{ij} f_{lm} \cdots f_{rs}}_{i < l < \cdots < r} + \cdots. \quad 27. \end{aligned}$$

Because each Mayer function  $f_{ij}$  is associated with a two-segment excluded volume interaction, a term with  $k$   $f_{ij}$  factors is referred to as a  $k$ -contact term. In this way, the partition function  $Q$  in Equation 1 is given by the series

$$\begin{aligned} Q &= 1 - \beta_0 \sum_{i < j} P_0(\mathbf{0}_{ij}) + \beta_0^2 \sum_{\substack{i < j < m \\ i < l}} P_0(\mathbf{0}_{ij}) P_0(\mathbf{0}_{lm}) + \cdots \\ &+ (-\beta_0)^k \sum_{\substack{i < j < m < r \\ i < l < \cdots < r}} \sum_{\substack{r < s \\ k \text{ os}}} P_0(\mathbf{0}_{ij}, \mathbf{0}_{lm}, \dots, \mathbf{0}_{rs}) + \cdots, \quad 28. \end{aligned}$$

obtained by substitution of Equations 27 and 6 into Equation 5 then into Equation 2 for  $P(\{\mathbf{r}(\tau)\})$ . Because each Mayer function  $f_{ij}$  contains a  $\delta$ -function of the spatial separation between  $\mathbf{r}(\tau_i)$  and  $\mathbf{r}(\tau_j)$ , straightforward integrations of the  $f_{ij}$ s in Equation 27 with the Gaussian bond probability (Equation 6) give rise to factors of  $P_0$ s in Equation 28, which are random-flight (unperturbed,  $\beta_0 = 0$ ) probabilities for the chain to adopt conformations with specific self-contacts. For example,  $P_0(\mathbf{0}_{ij})$  is the unperturbed probability that segments  $i$  and  $j$  are in contact, given by the Jacobson-Stockmayer (72) factor

$$P_0(\mathbf{0}_{ij}) = \left[ \frac{d}{2\pi l^2 |i-j|} \right]^{d/2} \quad 29.$$

Similarly  $P_0(\mathbf{0}_{ij}, \mathbf{0}_{lm}, \dots, \mathbf{0}_{rs})$  is the unperturbed probability for the chain to have the contact pairs  $(i, j)$ ,  $(l, m)$ , . . . ,  $(r, s)$ , which may readily be computed by the Wang-Uhlenbeck theorem (142). The terms in expansions such as Equation 28 are often represented by cluster diagrams (Figure 3), a convenient device for combinatorics bookkeeping. In Figure 3, a solid line represents the chain and each dotted line indicates the pair of segments that interact with repulsion  $\beta_0$ . Theories of this type are called two-parameter models because they depend only on the parameters  $nl^2$  and  $n^2\beta_0$ , the unperturbed mean square end-to-end distance and the total excluded volume among chain segments, respectively (148). The main result is that the partition function in Equation 28 now contains not only the local interactions, but also contains, in series expansion, the two-, three-, four-self-contact contributions, etc, from the nonlocal interactions.

The transition from the discrete bonds to the continuum path integral formulation is straightforward. The procedure consists of letting the segment lengths shrink to zero ( $\Delta\tau_j \equiv \tau_j - \tau_{j-1} \rightarrow 0$ ) and the number of segments increase to infinity while keeping the total contour length  $N \equiv \sum_{j=1}^n \Delta\tau_j$  and the step length  $l \equiv \langle R_n^2 \rangle / N$  finite. Despite the fact that real chains are comprised of discrete segments, adoption of a continuum model has advantages. Even in the discrete two-parameter model, final summations in expressions such as Equation 28 are often approximated by integrals (148). Also, the continuum path integral has been applied extensively in quantum mechanics and quantum field theory (40, 41, 57, 117), so a whole arsenal of techniques is available, including renormalization group methods (54, 147).

Using the bond probability (Equation 6) and excluded volume interaction (Equation 26), the partition function is given by

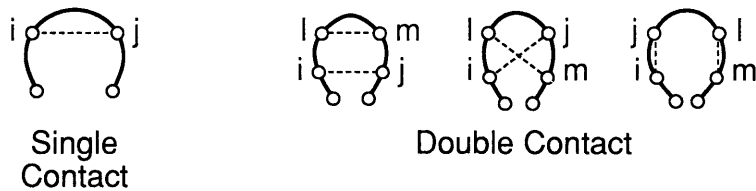


Figure 3 Cluster diagrams for the single- and double-contact terms. Note the three possibilities in the order of arrangement of the four indices  $i, j, l$ , and  $m$  under the conditions  $i < j, l < m$ , and  $i < l$ .



