content $\langle \theta_h \rangle$, and the last one is plotted against $\bar{\theta}_h$. The results for the mean-square radii of gyration $\langle S^2 \rangle_H$, $\langle S_0^2 \rangle^*$, and $\langle S_0^2 \rangle$ are also shown in a similar manner in Figure 5. It should be noted that $\langle R_0^2 \rangle^*$ and $\langle S_0^2 \rangle^*$ calculated for a finite number of samples by the Monte Carlo method were, in good agreement with $\langle R_0^2 \rangle$ and $\langle S_0^2 \rangle$, respectively, from the exact calculation for chain length $N = 151$. It was also found that the mean helical content $\langle \theta_h \rangle$ obtained by the Monte Carlo generation for $N = 151$ under the conditional probabilities for infinite chain length was in good agreement with F_h or $\bar{\theta}_h$ calculated exactly by the matrix multiplication method for $N = 151$.

Figure 6 illustrates the dependence of the mean-square end-to-end distances in the helix–coil transition region relative to that in the randomly coiling conformation, $\langle R^2 \rangle_H / \langle R^2 \rangle_{H;coil}$, $\langle R_0^2 \rangle^* / \langle R_0^2 \rangle^*_{coil}$, and $\langle R_0^2 \rangle / \langle R_0^2 \rangle_{coil}$, on the helical content $\langle \theta_h \rangle$ or $\bar{\theta}_h$. In Figure 7 relations for the mean-square radius of gyration are shown in a similar fashion. The dimensions of $\langle R^2 \rangle_{H;coil}$ and $\langle S^2 \rangle_{H;coil}$ are those obtained in our previous study[24] on non-self-intersecting chains for a randomly coiling polypeptide, and hence include the excluded-volume effect. The unperturbed dimensions in the random-coil form, $\langle R_0^2 \rangle^*_{coil}$, $\langle S_0^2 \rangle^*_{coil}$, $\langle R_0^2 \rangle_{coil}$, and $\langle S_0^2 \rangle_{coil}$, are also cited from the same work.[24] As shown by Nagai,[20] the unperturbed chain dimensions during the transition from coil to helix pass through a minimum. In

the model used here we cannot detect the minimum of the unperturbed dimensions in the figure, although they are present. Therefore, we could not compare the dimensions in the unperturbed state with those for non-self-intersecting chains in this respect.

The results on the mean-square end-to-end distance $\langle R^2 \rangle_{V,D}$, which was evaluated for $k = 100$, in eq 29 by taking into account the contribution of conformational energies calculated from eq 1 and 2, are plotted against the mean helical content $|\theta_h|$ in Figure 8. The unperturbed dimensions $\langle R_0^2 \rangle$ obtained from an exact calculation with eq 30 are also shown as the solid curve in the same figure. The results for the mean-square radius of gyration are shown in a similar manner in Figure 9. The excluded-volume effect during the helix–coil transition could be qualitatively observed in Figure 8, and more distinctly for the case of $\langle S^2 \rangle_{V,D}$ in Figure 9. Some points for $\langle S^2 \rangle_{V,D}$ are located below the solid curve, *i.e.*, the unperturbed dimension, $\langle S_0^2 \rangle$, in the region of helical content near unity; this results from the fact that the rodlike conformation abruptly bends at the middle point of the chain. The position of the junction and the orientation of bending significantly affect $\langle R^2 \rangle$ and $\langle S^2 \rangle$ in the region of high helical content. The results on $\langle R^2 \rangle_{V,D}$ and $\langle S^2 \rangle_{V,D}$, as seen in Figure 8 and 9, were therefore scattered. More refined results may be obtained by increasing the number of samples.

Computer Simulation of Polymer Conformation. II. Distribution Function for Polymers with Excluded Volume[1,2]

Steven D. Stellman*[3] and Paul J. Gans

New York University, Department of Chemistry, New York, New York 10003. Received July 1, 1972

ABSTRACT: Numerical distributions of end-to-end distances were generated by a Monte Carlo method for hard-sphere off-lattice polymers of length $N = 20, 40, 60, 80, 98$, and 298 atoms. Comparison by χ^2 tests against five recently proposed theoretical distribution functions showed that for $N = 80$ and $N = 98$, the data could be described, with 95% confidence, by the equation $f(r) = \exp[-(ar^2 + br + c)]$, where a and b are fitted parameters and c is a normalization constant. For $N = 298$, limitations of sample size lead to lower confidence limits (about 80%), but good fit. The above equation, and not its gaussian counterpart $\exp(-cr^2)$, is probably the limiting distribution function. The function accurately predicts the 1st through 12th observed moments at all chain lengths.

The distribution function for the end-to-end distance of a polymer with excluded volume has been the subject of many recent investigations.[4-15] Nevertheless, its theory remains very much an open question. The importance of

knowing the correct function arises from the relationship of its moments, $\langle r^2 \rangle$, $\langle r^4 \rangle$, etc., to observable physical quantities. Some geometric properties which could be predicted with the aid of the correct distribution function are the mean-squared radius of gyration,[16] light scattering curves,[17] average dipole moment,[18] and viscosity.[19]

(1) Presented in part at the IUPAC Symposium on Macromolecules, July 2–7, 1972, Helsinki.

(2) Paper I: S. D. Stellman and P. J. Gans, *Macromolecules*, 5, 516 (1972).

(3) To whom correspondence should be addressed at the Department of Biochemical Sciences, Frick Chemical Laboratory, Princeton University, Princeton, N. J. 08540.

(4) H. Yamakawa, *J. Chem. Phys.*, 48, 3845 (1968).

(5) P. J. Flory, "Principles of Polymer Chemistry," Cornell University Press, Ithaca, N. Y., 1953, Chapter 10.

(6) J. Mazur, *J. Chem. Phys.*, 43, 4354 (1965).

(7) J. Mazur, *J. Res. Nat. Bur. Stand., Sect. A*, 69, 355 (1965).

(8) C. Domb, J. Gillis, and G. Wilmers, *Proc. Phys. Soc.*, 85, 625 (1965).

(9) S. F. Edwards, *ibid.*, 85, 613 (1965).

(10) Z. Alexandrowicz, *J. Chem. Phys.*, 46, 3789 (1967).

(11) (a) H. Reiss, *ibid.*, 47, 186 (1967); (b) J. Curro and P. J. Blatz, *ibid.*, 48, 2832 (1968).

(12) R. Yeh and A. Isihara, *ibid.*, 51, 1215 (1969).

(13) T. Ree Chay, *ibid.*, 52, 1025 (1970).

(14) G. Allegra and G. Avitabile, *ibid.*, 56, 2385 (1972).

(15) F. T. Hioe and F. T. Wall, *J. Phys. Chem.*, 74, 4410 (1970).

(16) M. Fixman, *J. Chem. Phys.*, 36, 306 (1962).

(17) D. McIntyre and F. Gornick, Ed., "Light Scattering from Dilute Polymer Solutions," International Science Review Series, Vol. 3, Gordon and Breach, New York, N. Y., 1963.

(18) M. V. Volkenshtein, "Configurational Statistics of Polymeric Chains," Interscience, New York, N. Y., 1963, pp 331–352.

(19) P. J. Flory and T. G. Fox, *J. Amer. Chem. Soc.*, 73, 1904 (1951); M. V. Volkenshtein, ref 18, p 376.

In recent years, many hypothetical functions have been proposed on various theoretical or empirical grounds. Workers have converged on this problem mainly by three paths: analytical theory, exact enumeration, and Monte Carlo simulation. The goal of the theoretical approach is presently to obtain an asymptotic form of the distribution function, in terms of the expansion coefficient α and excluded-volume parameter z. Enumeration and Monte Carlo studies attack the problem from the opposite end, *viz.*, extrapolation of shorter chain results ($N = 1$ to *ca.* 5000, where N is the chain length) to chain lengths more representative of real polymers. In all cases, the goal is to obtain a closed distribution function of reasonable mathematical form on which to base predictions of polymer dimensions.

