6076 *Macromolecules* **1993,** *26,* 6076–6080

Calculation of Conformational Free Energy and Entropy of Chain Molecules with Long-Range Interactions

Olivier Collet and Samüel Premilat*

Laboratoire de Biophysique Moleculaire, Universite de Nancy I, Faculte des Sciences, URA CNRS 494, Bd des Aiguillettes, BP No. 239, 54506 Vandoeuvre les Nancy, France

Received April 29, 1993; Revised Manuscript Received July 26, 1993⊛

ABSTRACT: The probability that a macromolecular chain occurs in a given conformation is expressed as a product of mean probabilities obtained from Monte Carlo calculations. It permits us to develop a method for the determination of the conformational entropy of chains with long-range interactions. This method has been applied successfully to polymethylene chains with respectively short-range or short-range and hard-sphere interactions. Good results are also obtained for short chains with van der Waals interactions. For the three different potentials used, conformational free energy, energy, and entropy are linear functions of the number of bonds in the chain. The present method gives results in agreement with Monte Carlo sampling when the ratio of the long-range energy to the short-range one is smaller than about 1.5. This method of calculation has also been applied to the determination of chain geometrical dimensions, but it appears that it does not work for that purpose.

I. Introduction

Conformations of macromolecular chains depend very much on entropy which is associated with the order and, as a component of the free energy, with the stability of the system. Go and Scheraga[1] developed a method to calculate the entropy of macromolecules undergoing harmonic fluctuations around their stable states. This method was applied on several short polypeptide chains.[2] Karplus and Kushick[3] suggested calculating the covariances of the internal coordinates directly from molecular dynamics or Monte Carlo simulations rather than by normal modes analysis and applied their method to the molecular dynamics simulation of decaglycine and butane. More recently, a method for the calculation of the entropy of chain models based on a scanning procedure[4] has been proposed by Meirovitch and applied to models of self-avoiding walks on a square or a cubic lattice. This approach has been extended and applied to general chain models with excluded volume, to self-attractive random walks,[5] and to polypeptide chains.[6]

In the present work we propose to calculate values of the conformational entropy of macromolecular chains submitted to short-range interactions. The entropy can then be calculated by using transition probability matrices.[7] The application of this method is then extended to models of polymers submitted to long-range interactions (van der Waals or hard-sphere) in a manner similar to that proposed by Mattice[8,9] to reproduce geometrical properties of perturbed chains. In order to establish the range of validity of the approach, values of entropy are also determined by using Bennett's method[10] or an umbrella sampling algorithm recently proposed and used for the calculation of the conformational free energy of short polypeptide chains.[11] In the present work calculations are performed on a three rotational isomeric state model (t;g+;g−) of polymethylene.

II. Methods of Calculation

Atomic coordinates of polymethylene chains are computed using the method of Eyring[12] with a standard[7] C–C bond length of 1.53 Å and a fixed valence angle of 109°.

Three different situations of the chain molecule are considered in the following applications. They correspond respectively to three different potentials used for the conformational energy calculation. Firstly, when only short-range interactions are introduced, the chain is at the Θ point of Flory.[7] But if there are attractive interactions between the molecule and a good solvent, the effects of the latter can compensate for the attraction between atoms of the chain. The molecule therefore adopts extended conformations. A hard-sphere potential, with only a repulsive term, is well adapted, in that case, to simulate these long-range effects on the chain. On the contrary, when the chain atom groups and a bad solvent repel each other, intramolecular long-range interactions are dominant in the chain. A van der Waals potential can be used to simulate such a behavior of the chain molecule which will then adopt more folded conformations.

II.1. Chains with Short-Range Interactions. According to the statistical mechanics developed by Flory for the study of chain molecules with interdependent rotational potentials,[7] the short-range conformational energy of a chain molecule composed of n units can be expressed as a sum of first-neighbor pair energies $E_{\mathrm{kl};i}$:

$$E_{\mathrm{sh}} = \sum_{i=2}^{n-1} E_{\mathrm{kl};i}$$

where k is the isomeric state of bond $i-1$ and l that bond i.

Let us introduce the matrix **U** necessary for the determination of the partition function of configuration,

$$Z_{\mathrm{sh}} = J^{*}\mathbf{U}^{n-2}J$$

with

$$J^{*} = (1, 0, 0)$$

$$J = \begin{vmatrix} 1 \\ 1 \\ 1 \end{vmatrix}$$

and the 3 × 3 matrix **U**:

$$\begin{vmatrix} 1 & \sigma & \sigma \\ 1 & \sigma & \sigma \mathrm{w} \\ 1 & \sigma \mathrm{w} & \sigma \end{vmatrix}$$

We presently use[13,14] $RT = 0.866$ kcal/mol ($T = 433$ K), $\sigma = 0.56$, and $\sigma\mathrm{w} = 0.03$ and thus $E_{\mathrm{g}} = -RT \ln \sigma = 0.5$ kcal/mol; $E_{\mathrm{w}} = -RT \ln \sigma\mathrm{w} = 3.0$ kcal/mol.