$$Q = \mathcal{N} \int d\mathbf{r}(\tau_0) \int d\mathbf{r}(\tau_1) \dots \int d\mathbf{r}(\tau_n) \exp \left\{ -\frac{d}{2l^2} \sum_{j=1}^n |\mathbf{r}(\tau_j) - \mathbf{r}(\tau_{j-1})|^2 - \frac{\beta_0}{2} \sum_{\substack{i=0 \\ i \neq j}}^n \sum_{j=0}^n \delta[\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j)] \right\}, \quad 30.$$

where the factors of  $(d/2\pi l^2)^{d/2}$  are absorbed into the normalization  $\mathcal{N}$ , and the summation  $\sum_{i < j}$  in Equation 24 is replaced by the symmetric summation  $1/2 \sum_i \sum_j$ . Because  $l = \Delta\tau_j$  for discrete chains (with  $N = nl$ ), and because, by the definition of  $l$ , the mean-square length of a generalized segment is  $l\Delta\tau_j$ , the continuum limit of the bond probability term in the exponent is

$$\begin{aligned} \lim_{\Delta\tau_j \rightarrow 0} -\frac{d}{2l^2} \sum_{j=1}^n |\mathbf{r}(\tau_j) - \mathbf{r}(\tau_{j-1})|^2 &= \lim_{\Delta\tau_j \rightarrow 0} -\frac{d}{2l} \sum_{j=1}^n (\Delta\tau_j) \left| \frac{\mathbf{r}(\tau_j) - \mathbf{r}(\tau_{j-1})}{\Delta\tau_j} \right|^2 \\ &= -\frac{d}{2l} \int_0^N d\tau \left| \frac{d\mathbf{r}(\tau)}{d\tau} \right|^2. \end{aligned} \quad 31.$$

In the excluded volume interaction term in Equation 30,  $\beta_0$  is the repulsive strength per chain segment. Therefore, to generalize to the continuum case,  $\beta_0/l^2$  is identified as the repulsive strength per square unit of continuous contour length  $\tau$ . Hence, by substituting the equality  $l = \Delta\tau_j$  for discrete chains, the interaction term may be rewritten as

$$-\frac{\beta_0}{2l^2} \sum_{\substack{i=0 \\ i \neq j}}^n \Delta\tau_i \sum_{j=0}^n \Delta\tau_j \delta[\mathbf{r}(\tau_i) - \mathbf{r}(\tau_j)]. \quad 32.$$

The continuum limit is obtained by replacing the double summation by integrals,

$$-\frac{\beta_0}{2l^2} \underbrace{\int_0^N d\tau \int_0^N d\tau'}_{|\tau - \tau'| \geq a} \delta[\mathbf{r}(\tau) - \mathbf{r}(\tau')], \quad 33.$$

where the phenomenological constant  $a (\simeq l)$  is introduced to prevent the unphysical situation of a chain segment interacting with itself. The *path integral* or *functional integral*  $\int \mathcal{D}[\mathbf{r}(\tau)]$  is defined as the  $|\tau_j - \tau_{j-1}| \rightarrow 0$  limit of the multiple integrals over  $d\mathbf{r}(\tau)$ s,

$$\int \mathcal{D}[\mathbf{r}(\tau)] \equiv \lim_{\substack{n \rightarrow \infty \\ \Delta\tau_j \rightarrow 0}} \prod_{j=0}^n \int d\mathbf{r}(\tau_j). \quad 34.$$

This is sometimes schematically represented as the product  $\prod_{\tau} \int d\mathbf{r}(\tau)$  over continuous values of  $\tau$ . Collecting continuum limits from Equations 31, 33, and 34, we arrive at the path-integral partition function

$$Q = \mathcal{N} \int \mathcal{D}[\mathbf{c}(\tau)] \exp \{ -H[\mathbf{c}(\tau)] \},$$

$$H[\mathbf{c}(\tau)] = \frac{1}{2} \int_0^N dt \left| \frac{d\mathbf{c}(\tau)}{dt} \right|^2 + \frac{v_0}{2} \underbrace{\int_0^N dt \int_0^N dt' \delta[\mathbf{c}(\tau) - \mathbf{c}(\tau')]}_{|\tau - \tau'| \geq a}. \quad 35.$$

Here  $v_0 \equiv \beta_0 d^{d/2} / l^{2+d/2}$  is the repulsive excluded volume strength. The variable  $\mathbf{c}(\tau) \equiv (d/l)^{1/2} \mathbf{r}(\tau)$  is adopted instead of  $\mathbf{r}(\tau)$  because then the path integral does not contain the spatial dimension  $d$  explicitly, which is often useful.

When  $v_0$  is small—for chains near the theta condition—conformational properties may be deduced from the continuum path integral (Equation 35) by systematic perturbation expansion in powers of  $v_0$ . For instance, Muthukumar & Nickel (91) determined the mean-square end-to-end distance  $\langle R_n^2 \rangle$  through the sixth order in  $v_0$  for polymers with excluded volume in three dimensions. A detailed discussion of the Feynman-diagrammatic (40, 41) techniques of the perturbation procedure can be found elsewhere (9, 54).

Thus the two main avenues to refined modeling of nonlocal interactions in polymer chains beyond the Flory approximation are lattice simulations and continuum path integral methods. Each method has its advantages. Lattices offer more realistic treatment of local excluded volume and effects of finite monomer size. Finite-volume monomer units and finite-ranged infinite repulsive potentials are difficult to implement in analytic models: the path integral model treats chains as mathematical curves with vanishing volume, and excluded volume is modeled by a two-body  $\delta$ -function repulsion of vanishing range. On the other hand, for properties that are not sensitive to local details, the path integral approach is not restricted to short chains nor is it constrained as a lattice is to fixed bond angles and periodic spatial arrangements of monomers. A review by McKenzie provides an insightful discussion of the similarities and differences between lattices and analytic continuum models (88).

### *Homopolymer Collapse Theories*

As noted earlier, the Flory model predicts that homopolymers will collapse to compact states in poor solvents (129). However, until the work of