In paper I of this series,[2] details of a Monte Carlo method for studies of a hard-sphere off-lattice polymer model were given, and the dependence of the mean-squared end-to-end distance and radius of gyration were presented and discussed. In the present work the actual distributions of the end-to-end distance for various chain lengths are compared with several proposed distribution functions, for the purpose of choosing a "best" function consistent with our model and data. Comparison of trial to observed distributions is made by χ^2 tests and by comparison of predicted to observed moments. We further examine the possibility that one or more of these functions predict the limiting distribution for the excluded volume problem, the principal assumption being that if any of the proposed forms is a true asymptotic one, then χ^2 ought to decrease as the chain length N increases.

Two sets of comparisons were made. The first was done for the functions exactly as proposed by their authors, so that their applicability to our model could be directly tested. A second set of calculations was made on the same functions, but with variation of the function parameters. In this case we are testing the *form* of the equation, and search for those values of the parameters which minimize χ^2.

Trial Distribution Functions

The trial distribution functions selected for this study are listed along with their sources in Table I, and are numbered I–V. One or two examples of each of the three approaches, analytical, exact enumeration, and Monte Carlo, were selected.

The analytical method first proposed by Kuhn[20] uses a gaussian distribution function

$$f(r) = c_1 \exp(-c_2 r^2) \tag{1}$$

where r is the end-to-end distance and $4\pi r^2 f(r) dr$ is the probability that r falls in the range $r, r + dr$. The normalization constant c_1 is equal to $(3/2\pi N l^2)^{3/2}$, and $c_2 = 3/2Nl^2$ for a freely jointed chain. For chains with restricted bond angles but without excluded volume, the quantity Nl^2 is usually replaced by the distribution function second moment. However, for such models Flory[21,22] has shown that

$$\lim_{N \to \infty} \langle r^2 \rangle / N l^2$$

is finite, and that such chains can still be treated with gaussian statistics.

It has long been known that the nonbonded interactions in polymers destroy the gaussian character of the distribution function. The most successful analytical approaches have

TABLE I
DISTRIBUTION FUNCTIONS USED IN THIS STUDY

Eq	Function	Parameters[a]	Ref
I	$f(r) = \exp(-c_2 r^2)$	c_2	20
II	$f(r) = \exp(-\alpha_n{}^2 r^2/h^{*2} + 2(\alpha_n{}^2 - 1)r/h^*)$	α_n, h^*	10
III	$f(r) = \exp[-d_2(r - d_1)^2]$	d_1, d_2	9
IV	$f(r) = (r/\sigma)^{1/2} \exp[-(r/\sigma)^{5/2}]$	σ	8
V	$f(r) = \exp(-\alpha_N r^t)$	t, α	7

[a] Defined in the text.

treated this excluded-volume effect as a perturbation of the normal random walk. Alexandrowicz[10] has obtained his equation (eq II in Table I) by treating the distribution of internal segmental distances as gaussian, and expanding and averaging the intramolecular energy factor, $\exp[-\beta \Sigma_{k<l} \cdot \Sigma \delta(h_{lk})]$, where h_{lk} is the separation of segments l and k, and β is the binary cluster integral. The equation obtained by Edwards[9] (eq III of Table I) was derived by treating the interaction of the polymer with itself as a self-consistent field, obtaining an equation for $p(r, L)$, the probability of a polymer of length L having end-to-end distance r, also as a function of the intramolecular energy expression.

For our numerical study we can consider Alexandrowicz's and Edwards' distribution functions to have the single form

$$f(r) = k \exp[-(ar^2 + br + c)] \tag{2}$$

differing only in the assignment of physical meaning to the constants a, b, and c as follows: *gaussian*, $a = c_2$, $b = 0$, $c = 0$; *Alexandrowicz*, $a = c_2$, $b = -2(\alpha_n{}^2 - 1)/h^*$, $c = 0$; and *Edwards*, $a = 0.9c_2$, $b = -2(0.9c_2)[(5Nl/3)^{3/5}(\beta/3\pi l)^{1/5}]$, $c = (0.9c_2)[(5Nl/3)^{6/5}(\beta/3\pi l)^{2/5}]$. Here c_2 is the same as in eq 1, h^* is the most probable end-to-end distance, *i.e.*, that finite distance for which $\partial f(r)/\partial r = 0$, and $\alpha_n{}^2$ is the squared expansion factor, $\alpha_n{}^2 = h^{*2}/(2/3Nl^2)$, the denominator being the most probable squared end-to-end distance for gaussian statistics. Also, β, the binary cluster integral, reduces to $4\pi l^3/3$ for the hard-sphere model adopted for our Monte Carlo studies.

In addition to the analytical approaches taken by Alexandrowicz and Edwards, a variation-principle treatment was performed by Reiss[11a] and carried out to zero-order approximation in a self-consistent field. The resulting density function was found to have the form

$$f(r) = k \exp[-(ar^2 + b/r)] \tag{3}$$

However, Reiss's calculations using this lowest order approximation predicted a second-moment dependence $\langle r^2 \rangle \propto N^{4/3}$, which is not consistent with the widely accepted value $\langle r^2 \rangle \propto N^{6/5}$. Furthermore, other inconsistencies later pointed out[11b] leave this form somewhat in doubt. Therefore, we have somewhat arbitrarily chosen to omit eq 3 in this study.

In contrast to analytical studies, which attempt to derive the distribution function from first principles, exact enumeration[23] and Monte Carlo[2,24] studies simulate mathematically the geometric properties of the polymer and deduce afterwards the function which must have given rise to the observed distribution. Enumeration methods are limited to assessment of random walks on a space-filling lattice, where the ex-

(20) W. Kuhn, *Kolloid-Z. Z. Polym.*, **68**, 2 (1934).
(21) P. J. Flory, "Statistical Mechanics of Chain Molecules," Interscience, New York, N. Y., 1969.
(22) P. J. Flory and R. L. Jernigan, *J. Chem. Phys.*, **42**, 3509 (1965).

(23) C. Domb, *Advan. Chem. Phys.*, **15**, 229 (1969), and references cited therein.
(24) P. J. Gans, *J. Chem. Phys.*, **42**, 4159 (1965), and references cited therein.

<div align="center">

TABLE II

OBSERVED MONTE CARLO DISTRIBUTIONS OF THE SQUARED END-TO-END DISTANCE, r^2, SHOWING THE UPPER LIMIT
OF EACH GROUP, r_u^2, AND THE FRACTION OF POLYMERS OBSERVED IN EACH GROUP, f_g

</div>

N		1	2	3	4	5	6	7	8	9	10	11	12
20	r_u^2	19.02	38.04	57.05	76.07	95.09	114.1	133.1	152.1	190.2	247.2	342.3	∞
	f_g	0.0393	0.0811	0.1105	0.1011	0.1096	0.1078	0.0962	0.0834	0.1300	0.1017	0.0373	0.0020
40	r_u^2	48.0	96.1	144.1	192.1	240.2	288.2	336.2	384.3	480.3	624.4	864.6	∞
	f_g	0.0411	0.0923	0.1111	0.1091	0.1142	0.0997	0.0838	0.0825	0.1177	0.0938	0.0475	0.0072
60	r_u^2	78.9	157.8	236.6	315.5	394.4	473.3	552.1	631.0	788.8	1025	1420	∞
	f_g	0.0464	0.0897	0.1132	0.1095	0.1088	0.0984	0.0955	0.0739	0.1146	0.0978	0.0432	0.0090
80	r_u^2	113.2	226.4	339.6	452.9	566.1	679.3	792.5	905.7	1132	1472	2038	∞
	f_g	0.0461	0.0837	0.1110	0.1180	0.1178	0.1043	0.0864	0.0743	0.1099	0.0894	0.0497	0.0094
98	r_u^2	144.2	288.5	432.7	576.9	721.2	865.4	1010	1154	1442	1875	2596	∞
	f_g	0.0476	0.0905	0.1159	0.1174	0.1082	0.1012	0.0925	0.0669	0.1074	0.0868	0.0529	0.0127
298	r_u^2	764	1145	1527	1909	2482	2863	3436	4009	4581	5536	6681	∞
	f_g	0.0842	0.0726	0.0767	0.0815	0.1112	0.0727	0.0966	0.0828	0.0709	0.0902	0.0740	0.0866

cluded volume is fulfilled by the constraint of self-avoidance of the walks. They are also limited to extremely short lengths (for 13 steps, nearly a billion self-avoiding walks are possible on the simple cubic lattice). The properties of longer walks must then be extrapolated from enumeration data.