The conformational free energy F_{sh}, the energy E_{sh}, and the entropy TS_{sh} are then given by the following relations:

⊛ Abstract published in *Advance ACS Abstracts*, October 1, 1993.

0024-9297/93/2226-6076$04.00/0

$$F_{\mathrm{sh}} = -RT \ln Z_{\mathrm{sh}} \tag{1}$$

$$\langle E_{\mathrm{sh}} \rangle = (RT^2/Z_{\mathrm{sh}}) J^* \left(\sum_{j=0}^{n-3} \mathbf{U}^j (\mathrm{d}\mathbf{U}/\mathrm{d}T) \mathbf{U}^{n-3-j} \right) J \tag{2}$$

$$TS_{\mathrm{sh}} = \langle E_{\mathrm{sh}} \rangle - F_{\mathrm{sh}} \tag{3}$$

If P_C is the probability that a chain with short-range interactions occurs in configuration C, we have

$$TS_{\mathrm{sh}} = -RT \sum_{\{C\}} P_C \ln P_C \tag{4}$$

where S_{sh} is the contribution of the short-range interactions to the conformational entropy of the molecule. The summation is made over all possible configurations of the chain.

Since here $E_{\mathrm{sh}} = \sum E_{\mathrm{kl};i}$, we can write

$$P_C = \prod_{i=2}^{n} \mathbf{Q}^c_{\mathrm{kl};i} / Z_Q$$

with

$$Z_Q = J^* \left(\prod_{i=2}^{n} \mathbf{Q}_i \right) J$$

where $\mathbf{Q}_i$ is a 3 × 3 matrix with the elements $\mathbf{Q}_{\mathrm{kl};i}$ which are the conditional probabilities that bond i is in state l while bond $i - 1$ is in state k. So we have[7] $\mathbf{Q}_{\mathrm{kl};i} = P_{\mathrm{kl};i}/P_{\mathrm{k};i-1}$ where $P_{\mathrm{k};i-1}$ is the a priori probability that bond $i - 1$ is in state k, and $P_{\mathrm{kl};i}$ is the probability that bond $i - 1$ is in state k and bond i is in state l. One therefore has

$$P_{\mathrm{k};i-1} = \sum_{\alpha = \mathrm{t,g+,g-}} P_{\alpha\mathrm{k};i-1}$$

So the entropic term can be written as

$$TS_{\mathrm{sh}} = -RT \sum_{\{C\}} \left(Z_Q^{-1} \prod_{i=2}^{n} \mathbf{Q}^c_{\mathrm{kl};i} \right) \ln \left(Z_Q^{-1} \prod_{j=2}^{n} \mathbf{Q}^c_{\mathrm{kl};j} \right)$$

$$= -RT \sum_{\{C\}} Z_Q^{-1} \left(\prod_{i=2}^{n} \mathbf{Q}^c_{\mathrm{kl};i} \right) \left(\sum_{j=2}^{n} \ln \mathbf{Q}^c_{\mathrm{kl};j} \right) + RT \ln Z_Q$$

Let us introduce a 3 × 3 matrix $\mathbf{Q}'_j$ with the elements ($\mathbf{Q}_{\mathrm{kl};j} \ln \mathbf{Q}_{\mathrm{kl};j}$). The sum of the mean values of the elements of the matrix [$\ln \mathbf{Q}_j$] over the chain configurations is then given by

$$L_j = \mathbf{Z}_Q^{-1} J^* \left(\prod_{i=2}^{j-1} \mathbf{Q}_i \right) \mathbf{Q}'_j \left(\prod_{i=j+1}^{n} \mathbf{Q}_i \right) J$$

so, we have

$$TS_{\mathrm{sh}} = -RT \sum_{j=2}^{n} L_j + RT \ln Z_Q \tag{5}$$

For chains submitted only to short-range interactions, the a priori probabilities $P_{\mathrm{kl};i}$ are easily obtained, following the method proposed by Flory,[7] from the matrix **U** and related matrices.

II.2. Conformational Entropy of Chains with Long-Range Interactions. Now, let us consider a chain with short- and long-range interactions. In order to extend the use of relation 5 in that case, we have to calculate the values of $P_{\mathrm{kl};i}$ and $\mathbf{Q}_{\mathrm{kl};i}$ for this chain. To do this, note that since the probability of obtaining a chain configuration {C} with a total energy E_C is $P_C = \exp(-E_C/RT)/Z_T$ (Z_T is the partition function of the chain in total interaction), we can write

$$P_{\mathrm{kl};i} = \sum_{\{C\}} \delta_{\mathrm{kl};i} \exp(-E_C/RT)/Z_T \tag{6}$$

with $\delta_{\mathrm{kl};i} = 1$ if bond $i - 1$ is in state k and bond i is in state l and equals 0 otherwise.

This relation can be employed completely in practice only if the number of bonds in the chain remains very low.

