

# JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Registered in U. S. Patent Office. © Copyright, 1967, by the American Chemical Society

VOLUME 89, NUMBER 26

DECEMBER 29, 1967

## Physical and Inorganic Chemistry

### Dipole Moments in Relation to Configuration of Polypeptide Chains

Paul J. Flory and Paul R. Schimmel

Contribution from the Department of Chemistry, Stanford University,  
Stanford, California 94305. Received July 27, 1967

**Abstract:** Methods are presented for calculating mean-square dipole moments,  $\langle \mu^2 \rangle$ , of polypeptide chains,  $^+\text{NH}_3\text{-CHRCO}(\text{NHCHRCO})_{x-1}\text{NHCHRCO}_2^-$ , averaged over all configurations of the chain skeleton. They are applicable to chains of any number ( $x + 1$ ) of residues, the residues being in any specified sequence. Contributions to  $\langle \mu^2 \rangle$  from the dipole moments of the amide groups are included. Compared to the contribution from separation of the terminal charges, these group dipole moments make only a very small contribution to  $\langle \mu^2 \rangle$  for the lower polypeptide homologs, owing to the alternating directions of the group moments transverse to the axis of the chain. Dipole moments of glycine peptides are calculated and compared with experimental determinations carried out by Wyman and co-workers on aqueous solutions of homologs for which  $x + 1 = 2-7$ . Agreement for di- and triglycyl peptides is excellent. Departures for  $x + 1 > 3$  reflect diminution of the distance between the termini of the chain by Coulombic attraction. A perturbation treatment of this effect brings theory into agreement with experiment. Dipole moments calculated for the 14 diastereoisomers of the di-, tri-, and tetrapeptides of alanine are compared with the moments obtained from dielectric increments determined by Kenner and co-workers on aqueous solutions of these species. The effects of stereosequence on the dipole moment are well reproduced by the calculations.

The average dimensions of polypeptide chains in the random-coil form have been treated in recent publications.<sup>1-4</sup> The conformational energies of glycol,<sup>2</sup> alanyl,<sup>1,2</sup> and prolyl<sup>4,5</sup> residues situated within a polypeptide molecule were calculated as functions of the angles of rotation,  $\varphi$  and  $\psi$ , about the single bonds, N-C $\alpha$  and C $\alpha$ -C, adjoining the  $\alpha$ -carbon atom. Semiempirical expressions for the interactions between pairs of nonbonded atoms and groups were used for this purpose,<sup>1,2,4,5</sup> and it was shown to be essential to take account of the Coulombic energy of interaction between neighboring amide groups arising from their large electric moments (3.7 D.).<sup>1,2,6</sup> The conformational energy of the polypeptide chain can be resolved into a

sum of contributions, one for each residue. The contribution for residue  $i$  depends upon the pair of rotation angles,  $\varphi_i, \psi_i$ , for that residue.<sup>1,2</sup> The energy,  $V(\varphi_i, \psi_i)$ , associated therewith is sensibly independent of the conformation of its neighbors as specified by  $\varphi_{i-1}, \psi_{i-1}, \varphi_{i+1}$ , and  $\psi_{i+1}$ .<sup>1,2</sup> This fact, easily confirmed by examination of models, greatly facilitates analysis of spatial configurations of the random coil.<sup>1-5</sup>

Owing to the planarity of the amide group, and its overwhelming preference for the *trans* conformation, the spatial configuration of the polypeptide chain may be described in terms of virtual bonds of fixed length (3.8 Å) joining successive  $\alpha$ -carbon atoms in the sequence of residues,<sup>1,2</sup> as shown in Figure 1. The portion of the structure spanned by a virtual bond is designated a unit, which is to be distinguished from an amino acid residue.<sup>7</sup> The orientation of a given virtual bond relative to its immediate predecessor is deter-

- (1) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2791 (1965).
- (2) D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).
- (3) W. G. Miller, D. A. Brant, and P. J. Flory, *ibid.*, **23**, 67 (1967).
- (4) P. R. Schimmel and P. J. Flory, *Proc. Natl. Acad. Sci. U. S. A.*, **58**, 52 (1967).
- (5) P. R. Schimmel and P. J. Flory, in preparation.
- (6) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 663 (1965).

- (7) J. T. Edsall, P. J. Flory, J. C. Kendrew, G. Nemethy, G. N. Ramachandran, and H. A. Scheraga, *Biopolymers*, **4**, 121 (1966).