To represent this type of study, we have chosen the functions proposed in the investigations of Domb, Gillis, and Wilmers,[8] who performed exact enumerations of self-avoiding walks of up to 13 steps on the simple cubic lattice (eq IV of Table I)

$$f(r) = k(r/\sigma)^{1/2} \exp[-(r/\sigma)^{5/2}] \qquad (4)$$

where σ is the standard deviation of the distribution. A function having the same form was also adopted by Hioe and Wall[15] to describe their enumeration results.

Finally, Monte Carlo studies of self-avoiding walks are similar to exact enumeration in that they produce an observed set of walks and attempt to fit proposed functions to the data. The first such Monte Carlo distributions reported by Wall, *et al.*,[25] could not be fit to any functions because of the severe scatter of the data. The scatter itself was due to the inefficient sampling method; the data obtained in the present study are far more adequate, owing to the improved sampling methods described in paper I.

In Monte Carlo studies of self-avoiding walks on a tetrahedral lattice, Mazur[7] proposed the function (eq V of Table I)

$$f(r) = k \exp(-\alpha_N r^t) \qquad (5)$$

where α_N is a function only of chain length (and is *not* the expansion factor), and the exponent t is a "shape factor." This factor adds an additional degree of freedom to the function of Domb, Gillis, and Wilmers, to which eq 5 reduces for the special case $t = 2.5$, $\alpha_N = \sigma^{-2.5}$. Extrapolating his Monte Carlo data to $N = \infty$, Mazur found the best value of t to be 3.2, and required the second moment of $f(r)$ to coincide with the observed value of $\langle r^2 \rangle$, so that

$$\alpha_N = [\Gamma(5/t)/\Gamma(3/t)\langle r^2 \rangle]^{t/2} \qquad (6)$$

By use of the incomplete Γ function,[26] one can show that

$$\Phi(r) = \Gamma(3/t, \alpha_N r^t)/\Gamma(3/t) \qquad (7)$$

where $\Phi(r)$ is the fraction of walks having end-to-end distance of r or more.

Finally, Schatzki[27] has reported Monte Carlo distribution function data for lengths of up to 100 steps on a tetrahedral lattice. These data were fit by least squares to a four-term orthogonal function expansion, four being the number of terms needed to reduce the sum of the squared residuals below about 0.5%. The numerical results are compared with our own in a later section.

Method of Calculation

Numerical distributions of end-to-end distances were obtained by the Monte Carlo method described in paper I of this series.[2] The polymer model used was a chain of N ($= 20, 40, 60, 80, 98, 298$) atoms with fixed bond length 1.54 Å, fixed tetrahedral bond angles, and continuously variable dihedral angles (off-lattice model). The excluded-volume effect was imposed by a hard sphere of diameter 1.54 Å centered on each backbone atom. The Monte Carlo experiment consisted of varying the dihedral angles at random and averaging the properties of those generated conformations which were free of hard-sphere overlaps. Importance sampling was employed to reduce sampling variance and increase efficiency, and sample weights were biased to reflect a canonical ensemble.

For each length N, data consisting of individual values of r^2, the end-to-end distance of a single polymer, were collected at each step of the simulation, arranged in order of increasing magnitude, and then divided into 12 groups of approximately equal numbers of data points. The Monte Carlo results are displayed in Table II.

Each function in turn was subjected to the following statistical analysis at each chain length. The null hypothesis, H_0, that the given function gave rise to the observed distribution of data, was tested by computing the statistic[28]

$$\chi_0{}^2 = \sum_{j=1}^{12} \frac{(d_j - e_j)^2}{e_j} \qquad (8)$$

and comparing $\chi_0{}^2$ to the value of χ^2 given by standard tables[29] at the 95% confidence level. In eq 8, d_j is the number of observed samples falling into the *j*th group (listed in

(25) F. T. Wall, L. A. Hiller, Jr., and W. F. Atchison, *ibid.*, **23**, 913, 2314 (1955); **26**, 1742 (1957).

(26) The incomplete Γ function is defined by the relation

$$\Gamma(y, x) = \int_x^\infty e^{-t} t^{y-1} dt$$

(27) T. F. Schatzki, *J. Polym. Sci.*, **57**, 337 (1962).

(28) The zero subscript signifies that $\chi_0{}^2$ is an estimator for the true χ^2 which would be observed with an infinite number of samples.

(29) E. Kreyszig, "Advanced Engineering Mathematics," Wiley, New York, N. Y., Chapter 18.

TABLE III
VALUES OF χ_0^2 FOR FUNCTIONS USING THEORETICAL PARAMETERS DEFINED IN TEXT

	I		II			III			IV		V		
N	$c_2{}^a$	χ_0^2	$h^{*\,b}$	α_n	χ_0^2	d_1	d_2	χ_0^2	$\sigma(r)^c$	χ_0^2	t	α_N	χ_0^2
20	1.581×10^{-2}	1,892	9.497	1.648	272.2	10.735	2.846×10^{-2}	11,500	8.197	3156	3.2	3.979×10^{-4}	223.1
40	7.906×10^{-3}	4,327	15.299	1.900	369.2	16.272	1.423×10^{-2}	8,980	13.487	1716	3.2	9.037×10^{-5}	57.58
60	5.271×10^{-3}	6,640	19.689	2.005	694.7	20.753	9.487×10^{-3}	9,580	17.524	1268	3.2	4.086×10^{-5}	37.52
80	3.953×10^{-3}	9,624	23.660	2.091	1002	24.663	7.115×10^{-3}	8,880	21.133	1083	3.2	2.292×10^{-5}	42.64
98	3.227×10^{-3}	12,600	26.747	2.138	1729	27.857	5.809×10^{-3}	10,000	24.137	904	3.2	1.556×10^{-5}	50.19
298	1.061×10^{-3}	97,000	52.949	2.435	7132				47.972	1326	3.2	1.794×10^{-6}	38.01

^{a} $c_1 = (c_2/\pi)^{3/2}$. ^{b} Most probable end-to-end distance estimated with eq 11. ^{c} $\sigma(r)$ estimated as $\sigma^{1/2}(r^2)$.

Table II), and e_j is the number of sample values expected in the jth group if the null hypothesis is true, i.e.

$$e_j = \frac{\displaystyle\int_{m_{j-1}}^{m_j} 4\pi r^2 f(r)\,dr}{\displaystyle\int_0^\infty 4\pi r^2 f(r)\,dr}, j = 1, \ldots, 12 \qquad (9)$$

where m_j is the upper limit of group j ($m_0 \equiv 0$) and $f(r) = H_0$.

For the 11 degrees of freedom here,[29] $\chi_{0.05,11}{}^2 = 19.68$. This means that if the computed χ_0^2 value exceeds 19.68, the null hypothesis, H_0, is to be rejected with 95% confidence. The tabulated values of the χ^2 statistic measure the statistical scatter which may reasonably be expected if the data are normally distributed about the assumed distribution function, H_0. Since we are explicitly concerned with this scatter, in the discussion which follows we will limit ourselves to noting and comparing trends within a function and to a lesser degree the relative magnitudes of χ_0^2. We note that when comparing values for a given function at different chain lengths, if χ_0^2 values are very close together, prediction of a trend may not be justified.

Results of χ^2 Tests

A. Direct χ^2 Tests. The values of χ_0^2 computed by the method described above are reported in Table III. The degree of fit may be observed in Figure 1, where we have plotted the observed distribution as a bar-type histogram, superimposed on plots of the five functions, for the point $N = 98$.