II.2.1. Chain with Short-Range and Hard-Sphere Interactions. The conformational energy is now given by $E_T = E_{\mathrm{sh}} + E_{\mathrm{Hs}}$, where E_{Hs} corresponds to the hard-sphere potential.[19] We apply the method of calculation defined by relations 5 and 6 to chains of up to 100 bonds. The free energy of the chain is now calculated with a Monte Carlo procedure based on a method proposed by Bennett.[10] Since the thermodynamic behavior of a molecular system submitted only to short-range interactions is well known,[7] the corresponding free energy of conformation F_{sh} is easily calculable. To estimate the free energy $F_{\mathrm{sh+L}}$ for this system when long-range interactions are also taken into account, let us define $\Delta F = F_{\mathrm{sh+L}} - F_{\mathrm{sh}}$. We hence have

$$\Delta F = RT \ln(Z_{\mathrm{sh}}/Z_{\mathrm{sh+L}})$$

and from the properties of the "Metropolis function" $M(x) = \min\{1,\exp(-x)\}$, we get[10]

$$\frac{Z_{\mathrm{sh}}}{Z_{\mathrm{sh+L}}} = \frac{\langle M(\beta(E_{\mathrm{sh}} - E_{\mathrm{sh+L}})) \rangle_{\mathrm{sh+L}}}{\langle M(\beta(E_{\mathrm{sh+L}} - E_{\mathrm{sh}})) \rangle_{\mathrm{sh}}}$$

$\beta = 1/RT$. Brackets denote Monte Carlo averages calculated over the system under the effects of short-range and hard-sphere interactions (sh + Hs) or only short-range interactions (sh). Let f be the relative number of configurations which are accepted during this sampling procedure; they correspond to a long-range energy $E_{\mathrm{Hs}} = 0$. So $1 - f$ are disallowed ($E_{\mathrm{Hs}} = +\infty$). Under such conditions, we have

$$\langle M(\beta(E_{\mathrm{sh}} - E_{\mathrm{sh+Hs}})) \rangle_{\mathrm{sh+Hs}} = \langle (1 - f) \min\{1,+\infty\} + f \min\{1,1\} \rangle_{\mathrm{sh+Hs}}$$

$$= 1$$

$$\langle M(\beta(E_{\mathrm{sh+Hs}} - E_{\mathrm{sh}})) \rangle_{\mathrm{sh}} = \langle (1 - f) \min\{1,0\} + f \min\{1,1\} \rangle_{\mathrm{sh}}$$

$$= f$$

Hence the ratio of the partition functions becomes:

$$Z_{\mathrm{sh}}/Z_{\mathrm{sh+Hs}} = 1/f$$

and finally, $\Delta F = -RT \ln f$ and $F_{\mathrm{sh+Hs}} = F_{\mathrm{sh}} - \mathrm{RT} \ln f$.

The average value of the energy $E_{\mathrm{sh+Hs}}$ is obtained from a MMC (Metropolis Monte Carlo)[15,16] calculation. The sampling procedure presently used is a simple adaptation of the one established for the study of polypeptide chains.[17,18] By using the values of $F_{\mathrm{sh+Hs}}$ and $E_{\mathrm{sh+Hs}}$ we can determine $TS_{\mathrm{sh+Hs}}$. Note that with large values of n, relation 6 can no longer be used for a complete calculation. But $(n - 1)$ matrices P_i can be obtained by using the same Monte Carlo procedure from which the 3 × 3 elements $P_{\mathrm{kl};i}$ (with $2 \le i \le n - 1$) are now determined by

$$P_{\mathrm{kl};i} = \frac{1}{N} \sum_{j=1}^{N} (\delta_{\mathrm{kl};i})_j \tag{7}$$

N is the number of steps of the MMC sampling. We hence computed $TS_{\mathrm{sh+Hs}}$ using relations 5 and 7. Results obtained from these different methods of calculation can then be compared.

II.2.2. Chain with Short-Range and van der Waals Interactions. The validity of eqs 5 and 7 when long-range interactions are taken into account has also been tested on chains of up to 100 bonds submitted to short-range and long-range van der Waals interactions. The total energy is then $E_t = E_{\mathrm{sh}} + E_{\mathrm{VW}}$ where the van der Waals term is presently given the expression[20] $E_{\mathrm{VW}} = 864000/r^{12} - 940/r^6$.

A MMC sampling[17,18] is used to estimate the average energy $\langle E_t \rangle$ of the chain. The free energy difference between chains with short-range and van der Waals interactions (sh + VW) and chains with only short-range (sh) interactions is obtained from an umbrella sampling[11,21,22] based on the following relation giving the ratio of the partition functions of the molecule in these two situations:[23]

$$Z_{\mathrm{sh+VW}}/Z_{\mathrm{sh}} = \langle \exp(-\beta E_{\mathrm{VW}}) \rangle_{\mathrm{sh}}$$

The average is calculated over the chain conformations taking into account only short-range interactions. The free energy difference between these two chain models can therefore be estimated by

$$\Delta F = -RT \ln(Z_{\mathrm{sh+VW}}/Z_{\mathrm{sh}}) = -RT \ln \langle \exp(-\beta E_{\mathrm{VW}}) \rangle_{\mathrm{sh}}$$

But to compute such a mean value, the sample of chains should represent simultaneously the conformational space of the system under the two given situations. To achieve this, a weighting function W is used to introduce a bias on Boltzmann factors.[21,22] Moreover, by introducing $N-1$ intermediate states,[11] $Z_{\mathrm{sh+VW}}/Z_{\mathrm{sh}}$ can also be written as a product:

$$\prod_{j=0}^{N-1}(Z_{\mathrm{sh}+(j+1)\mathrm{VW}/N}/Z_{\mathrm{sh}+j\mathrm{VW}/N})$$

We hence get

$$\Delta F = -RT\sum_{j=0}^{N-1}\ln \langle \exp(-\beta E_{\mathrm{VW}}/N)\rangle_{\mathrm{sh}+(j/N)\mathrm{VW}}$$

The weighting function $W = \exp(-\beta E_{\mathrm{VW}}/2N)$ can then be used to improve the computation of ΔF and one has[11]

$$\Delta F = -RT\sum_{j=0}^{N-1} \ln \frac{\langle \exp(-\beta E_{\mathrm{VW}}/2N)\rangle_{\mathrm{sh}+(j+0.5)\mathrm{VW}/N}}{\langle \exp(+\beta E_{\mathrm{VW}}/2N)\rangle_{\mathrm{sh}+(j+0.5)\mathrm{VW}/N}}$$

Note that for a given value of j (from 1 to $N-1$), the average value of $\langle \exp(-\beta E_{\mathrm{VW}}/2N)\rangle$ is computed with Boltzmann factors calculated with the energy $E_{\mathrm{sh}} + (j + 0.5)E_{\mathrm{VW}}/N$.

III. Results and Discussion

III.1. Short Chain with Charged Ends. In order to test the validity of relations 5 and 6 in the case of a chain submitted to long-range effects, we applied them to a polymethylene chain of 10 bonds with the total energy $E_{\mathrm{T}} = E_{\mathrm{sh}} + E_{\mathrm{L}}$. Here $E_{\mathrm{L}} = E_{\mathrm{HS}} + E_{\mathrm{el}}$ and E_{HS} is a hard-sphere potential[19] equal to 0 when $r_{ij} > 2R_{\mathrm{VdW}}$ and $+\infty$ otherwise. r_{ij} is the distance between atoms i and j and R_{VdW} the van der Waals radius (presently equal to 1.6 Å).

The long-range effect is completed by an electrostatic potential (in kcal/mol): $E_{\mathrm{el}} = 332.0q_1q_2/\epsilon r$ where q_1 and q_2 are partial charges located on the ends of the chain. The dielectric constant ϵ equals to 3.5 and r is the end-to-end distance of the chain. It must be noted that such a chain model has no precise physical meaning; it is only used in order to probe the limit of validity of relations 5 and 6 when long-range interactions are used.

Note that this chain has only 3^8 different configurations, so an exact and complete calculation of the $P_{\mathrm{kl};i}$ is possible by using relation 6. The entropy can then be calculated by using relation 5, as it has been done for a chain with only short-range interactions. For this same small chain we also calculated the exact thermodynamic values directly from the relations giving Z_{T}, $F_{\mathrm{T}} = -RT\ln Z_{\mathrm{T}}$, E_{T}, and $TS_{\mathrm{T}} = E_{\mathrm{T}} - F_{\mathrm{T}}$. Results obtained from these two approaches can thus be compared.

The product of the two charges varies from $-1.0e^2$ to $+1.0e^2$, and for every value of q_1q_2, the entropy and the conformational energy E_{T} are calculated. The long-range contribution to the energy is then $E_{\mathrm{L}} = E_{\mathrm{T}} - E_{\mathrm{sh}}$ where E_{sh} is obtained from relation 2. Results obtained for a chain with short-range, hard-sphere, and electrostatic interactions are presented in Figure 1a and for chains with van der Waals interactions, in Figure 1b. It appears that both sets of curves are very similar.

The value of TS for the chain with only short-range and hard-sphere interactions (i.e. $q_1q_2 = 0$) is equal to 6.057 kcal/mol. When a small electrostatic potential is applied ($q_1q_2 = 0.1e^2$) the two methods of calculation give the same result for TS. When a very strong negative electrostatic potential is applied ($q_1q_2 = -1.0e^2$), the chain is completely folded, and since the electrostatic potential is so important, the chain is then maintained in a deep potential well of the configuration space. As a consequence, one gets a small value for the entropic term ($TS = 1.444$ kcal/mol). In contrast, when the electrostatic potential is very repulsive ($q_1q_2 = 1.0e^2$), its effect is less important because the end-to-end distance is generally very large in that case; the entropy of conformation is then only slightly smaller ($TS = 5.083$ kcal/mol) than that obtained with $q_1q_2 = 0$.