On the basis of the χ^2 criterion, Table III shows that none of the functions gave a statistically acceptable fit to the data. However, this single statistic does not tell the whole story, as examination of the individual functions shows.

Function I. Because all bond angles in our model were tetrahedral, the following equation was used in place of eq 1

$$f_I(r) = c_1{}' \exp(-3r^2/4Nl^2) \qquad (10)$$

From Table III it is seen that the attempted fit to the gaussian function begins poorly and becomes worse as N increases. Hence, the freely jointed chain model is utterly without value for describing real chain statistics.

Function II. The χ_0^2 for Alexandrowicz's function begins at about 272 and increases steadily with N. This behavior may be traced to the fact that function II depends on two parameters, one of which is fitted to the experimental mean, h^*, and the other of which is the analogous value for the gaussian case. But from the curve-fitting point of view, this value is purely arbitrary.
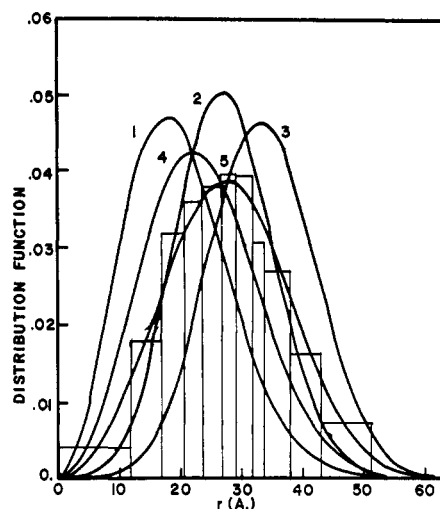

Figure 1. Normalized radial distribution functions, $4\pi r^2 f(r)$, vs. end-to-end distance r for the five functions, at data point $N = 98$. Observed distribution is shown as a histogram.

Since the χ_0^2 values themselves were much lower than for function I, we examined the theoretical constants more closely to see whether refinement of the derivation could lead to better fit. Three major assumptions were made in that derivation. The first was in its being carried to first order only. Secondly, the approximation $r \ll Nl^2$ was used, a condition which may mar the fit near the upper tail of the distribution. Finally, the function as reported by Alexandrowicz calls for the parameter h^*, the most probable end-to-end distance, rather than the ordinarily observed rms value, $\langle r^2 \rangle^{1/2}$. As we did not measure h^* directly, we were obliged to estimate its value before computing the degree of fit. This was done by using Alexandrowicz's eq 17–19, in which h^{*2} and $\langle r^2 \rangle$ respectively are expanded to two terms in z, the well-known excluded-volume parameter. Eliminating z between the two equations, one obtains

$$h^{*2} = \langle r^2 \rangle \left\{ \frac{1}{2}\pi^{1/2} + Nl^2 \left(\frac{2}{3} - \frac{1}{2}\pi^{1/2} \right) \right\} \Big/ \langle r^2 \rangle \qquad (11)$$

Substitution of the observed values of $\langle r^2 \rangle$ yields the h^* entries of Table III. It can be seen from eq 11 that the asymptotic dependence of the predicted second moment on N is undisturbed by this approximation.

Function III. The χ_0^2 values obtained with Edwards' parameters cluster about 10^4 and exhibit little variation with chain length. The relevant constants contain a variety of factors, many of which result from approximations and esti-

TABLE IV
VALUES OF χ_0^2 FOR FUNCTIONS USING BEST-FIT VALUES OF PARAMETERS

	I		II, III[a,b]					IV		V		
N	c_2	χ_0^2	h^*	α_n	a_{min}	b_{min}/a_{min}	χ_0^2	σ	χ_0^2	t	α_N	χ_0^2
20	1.265×10^{-2}	1160	9.827	1.748	3.164×10^{-2}	13.2206	173.91	9.608	302.2	3.73	9.071×10^{-5}	81.67
40	5.060×10^{-3}	721	15.448	1.622	1.102×10^{-2}	19.1561	53.76	15.210	78.39	3.51	3.267×10^{-5}	35.56
60	3.162×10^{-3}	667	19.650	1.590	6.547×10^{-3}	23.7513	55.50	19.493	64.44	3.41	1.954×10^{-5}	25.28
80	2.214×10^{-3}	634	23.526	1.590	4.568×10^{-3}	28.4370	18.48	23.250	37.45	3.31	1.538×10^{-5}	38.32
98	1.678×10^{-3}	475	26.341	1.510	3.286×10^{-3}	29.5800	13.62	26.290	14.83	3.04	2.890×10^{-5}	42.64
298	4.488×10^{-4}	255	51.281	1.432	7.798×10^{-4}	53.2749	32.81	51.980	31.82	3.015	4.180×10^{-6}	22.73

[a] Minimizing parameters for function III are given by: $d_1 = h^*(1 - \alpha_n^{-2})$, $d_2 = \alpha_n^2/h^{*2}$. [b] a_{min} and b_{min} are minimizing parameters for eq 2: $a = d_2$, $b = 2d_1d_2$.

mations carried out in the derivation. Function III does not seem to be the limiting distribution.

Function IV. The function proposed by Domb, *et al.*, shows a fit at $N = 20$ only moderately better than that of function III, but steadily improves to $N = 98$. Furthermore, the slight regression at $N = 298$ may be due only to insufficient samples, a point we shall take up again in connection with function II in the next section.

One major approximation was required with function IV. Equation 4 calls for σ, the standard deviation of the observed distribution. In our Monte Carlo studies, we observed not $\sigma(r)$, but $\sigma(r^2)$, so we have estimated $\sigma(r)$ as $[\sigma(r^2)]^{1/2}$. This could not have been too serious, for the χ_0^2 values listed in Table IV for the optimum σ's show the same trend as those in Table III, and differ from the latter by generally no more than 10%.

Function V. The best raw fit was obtained with the function and parameters of Mazur, as Figure 1 clearly illustrates. At $N = 20$, χ_0^2 begins at about the same level as for function II, but drops to a steady, low value near 50. This is encouraging, especially in view of the fact that one degree of freedom is lost due to forced fit to the observed second moment (eq 6). It is not possible at these lengths to test whether the fit will improve as N increases without bound.

B. Minimization of χ_0^2. Each of the five trial functions contains one or two parametric constants which are in principle functions of the physical attributes of the molecular geometry. However, it is possible to view them merely as adjustable parameters, and correspondingly to adjust their values so as to minimize χ_0^2, and hence to obtain the best fit possible with each functional form. We assume that this minimum value represents the best fit that can be hoped for at a given chain length, and, consequently, if such minimum values decrease as N increases, we can speculate on the asymptotic fit for the given function.

The minimization results are shown in Table IV and for the data points $N = 98$ in Figure 2. The observed distribution for $N = 98$ is also shown in Figure 2, as circles centered at the midpoints of the histogram bars which were shown in Figure 1. (The bars themselves were omitted for pictorial clarity.)

For the one-parameter functions (I and IV), the minimization was performed simply by recomputing χ_0^2 for different values of that one parameter. For two-parameter functions, it was necessary to compute a grid for $\chi_0^2(p_1, p_2)$, mapping the contours of the χ_0^2 space formed by.varying the two parameters p_1 and p_2, and gradually tightening the grid around the deepest observed minima. Standard numerical minimization procedures failed because the space is so "bumpy" and studded with dozens of local minima of nearly equal magnitude. We can assert that the values of minimizing parameters reported in Table IV are accurate to 2% or less. A
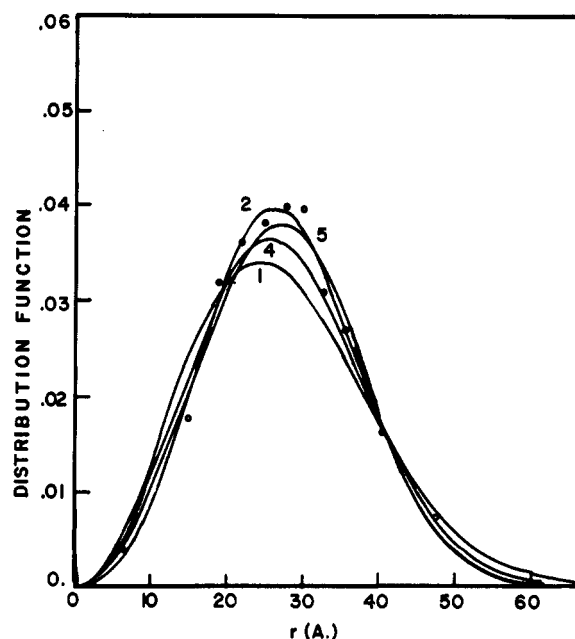

Figure 2. Normalized radial distribution function for the same functions with best-fit parameters (II = III). Observed distribution is shown as circles.

contour map for the minimization of function II for $N = 20$ is shown in Figure 3.