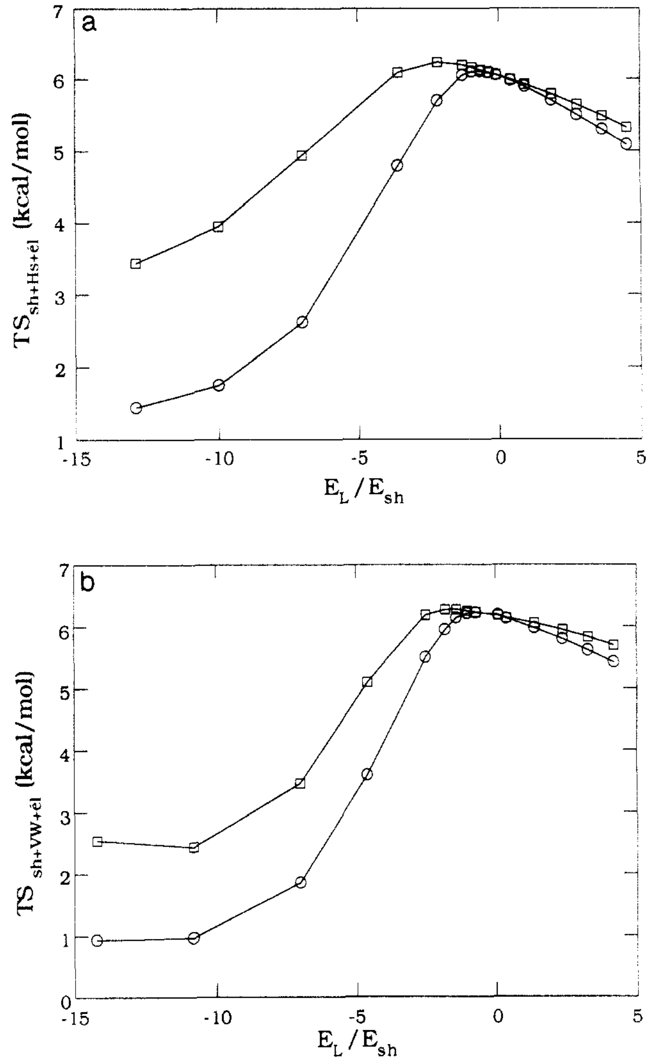

Figure 1. (a) Conformational entropy of polymethylene chains as a function of $E_{\mathrm{L}}/E_{\mathrm{sh}}$; $E_{\mathrm{L}} = E_{\mathrm{Hs}} + E_{\mathrm{el}}$ (E_{Hs} is the hard-sphere potential, and E_{el} the electrostatic energy due to varying charges located on the ends of the chains. Key: (○) complete calculation of the partition function; (□) relation 7 using matrices **Q** given by a complete calculation. (b) Conformational entropy of polymethylene chains with van der Waals and electrostatic interactions for E_{L}. Results are obtained as indicated in (a).

In fact we can define the probabilities $P_{\mathrm{kl};i}$ as mean long-range probabilities. So the long-range interactions can be considered as a perturbation of the short-range interactions. One can thus introduce a "perturbation factor", $P = E_{\mathrm{L}}/E_{\mathrm{sh}}$, and it appears that results obtained from the two methods of calculation present an increasing disagreement when the factor P decreases with values lower than above -1.5.

III.2. Chain with Short-Range and Hard-Sphere Interactions. Results obtained in this case show that relation 5 can be used even for the longer chains. In Figure 2 we can see that $\langle E_{\mathrm{sh+Hs}}\rangle$ is a linear function of n. Calculations performed with matrix **U**, for $N > 20$, gave $\langle E_{\mathrm{sh}}\rangle \approx 0.220(n-2)$ (kcal/mol) and Monte Carlo sampling permitted the following relation to be established: $\langle E_{\mathrm{sh+Hs}}\rangle = 0.196(n-2)$ (kcal/mol).

With the approximation $\Delta E = \langle E_{\mathrm{sh+Hs}}\rangle - \langle E_{\mathrm{sh}}\rangle$, we obtained $\Delta E = 0.024(n-2)$ (kcal/mol) which corresponds to $\Delta E/\langle E_{\mathrm{sh}}\rangle = 0.109$. The perturbation factor being very small, results given by eq 5 are actually in very good agreement with values obtained from Monte Carlo calculations. An example of such an agreement can be seen in Figure 3a where the entropic term is represented.