In addition to the described minimization, the 1st through 12th moments of the functions were computed using both theoretical and best-fit parameters, according to the following equation for the pth moment of $f(r)$

$$\langle r^p \rangle = \frac{\displaystyle\int_0^\infty 4\pi r^{p+2} f(r)\, dr}{\displaystyle\int_0^\infty 4\pi r^2 f(r)\, dr} \tag{12}$$

Once again, the results are best explained by considering the functions individually.

Function I. The gaussian distribution has only one parameter to be fit, c_2, since c_1 is merely a normalization constant. The effect of adjusting c_2 is to broaden the frequency–function curve by fitting the function mean to the observed $\langle r^2 \rangle$. Thus, from Table IV we find that

$$1/c_{2,min} = 2.12 N^{1.226} \propto \langle r^2 \rangle \tag{13}$$

while for the normal random walk $1/c_2$ is also proportional to $\langle r^2 \rangle$.

While it appears that the degree of fit improves with increasing chain length N, this improvement is very slow, and
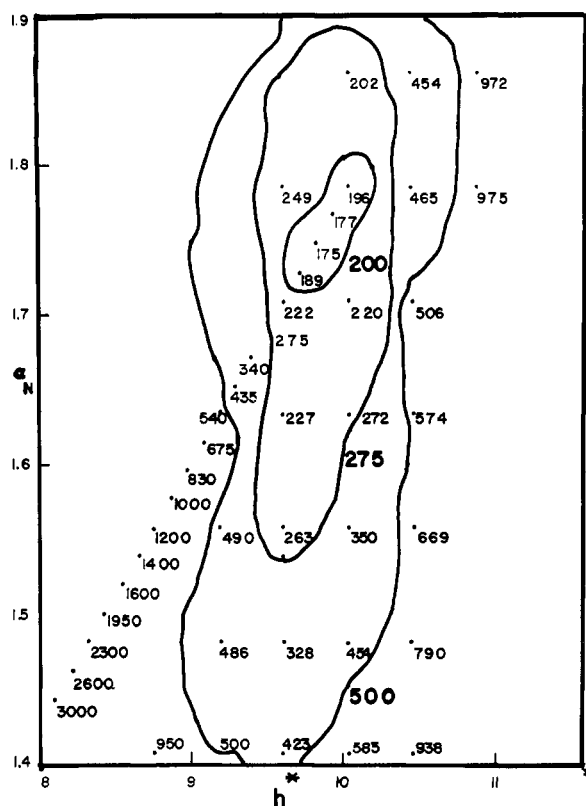
Figure 3. Contour grid for $\chi_0{}^2$ for function II, as a function of parameters α_n and h^*.

we cannot predict whether $\chi_0{}^2$ will ever drop below $\chi_{0.05,11}{}^2$ in the limit of infinite N. This point is important, because it is precisely this condition which is implied by any theoretical treatment which purports to reduce to the gaussian case in the limit of large N. The issue will be settled, however, in the discussion of higher moments in a later section.

Functions II and III. Since these two functions are equivalent, we list a single set of $\chi_0{}^2$ values in Table IV. The best-fit parameters are presented in two ways, first as values of $h^*_{\min}$ and $\alpha_{n,\min}$ for eq II of Table I, and also for comparison with eq 2.

Table IV shows not only that goodness of fit improves with N, but that by the time $N = 80$, $\chi_0{}^2$ has dropped below the critical value of 19.68, and thus justifies acceptance of eq 2 as the parent distribution of the observed data. Furthermore, the slightly higher value of $\chi_0{}^2$ at $N = 298$ does not necessarily preclude eq 2 as the limiting distribution. Judging from the large variance of the actual distribution (the standard deviation at $N = 298$ was 19.7, or 35% of the observed mean, indicating a quite broad peak), it may well be that insufficient samples were taken at $N = 298$. In support of this conjecture, it was observed that $\chi_0{}^2$ decreased from 52.51 for 6740 samples to the tabulated value of 32.81 when an additional 14,300 samples were obtained. This may also be an explanation for correspondingly higher values of $\chi_0{}^2$ (raw and best fit) for function IV at $N = 298$.

The values of $h^*_{N,\min}{}^2$ which minimize $\chi_0{}^2$ at each N are not equal to the $\langle r^2 \rangle_N$ but, as Alexandrowicz predicted, are directly proportional to them. From the data it was observed that

$$\langle r^2 \rangle_N = aN^b \qquad (14)$$

and that

$$h^*_{N,\min}{}^2 = a'N^{b'}$$

N	t	α_N	$\chi_0{}^2(t, \alpha_N)$
20	4.6009	7.957×10^{-6}	31.88
40	3.8154	1.201×10^{-5}	54.55
60	3.5289	1.276×10^{-5}	29.67
80	3.4981	7.561×10^{-6}	46.49
98	3.2396	1.336×10^{-5}	54.43
298	2.9895	4.662×10^{-6}	23.66

It was found by least-squares analysis that $b' = 1.195$, with $\sigma(b') = 0.021$. This lies close enough to our observed values $b = 1.221$, $\sigma(b) = 0.021$, to justify our concluding that

$$h^*_{N,\min}{}^2 = \frac{a}{a'}\langle r^2 \rangle$$

for this region of N. However, the ratio $\langle r^2 \rangle / h^{*2}$ actually increases very slowly with N, so it is not safe to assume this relationship beyond the given range.

Function IV. The fit obtained with the function proposed by Domb, *et al.*, improves as chain length increases, and becomes statistically acceptable at $N = 98$. For $N = 298$ the fit worsens, however. Therefore, on this information alone it is not possible to judge whether function IV is an acceptable limiting distribution, since the high value of $\chi_0{}^2$ at $N = 298$ could be attributed to insufficient samples. It will be necessary to calculate higher moments of the distribution for comparison with observed values. This is done in the following section.

Function V. Mazur's function tests an interesting hypothesis, because it predicts a distribution whose shape is independent of N. That is, except for a scaling factor which depends only on $\langle r^2 \rangle_N$, the distribution of end-to-end distances is an unvarying function for all N. We need only show that if the group limits are preset as fractions of $\langle r^2 \rangle_N$, then function II predicts group occupation numbers which are independent of N. From the definition of α_N given by eq 6, for any two chain lengths N_1 and N_2

$$\alpha_{N_1}/\alpha_{N_2} = (\langle r^2 \rangle_{N_2}/\langle r^2 \rangle_{N_1})^{t/2} \qquad (15)$$

$$\Phi_i(r)_{N_1} = \Gamma\left(\frac{3}{t}, \alpha_{N_1}x_{i,N_1}{}^{t/2}\right)\bigg/\Gamma\left(\frac{3}{t}\right)$$

where x_{i,N_1} is the upper limit for group i at chain length N_1. If we set

$$x_{i,N_1} = k\langle r^2 \rangle_{N_1}$$

and

$$x_{i,N_2} = k\langle r^2 \rangle_{N_2}$$

with k a constant, then

$$\alpha_{N_1}(k\langle r^2 \rangle_{N_1})^{t/2} = \alpha_{N_2}(k\langle r^2 \rangle_{N_2})^{t/2}$$

by eq 15. Since this quantity is the lower limit of integration in the evaluation of the incomplete Γ function, the assertion is proved.

As may be observed from Table III, the value of $\chi_0{}^2$ drops with increasing N to a minimum of 37.52 at $N = 60$, increases slightly, then falls to 38.01 at $N = 298$. Although these values never fall below 19.68, they compare favorably with those given by function II, and hence neither justify nor negate the hypothesis that the functional shape depends on N. To be conclusive, further data at $N = 298$ and higher are needed.

<div align="center">

TABLE VI

OBSERVED FIRST AND SECOND MOMENTS, ALONG WITH THOSE COMPUTED FOR GIVEN FUNCTIONS
WITH BOTH THEORETICAL AND BEST-FIT VALUES OF PARAMETERS

</div>

N	Moment	Obsd	I	II	III	IV	V^a
					—Function—		
With Theoretical Parameters							
20	1	10.170	8.97	9.898	13.577	8.604	10.064
	2	114.11	94.86	109.21	198.46	83.437	
40	1	16.084	12.69	15.725	20.086	14.158	15.994
	2	288.19	189.73	271.19	432.25	225.88	
60	1	20.565	15.54	20.152	25.280	18.395	20.497
	2	473.26	284.59	442.72	682.74	381.34	
80	1	24.627	17.95	24.146	29.772	22.184	24.555
	2	679.28	379.46	632.68	945.10	554.59	
98	1	27.723	19.86	27.257	33.420	25.337	27.715
	2	865.39	464.83	804.32	1189.21	723.47	
298	1	54.303	34.64	53.595		50.357	54.437
	2	3339.21	1413.47	3068.45		2857.8	
With Best-Fit Parameters							
20	1		10.033	10.178	b	10.086	10.073
	2		118.58	114.70		114.63	113.15
40	1		15.863	16.130		15.966	16.032
	2		296.44	290.53		287.28	287.75
60	1		20.067	20.566		20.462	20.504
	2		474.38	473.36		471.85	471.55
80	1		23.981	24.622		24.406	24.502
	2		677.51	678.53		671.27	674.77
98	1		27.546	27.747		27.597	27.649
	2		893.92	866.79		858.29	864.59
298	1		53.263	54.408		54.565	54.200
	2		3342.25	3353.10		3355.2	3324.37

^{a} Fit to $\langle r^2 \rangle$. ^{b} Same as function II.

As an alternative approach, we may inquire whether an improvement is obtained by fitting f(r) to two observed moments. Using the experimental first and second moments, it was possible to derive new values for the parameters t and α in function V, using the method of moments.[30] For function V, the pth moment is given by

$$\mu_p \equiv \langle r^p \rangle = \Gamma\!\left(\frac{3+p}{t}\right)\bigg/\alpha^{p/t}\Gamma\!\left(\frac{3}{t}\right) \qquad (16)$$

Therefore

$$\mu_1 = \Gamma\!\left(\frac{4}{t}\right)\bigg/\alpha^{1/t}\Gamma\!\left(\frac{3}{t}\right)$$

$$\mu_2 = \Gamma\!\left(\frac{5}{t}\right)\bigg/\alpha^{2/t}\Gamma\!\left(\frac{3}{t}\right)$$

Our Monte Carlo experiments provided average values of the first and second moments (see Table VI), m_1 and m_2, which are unbiased estimators for μ_1 and μ_2, respectively. Substituting m_1 for μ_1 and m_2 for μ_2, and solving for α, we have

$$\alpha = \left[\frac{\Gamma\!\left(\dfrac{4}{t}\right)}{m_1\Gamma\!\left(\dfrac{3}{t}\right)}\right]^{t} = \left[\frac{\Gamma\!\left(\dfrac{5}{t}\right)}{m_2\Gamma\!\left(\dfrac{3}{t}\right)}\right]^{t/2} \qquad (17)$$

To find t for the given distribution, then, one need only subtract the $t/2$ term in eq 17 from the t term, set the difference equal to zero, and solve the resulting equation for t. α is then found by substituting t back into eq 17. This procedure was done numerically by the method of secants,[31] with results

(30) J. E. Freund, "Mathematical Statistics," Prentice-Hall, Englewood Cliffs, N. J., Chapter 9.

(31) A. Ralston, "A First Course in Numerical Analysis," McGraw-Hill, New York, N. Y., Chapter 8.

shown in Table V. For the most part, these results improve on the theoretical fit (Table III), but, with the exception of $N = 20$, do not do quite so well as the empirical best-fit obtained by minimization.

Moments

The ultimate test of acceptibility for a given probability density function such as eq I–V is how well the given function reproduces the independently observed moments of the distribution. Table VI shows the observed first and second moments, as well as values predicted by the functions for their theoretical and best-fit forms. Interest in polymer statistics centers around the second moments, so we will discuss this point separately from the higher moments.

A. Second Moments. The Exponent b. The exponential dependence of $\langle r^2 \rangle$ on chain length, given by eq 14, has not been critically studied in Monte Carlo experiments until recently[2,32] because the large standard deviation of $\langle r^2 \rangle$ at each N (see paper I) necessitates an exceptionally large sample size to assure the goodness of $\langle r^2 \rangle_{\text{obsd}}$ as an estimator of the true $\langle r^2 \rangle$.

Table VII shows the results of a least-squares analysis on the second moments of Table VI. As a general rule, the smaller the standard deviation, $\sigma(b)$, the more reliable the exponent b and the narrower the confidence limits which can be placed on it. The figures shown represent a five-point fit. We chose to exclude the point $N = 20$ because of the well-known physical "induction period," the region of small N for which the $\langle r^2 \rangle$ relationship is erratic and the value of b itself changes rapidly with N. (Numerical justification of this was already given in paper I.)

(32) E. Loftus and P. J. Gans, *J. Chem. Phys.*, **49**, 3828 (1968).

TABLE VII

LEAST-SQUARES VALUES OF FACTOR a AND EXPONENT b IN THE EQUATION $\langle r^2 \rangle = aN^b$, WITH STANDARD DEVIATION IN b, $\sigma(b)$

	Observed	I	II	III	IV	V
			With Theoretical Parameters			
a	3.205	4.743	3.148	6.793	0.779	3.207
b	1.221	1.000	1.209	1.126	1.262	1.220
$\sigma(b)$	2.07×10^{-2}	5.06×10^{-4}	1.76×10^{-2}	1.74×10^{-3}	1.06×10^{-2}	2.08×10^{-2}
			With Best-Fit Parameters			
a		3.402	3.235	a	1.146	3.209
b		1.210	1.219		1.224	1.220
$\sigma(b)$		5.55×10^{-2}	1.44×10^{-2}		3.85×10^{-4}	1.74×10^{-2}

a Same as II.

TABLE VIII

VALUES OF REDUCED-MOMENT FUNCTION $\delta(4, 2)$ OBSERVED, AND PREDICTED BY FUNCTIONS I–V WITH BEST-FIT PARAMETERS

		Function			
N	Obsd	I	II, III	IV	V
20	1.3467	1.6667	1.3933	1.4618	1.3897
40	1.3984	1.6667	1.4282	1.4618	1.4075
60	1.4210	1.6667	1.4376	1.4618	1.4165
80	1.4321	1.6667	1.4376	1.4618	1.4262
98	1.4531	1.6667	1.4624	1.4618	1.4560
298	1.4749	1.6667	1.4881	1.4618	1.4591

The interesting features of Table VII may be summarized as follows. It has been widely suggested that $b = {}^6/_5$ for three-dimensional lattices.[23] In fact, functions IV and V were obtained in consideration of the simple cubic and tetrahedral lattices, respectively, while functions I, II, and III were obtained without reference to any lattice. With the exception of the admittedly inapplicable gaussian function, and function III-theoretical, both the observed and predicted values of b are significantly higher than ${}^6/_5$ for all other functions, indicating a value of $b \geq 1.22$, in good agreement with the results of paper I and with the value found by Loftus and Gans.[32] The latter authors used a Monte Carlo model similar to ours, but an entirely different sampling procedure. By contrast, our Monte Carlo results were obtained for free dihedral rotation (not on a lattice), and our highest chain length is 298, not infinity. Nevertheless, the standard deviation in b of 1% or less indicates to us a strict adherence to eq 14, with no change in slope even for much higher N.