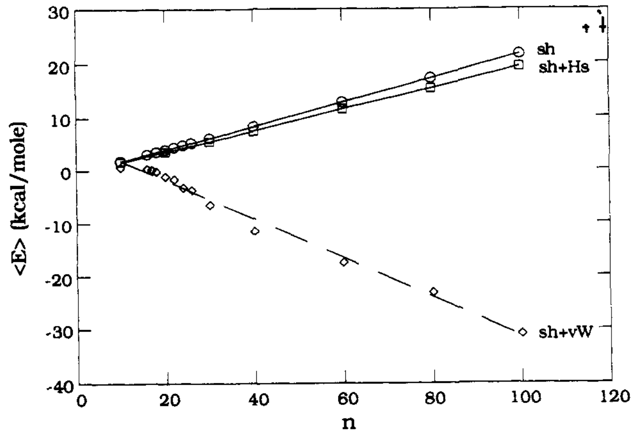

Figure 2. Conformational energy as function of the number n of bonds: (○) Calculated with relation 2 for chains with short-range interactions; (□) results from Monte Carlo calculations for chains with hard-sphere interactions; (◇) results from Monte Carlo calculations for chains with van der Waals interactions.

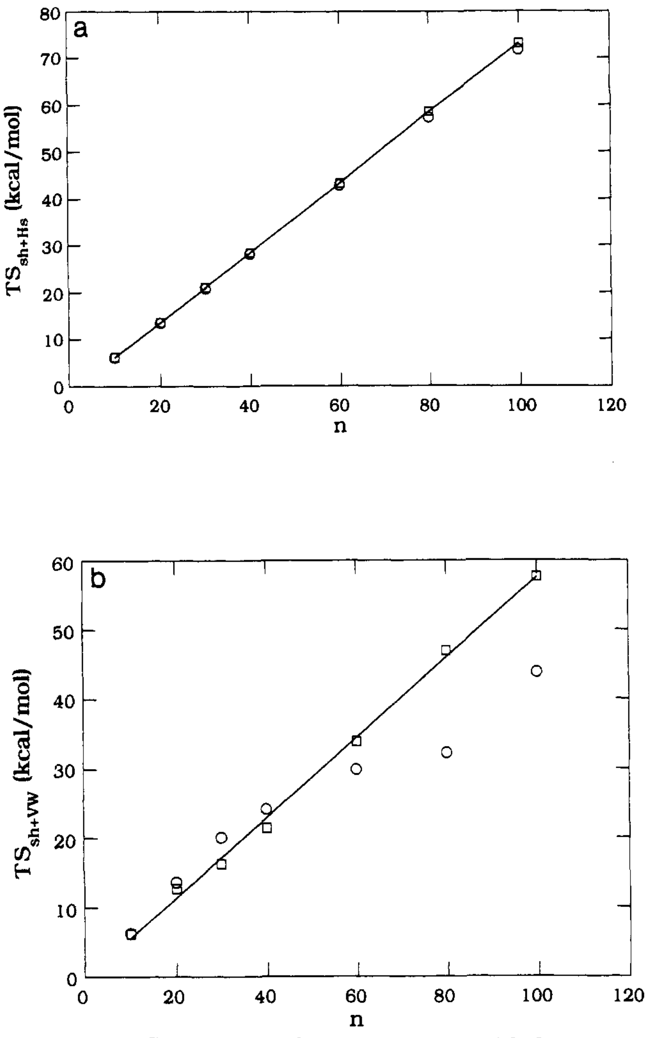

Figure 3. (a) Conformational entropy of chains with short-range and hard-sphere interactions as a function of chain length n: (□) obtained from the energies E (Monte Carlo) and F (Bennett's method); (○) obtained by relation 7 using matrices **Q** (obtained from Monte Carlo). (b) Conformational entropy of chains with short-range and van der Waals interactions as a function of the number of bonds n: (□) Monte Carlo calculation of E and umbrella sampling for F; (○) results from relation 7 using matrices **Q** (from Monte Carlo sampling).

III.3. Chain with Short-Range and van der Waals Interactions. Values of $\langle E_{sh}\rangle$ and $\langle E_{sh+VW}\rangle$ obtained respectively from the two methods of calculation are given in Figure 2. It appears that for $N > 20$, the total energy is negative; one has $\langle E_{sh+VW}\rangle = -0.37(n - 2) - 4.96$ (kcal/mol) for large n.

One should note that in a chain of n units submitted to a van der Waals potential there are $n(n - 1)/2$ terms of interaction. But, as we see, a linear function of n is obtained for the energy. This comes from the fact that with van der Waals potentials an almost constant small number of atoms does actually interact with any atom of the chain and the energy becomes therefore a linear function of n.

Moreover, one can see that the long-range interactions are more important than the short-range ones. So, the perturbation factor P is smaller than −1, and the larger is n, the smaller is P. Consequently, results given by relation 5 are valid for small n but, as we can see in Figure 3b, when $n > 60$, values of the entropy are in poor agreement with those given by an umbrella sampling. It appears that in the presence of long-range interactions, the mean probability approximation can be used to compute the conformational entropy only if $|E_l| \leq E_{sh}$.