The functional dependence of $\langle r^2 \rangle$ on N for functions I, IV, and V is easily derived. The three equations are

$$\langle r^2 \rangle_\mathrm{I} = 1.5/c_2$$

$$\langle r^2 \rangle_\mathrm{IV} = \sigma^2 \Gamma(11/5)/\Gamma(7/5)$$

$$\langle r^2 \rangle_\mathrm{V} = \frac{\Gamma(5/t)}{\alpha_N{}^{2/t}\Gamma(3/t)}$$

It is striking that function IV predicts $\langle r^2 \rangle$ to be proportional to the variance, a fact borne out by our data. The corresponding expressions for functions II and III are far more complicated and are not given here.

As Table VI shows, all five functions accurately reproduced the observed second moments, even though the parameters were adjusted to fit the overall distribution function shape, and not the second moments explicitly. In order to achieve greater discrimination among the functions we need to examine the higher moments.

B. Higher Moments. To avoid dealing with extremely large numbers, we adopt Mazur's practive and define a reduced moment function

$$\delta(p, s) = \frac{\langle r^p \rangle}{\langle r^s \rangle^{p/s}} \tag{18}$$

Observed and computed values of $\delta(4,2)$ are given in Table VIII. The observed $\delta(4, 2)$ increases with N, but appears to approach an upper limit which unfortunately our data are insufficient to extrapolate. Functions I and IV both predict this limiting behavior, since $\langle r^4 \rangle_\mathrm{I} = 15/4c_2{}^2$ and $\langle r^4 \rangle_\mathrm{IV} = \Gamma(3)\sigma^4/\Gamma(7/5)$. Hence

$$\delta_\mathrm{I}(4, 2) = 5/3 = 1.6667$$

and

$$\delta_\mathrm{IV}(4, 2) = \frac{\Gamma(3)\Gamma(7/5)}{[\Gamma(11/5)]^2} = 1.4618$$

By working out several examples, one quickly observes that all the δ's for eq IV are constants, completely independent of σ. This will in general be true for one-parameter functions.

The values of $\delta(4, 2)$ predicted by functions II and V are functions of their respective parameters, of course, but δ_V $(4, 2)$ is a function only of t, and not α_N.

$$\delta_\mathrm{V}(4, 2) = \Gamma(3/t)\Gamma(7/t)/\Gamma^2(5/t)$$

If t were always equal to 3.2, as Mazur states, then $\delta_\mathrm{V}(4, 2) = 1.4376$ for all N. Conversely, the limiting values of t computed from the constant values of $\delta(4, 2)$ given by functions I and IV should be 2.0 and 2.5, respectively. We can conclude not only that t must be significantly less than 3.2, but also that *at least two parameters* are needed in order to describe the observed nonconstant $\delta(4, 2)$ and the higher reduced moments as well.

The higher order reduced moments are given in Table IX. The best-fit functions II–III predicted these moments quite well, as the table shows. The moments are plotted against N^{-1} in Figure 4. Table IX also shows the constant reduced moments predicted by function IV. It can be seen that in every case they fall below the observed values for $N = 298$ (and hence even farther below any possible extrapolated limit), and below the values predicted by II–III.

C. Approach to Gaussian Behavior. A number of observations have already been made which seem to rule out the gaussian function, I, as the limiting function for an infinite chain. The only arguments which might be raised in favor of function I are that $\chi_0{}^2$ drops slowly with increasing N and that t of function V also drops slowly. However, as we have seen, a far more likely limiting value for t is some-

TABLE IX

VALUES OF THE REDUCED MOMENT FUNCTION $\delta(p, s)$ OBSERVED IN MONTE CARLO POLYMER GENERATION, AND PREDICTED BY FUNCTIONS II, III, AND IV USING BEST-FIT PARAMETERS

N	$\delta(2, 1)$	$\delta(4, 2)$	$\delta(6, 3)$	$\delta(6, 2)$	$\delta(8, 4)$	$\delta(10, 5)$
			Observed			
20	1.103	1.347	1.710	2.178	2.212	2.898
40	1.114	1.398	1.846	2.420	2.491	3.393
60	1.119	1.421	1.919	2.547	2.693	3.883
80	1.120	1.432	1.969	2.625	2.853	4.341
98	1.126	1.453	2.007	2.710	2.873	4.179
298	1.132	1.475	2.068	2.827	3.059	4.716
			Predicted (II–III)			
20	1.106	1.394	1.886	2.457	2.683	3.953
40	1.117	1.428	1.971	2.618	2.858	4.287
60	1.119	1.437	1.995	2.662	2.903	4.382
80	1.119	1.437	1.993	2.662	2.903	4.387
98	1.125	1.462	2.052	2.782	3.028	4.632
298	1.133	1.488	2.116	2.909	3.164	4.902
			Predicted (IV)			
All N	1.113	1.462	2.038	2.763	2.968	4.452

Figure 4. Various observed reduced moments plotted as a function of reciprocal chain length.

thing between 2.5 and 3.0, rather than the gaussian value of 2.0.

The arguments against function I are strong. First, observed and predicted second moments indicate that $\langle r^2 \rangle \propto N^b$, where $b > {}^6/_5 > 1$, that is, that the limit as $N \to \infty$ of $\langle r^2 \rangle / Nl^2$ diverges. Secondly, if eq 2 is rewritten

$$f(r) = k \exp\left[-a\left(r^2 - \frac{b}{a}r\right)\right]$$

then even though the coefficient b decreases, the ratio b/a actually increases. As Table IV shows, this perturbing influence, which results physically from the introduction of excluded volume, increases approximately as $N^{1/2}$.

D. Comparison with Other Numerical Studies. In addition to those of Mazur, Monte Carlo studies of the distribution function on a tetrahedral lattice have been performed by Schatzki[27] and by Gans.[33] Schatzki's data for walks of up to 100 steps were fit by standard least-squares techniques to a complete set of orthogonal functions, $\phi_m(r)$, with coefficients a_m, based on an expansion of $f(r)$ in terms of Hermite polynomials

$$f^{1/2}(r) = \sum_{m=0}^{M-1} a_m(N)\phi_m(r) \tag{19}$$

The order of the Hermite functions, $2m + 1$, was chosen to make $f_0^{1/2}(r) = \phi_0(r)$ (i.e., $M = 1$, $a_0 = 1$) coincide with the gaussian distribution, eq 1. The number of terms, M, was chosen as that value beyond which the sum of the squared residuals, $\Delta^2(M)$, changed by less than 0.01%, and turned out to be $M = 4$.

Least-squares residuals and χ^2 values are closely related statistics, as eq 8 shows, the latter measuring the probability that observed data arose from a hypothetical distribution. For walks of 20, 40, 60, 80, and 100 steps, Schatzki's least-squares residuals were 0.037, 0.005, 0.001, 0.001, and 0.001, respectively. Some corresponding χ_0^2 values, computed from his data by us, are 40:503, 60:102, 80:21.96, and 100:36.86, reflecting the trend of the Δ^2 values. (The independently observed data of Gans at $N = 100$ gave $\chi_0^2 = 41.37$.) The exception to this trend is $N = 100$, probably due, as
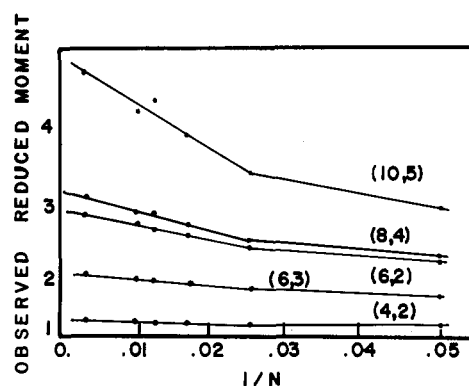
Schatzki himself suggests, to the smallest number of samples (12,000), and hence insufficient representation in the tails of the distribution. None of these χ_0^2 values fall below the statistically acceptable level, although the trend is clearly towards good fit for longer walks. Whether this lack of fit is due to the number M of terms taken, insufficient samples, or inherent numerical error in the least-squares procedure we cannot tell.

The reason for mentioning this study is to assess the predictive properties of eq 19. Assuming good fit above $N = 100$, the data would be completely described by the constants $a_m = a_m(M, N)$. As N increased, it was observed that a_3 quickly vanished, and that a_2 also approached zero, though much more slowly. Hence, above a certain point, N_0, only two terms of eq 19 are needed, the first corresponding to a gaussian term and the second amounting to a nonnegligible perturbation. In fact, a_0 was found to fluctuate about 0.97 for small N and then to decrease nonlinearly from 0.974 at $N = 35$ to 0.948 at $N = 100$. In addition, a_1 was found to increase with N. In other words, above about 35 steps, both the gaussian term, a_0, and the perturbation term, a_1, depart significantly from their gaussian values, namely unity and zero, respectively.

The second moments could be computed from the published a_m and eq 19, and were found to fit $\langle r^2 \rangle = 3.99N^b$ (in units of $3^{1/2}$), where $b = 1.191$, $\sigma(b) = 0.013$. In particular, $\langle r^2 \rangle_{100} = 960.47$ for Schatzki's data, in excellent agreement with the value 953.98 observed by Gans.[33] Our own data at $N = 98$ were found to fit eq 19 with Schatzki's a_m (4,100), but with a slight reduction of his parameter ϵ. The corresponding χ_0^2 value was 54.8, and $\langle r^2 \rangle$ was computed to be 862.98 Å^2, which agrees well with our observed values.

In the absence of analytical expressions for a_0 and a_1, this empirical distribution can be of value only in confirming existing independently observed data, such as those of Mazur or Gans, since at present the only way to obtain the a_m is actually to perform the analysis on collected data. Obviously, once the data have been obtained, the properties are already available and further analysis in unnecessary.

Let us note that any continuous function can be expanded in terms of a complete orthogonal set, so that eq 19 is truly a general form for a distribution function. However, even if a two-, three-, or four-parameter fit is statistically "good," our purpose is better served if we can attach physical meaning to those parameters.

Discussion

Returning to the success of function II–III (best fit), successful fit of distributions and higher moments by two-pa-

(33) P. J. Gans, unpublished data.

rameter functions gives numerical support for the two-parameter theory of polymer solutions.[34] The form of the best function is a perturbed gaussian, $\exp(-ar^2 - br - c)$, with c a normalizing factor. Because the numerical Monte Carlo distribution is observed for a model with explicit long-range correlation, there is no need to resort to the usual series expansion[35] for $\langle r^2 \rangle / \langle r^2 \rangle_0$ in terms of the parameter $z = (c_2/\pi)^{1/2}\beta$. The major analytical problem with such an expansion arises from the slow convergence of the series, and implies validity of the expansion only near the θ point.

The form of eq 2 also suggests that a should be a gaussian-

(34) H. Yamakawa, "Modern Theory of Polymer Solution," Harper and Row, New York, N. Y., 1971.
(35) H. Yamakawa, A. Aoki, and G. Tanaka, *J. Chem. Phys.*, **45**, 1938 (1966).

like parameter which pertains to the molecular geometry, and b a perturbation parameter which contains the "hardness" of the excluded-volume effect. In fact, data from Table IV obey the relation

$$1/a_{\min} = 0.645 N^{4/3}$$

Thus, the best-fit values of the parameter a are proportional to $N^{4/3}$, which is the known N dependence of first-order perturbation theory.

A Comparison of the Effects of Salt and Temperature on Charged and Uncharged Polypeptides in Water[1]

Wayne L. Mattice,* Jung-Teh Lo, and Leo Mandelkern

Department of Biochemistry, Louisiana State University, Baton Rouge, Louisiana   70803, and Department of Chemistry and Institute of Molecular Biophysics, The Florida State University, Tallahassee, Florida   32306.   Received May 3, 1972

ABSTRACT: The hydrodynamic properties of poly(N^5-ω-hydroxyethyl-L-glutamine) and ionized poly(L-glutamic acid) in water differ in a manner which reflects the polyelectrolyte properties of poly(L-glutamic acid). However, the circular dichroism spectra of these two polypeptides are affected in a similar manner by increasing temperature and salt concentration. The ability of increasing temperature and increasing potassium chloride concentration to reduce the intensity of the positive circular dichroism band are cumulative for both poly(N^5-ω-hydroxyethyl-L-glutamine) and ionized poly(L-glutamic acid). On the other hand, in the helix–coil transition of poly(N^5-ω-hydroxypropyl-L-glutamine), the effects of increasing temperature and increasing potassium chloride concentration offset each other. It is concluded that the effects of salts on the circular dichroism of charged polypeptides cannot be accounted for by the formation of α helices and that the salt effect must be more complex than a simple shielding of electrostatic charges on the side chains.

The conformational properties of fully ionized homopolypeptides, notably poly(L-glutamic acid) and poly(L-lysine), have been of interest recently because of the suggestions that they may become partially ordered in dilute aqueous solution when the concentration of added electrolyte is properly chosen.[2-7]

On theoretical grounds Krimm and Mark[2] predicted that, in the absence of added salt, the repulsions between the charges in the side chains would lead to the formation of short ordered regions in the polypeptide backbone with values of the torsion angles ϕ and ψ[8] similar to those observed

* Address correspondence to this author at Louisiana State University.
(1) This work was supported by a contract with the Division of Biology and Medicine, Atomic Energy Commission (L. M.), a Public Health Service Postdoctoral Fellowship from the National Institute of General Medical Sciences (W. L. M.), and a Frederick Gardner Cottrell grant in aid (W. L. M.).
(2) S. Krimm and J. E. Mark, *Proc. Nat. Acad. Sci. U. S.*, **60**, 1122 (1968).
(3) M. L. Tiffany and S. Krimm, *Biopolymers*, **6**, 1379 (1968).
(4) M. L. Tiffany and S. Krimm, *ibid.*, **8**, 347 (1969).
(5) S. Krimm, J. E. Mark, and M. L. Tiffany, *ibid.*, **8**, 695 (1969).
(6) Y. P. Myer, *Macromolecules*, **2**, 624 (1969).
(7) D. G. Dearborn and D. B. Wetlaufer, *Biochem. Biophys. Res. Commun.*, **39**, 314 (1970).
(8) IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970); *J. Mol. Biol.*, **52**, 1 (1970).

for collagen[9,10] and poly(L-proline) II in the solid state.[11] In the absence of added salt, the circular dichroism of fully ionized poly(L-glutamic acid) and poly(L-lysine), as well as native collagen and poly(L-proline) with all peptide bonds in the trans configuration, exhibits a weak positive band followed by a strong negative band about 20 nm toward shorter wavelengths.[3-5] In the presence of sufficiently high salt concentrations, the circular dichroism exhibits only a single moderately strong negative band near 200 nm.[3-5] The salt was believed to reduce the electrostatic interaction of the side chains, leading to the formation of a true random coil and the simultaneous loss of the positive circular dichroism band.[3-5] Recently electron micrographs of poly(L-glutamic acid) precipitated from calcium acetate have been interpreted as being consistent with the extended helical conformation proposed by Krimm and coworkers.[12]

Myer[6] and Dearborn and Wetlaufer[7] have observed similar effects of salts on the circular dichroism of ionized homopolypeptides, but they have offered a different interpretation. Curve fitting could hypothetically account for the

(9) A. Rich and F. H. C. Crick, *J. Mol. Biol.*, **3**, 483 (1961).
(10) G. N. Ramachandran in "Treatise on Collagen," Vol. 1, G. N. Ramachandran, Ed., Academic Press, New York, N. Y., 1967, p 103.
(11) V. Sasisekharan, *Acta Crystallogr.*, **12**, 897 (1959).
(12) H. D. Keith, *Biopolymers*, **10**, 1099 (1971).