III.4. Conformational Free Energy and Entropy. The free energy of the chain with short-range interactions obtained from calculation with matrices **U** is a linear function of n (Figure 4): $F_{sh} = -0.54(n - 2)$ (kcal/mol). Umbrella sampling and Bennett's method gave us the values of $\Delta F = -RT \ln (Z_{sh+L}/Z_{sh})$ for the two kinds of long-range interactions (van der Waals and hard spheres) and for chains of different bond number n. Results thus obtained can be written as

$$\Delta F_{VW} = F_{sh+VW} - F_{sh} = -0.36(n - 2) \text{ (kcal/mol)}$$

$$\Delta F_{HS} = F_{sh+HS} - F_{sh} = +0.0003(n - 2) \text{ (kcal/mol)}$$

Consequently, the free energy of the chains in total interaction depends linearly of n (Figure 4):

$$F_{sh+HS} = -0.54(n - 2) \text{ (kcal/mol)}$$

$$F_{sh+VW} = -0.90(n - 2) \text{ (kcal/mol)}$$

The conformational entropy of the chain with short-range interactions calculated with the matrix **U** can be put in the following form (Figure 5): $TS_{sh} = 0.765(n - 2)$. Values of the entropy obtained from the Monte Carlo sampling can also be written as linear functions of n (Figure 5):

$$TS_{sh+VW} = 0.587(n - 2) \quad \text{and} \quad TS_{sh+HS} = 0.734(n - 2)$$

III.5. Chain Dimensions. The conformational behavior of chains respectively submitted to the three kinds of potentials can also be characterized by values of their characteristic ratio,[10] $Cn = \langle r^2\rangle/nl^2$ with $\langle r^2\rangle$ being the mean square of the end-to-end distance of the chain with bond length l. Results, presently obtained from Monte Carlo calculations are shown in Figure 6. For the chain with short-range interactions, results are in very good agreement with those given by Flory and Jernigan.[14] With the hard-sphere potential, as atoms are not allowed to overlap, we thus observe that average end-to-end distances are larger than for chains with only short-range interactions. One can see that for the chains with short-range and van der Waals interactions, a maximum of Cn occurs for $n \approx 18$. Moreover, for $n < 14$, Cn values are near to those of chains with only short-range interactions. But this is no longer observed for $n > 20$. Such an effect can be explained by considering the values of the energy of the chain with short-range and van der Waals interactions. For $N < 18$, the total conformational energy is positive

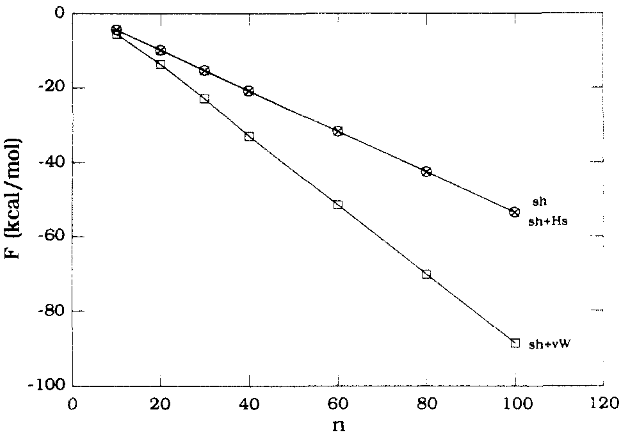

Figure 4. Conformational free energy as a function of chain length n: (○) short-range interaction and relation 1; (×) hard-sphere and Bennett's method; (□) van der Waals and umbrella sampling. Values for chains with short-range or hard-sphere interactions present differences of less than 1%.

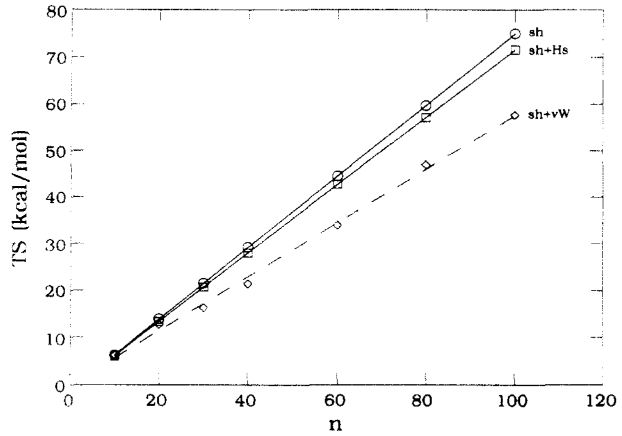

Figure 5. Conformational entropy of chains as a function of chain length n: (○) short-range interactions using relation 3; (□) results obtained from Monte Carlo calculations for hard-sphere interactions; (◇) results obtained from Monte Carlo calculations for van der Waals interactions.

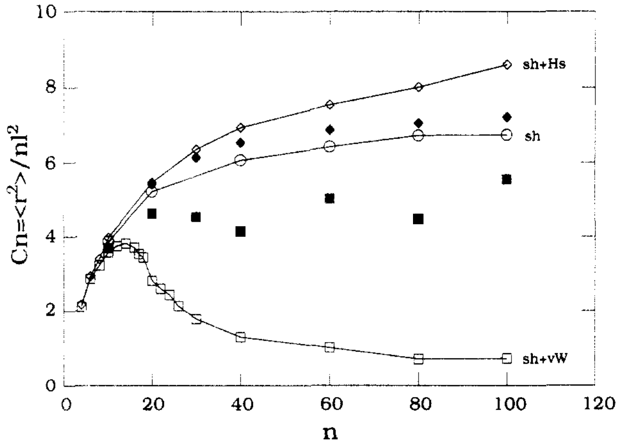

Figure 6. Flory ratio $Cn = \langle r^2 \rangle / nl^2$ for polymethylene chains as a function of the chain length n: (○) short-range; (◇) hard-sphere interactions with Monte Carlo calculations; (◆) hard-sphere interactions calculated with **Q** matrices; (□) van der Waals interactions with Monte Carlo calculations; (■) van der Waals interactions calculated with **Q** matrices.

and this means that $E_{sh} > |E_{VW}|$. The van der Waals contribution to the energy increases with n. Cn also increases with n up to $n = 18$, which is the value of n where E_{sh} equals $|E_{VW}|$ and the maximum for Cn occurs (Figure 6). When $n > 18$, the total conformational energy becomes negative ($E_{sh} < |E_{VW}|$); the attractive energy term being the most important, we observe a folding of the chain. For increasing values of n, this effect of the van der Waals energy is even more and more important and the decreasing values of Cn correspond to collapsed chains.

Matrices $\mathbf{Q}_i$ obtained from Monte Carlo computations have also been applied to the calculation of values of Cn for the three kinds of potential presently used. As indicated in Figure 6, results obtained from calculations performed with matrices $\mathbf{Q}$ disagree with those obtained from Metropolis Monte Carlo sampling, except, as said above, for the chain with short-range interactions.

IV. Conclusion

A limitation on the present method comes from the fact that there is no simple relation between the energy of a chain and the value of Cn. As a consequence, it appears that matrices $\mathbf{Q}_i$ obtained from Monte Carlo calculations are not adapted to compute geometrical parameters of chains with long-range interactions. To get better results, one should rather apply the procedure proposed by Mattice et al.[8,9] who used known asymptotic relations to perform adequate modifications of the matrix elements.

Conversely, for quantities directly related to the chain conformational energy such as conformational free energy and entropy, calculations performed using matrices $\mathbf{Q}_i$ (determined from Monte Carlo calculations when long-range interactions are included), give results in fairly good agreement with values obtained from umbrella sampling. The main limitation in the use of these matrices may come from relatively too important long-range interactions compared to short-range ones. But, results presently obtained show that, with potentials generally used and obviously for hard-sphere potentials, such a limitation should not be very drastic. One should essentially note that the application of the present method of calculation with matrices $\mathbf{Q}_i$ is very fast compared to the time-consuming Monte Carlo procedures. It could even be much more efficient when applied to macromolecular chains composed of units more complex than those of polymethylene.

References and Notes

(1) Go, N.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 535.
(2) Go, M.; Go, N.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 137.
(3) Karplus, M.; Kushick, J. N. *Macromolecules* **1981**, *14*, 325.
(4) Meirovitch, H. *Macromolecules* **1983**, *16*, 249.
(5) Meirovitch, H. *Phys. Rev.* **1985**, *A32*, 3709.
(6) Meirovitch, H.; Vasquez, M.; Scheraga, H. A. *Biopolymers* **1988**, *27*, 1180.
(7) Flory, P. J. In *Statistical Mechanics of Chain Molecules*; Interscience: New York, 1969.
(8) Mattice, W. L.; Santiago, G. *Macromolecules* **1980**, *13*, 1560.
(9) Mattice, W. L. *Macromolecules* **1981**, *14*, 1485.
(10) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245.
(11) Premilat, S. *J. Chem. Phys.* **1989**, *90*, 5128.
(12) Eyring, H. *Phys. Rev.* **1932**, *39*, 746.
(13) Abe, A.; Jernigan, R. L.; Flory, P. J. *J. Am. Chem. Soc.* **1966**, *88*, 631.
(14) Flory, J. P.; Jernigan, R. L. *J. Chem. Phys.* **1964**, *42*, 3509.
(15) Metropolis, N. A.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
(16) Binder, K. In *Monte Carlo Methods in Statistical Physics*; Springer-Verlag: Berlin, 1979.
(17) Premilat, S.; Maigret, B. *J. Chem. Phys.* **1977**, *66*, 3418.
(18) Premilat, S. *J. Chem. Phys.* **1988**, *88*, 3385.
(19) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95.
(20) Pletnev, V. Z.; Popov, E. M.; Kadymova, F. A. *Theor. Chim. Acta* **1974**, *35*, 93.
(21) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187; *Chem. Phys. Lett.* **1974**, *28*, 578.
(22) Lee, C. Y.; Scott, H. L. *J. Chem. Phys.* **1980**, *73*, 4591.
(23) Beveridge, D. L.; DiCapua, F. M. In *Computer simulation of biomolecular systems*; van Gunsteren, W. F., Weiner, P. K., Eds.; Escom: Leiden, 1989; pp 1–26.