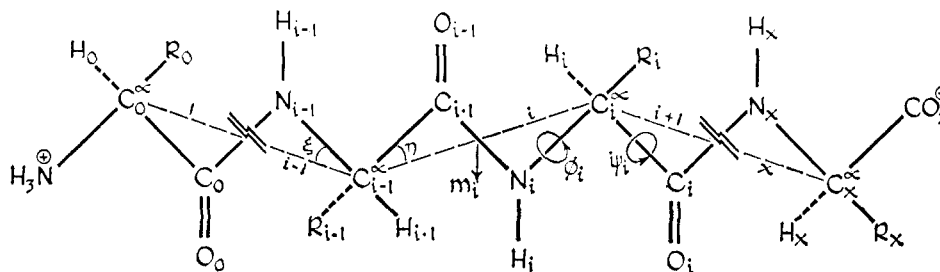


Figure 1. Schematic representation of the  $\alpha$ -L-polypeptide chain with all units *trans*, *i.e.*, with each rotation angle  $\varphi$  and  $\psi$  defined to be zero. Amide groups are rigidly fixed in the planar *trans* configuration. Subscripts denote serial order of amino acid residues and dashed lines are virtual bonds joining successive  $\alpha$ -carbon atoms.

mined by the pair of rotation angles  $\varphi$ ,  $\psi$  preceding the virtual bond in question.<sup>1,2</sup> A Cartesian coordinate system is defined for each virtual bond. The transformation  $\hat{T}_i$  of a vector from reference frame  $i + 1$  affixed to bond  $i + 1$  to reference frame  $i$  affixed to bond  $i$  depends on the rotation angles  $\varphi_i$  and  $\psi_i$ .<sup>1,2</sup> Analysis of the configuration requires the averages of these matrices denoted by  $\langle \hat{T}_i \rangle$  and obtained by averaging  $\hat{T}_i$  over  $\varphi_i$  and  $\psi_i$  with weighting by Boltzmann factors,  $\exp[-V(\varphi_i, \psi_i)/kT]$ . Calculations<sup>1,2</sup> of the characteristic ratio,  $\langle r^2 \rangle_0/xl_u^2$ , of the unperturbed mean-square end-to-end length of the chain to the number  $x$  of peptide units (*i.e.*, virtual bonds) multiplied by the square of the length,  $l_u (= 3.8 \text{ \AA})$ , of a virtual bond are in accord with experimental results<sup>8</sup> for several L-polypeptides. Introduction of D residues in admixture with L residues markedly reduces the value calculated for the characteristic ratio.<sup>3</sup> Available experimental evidence lends support to this prediction.<sup>3</sup> Thus, the average dimensions are sensitive to the stereoregularity of the asymmetric centers of the amino acid residues.<sup>3</sup>

The mean-square dipole moment,  $\langle \mu^2 \rangle$ , is another property which depends on the configuration of the chain. Unlike the mean-square dimension,  $\langle r^2 \rangle_0$ , it is subject to experimental determination for short oligomers as well as for long-chain polymers. Dipole moments of normal alkylene dibromides,<sup>9</sup>  $\text{Br}(\text{CH}_2)_n\text{Br}$ , of dihydropolyfluoromethylenes,<sup>10</sup>  $\text{H}(\text{CF}_2)_n\text{H}$ , and of polyoxyethylenes,<sup>11</sup>  $\text{RO}(\text{CH}_2\text{CH}_2\text{O})_x\text{OR}$ , have been treated according to newer methods<sup>12-17</sup> developed for chain molecules in which neighbor rotations are interdependent.

Some years ago Wyman and co-workers<sup>18</sup> investigated the dipole moments of glycine oligomers in the range from the dipeptide to the heptapeptide. The value of  $\langle \mu^2 \rangle$  was found to be approximately propor-

tional to the number of units. They compared their results with the mean-square separation of the terminal charges estimated assuming free rotation about all skeletal bonds, amide bonds included.<sup>18</sup> More recently the dielectric increments of solutions of the diastereoisomeric di-, tri-, and tetrapeptides of alanine have been measured by Beacham, Ivanov, Kenner, and Sheppard.<sup>19</sup> They found the dielectric increments to vary significantly with the stereo (D or L) composition and sequence of the alanyl residues.<sup>19</sup>

A general treatment of dipole moments of polypeptide chains is presented in this paper. Structural parameters, hindrances to rotation, and the stereochemical configurations of  $\alpha$ -carbon centers are taken into account. The Coulombic attraction between terminal groups  $^+\text{NH}_3$  and  $\text{CO}_2^-$  is treated by a perturbation method. Contributions to  $\langle \mu^2 \rangle$  from the dipole moments of the amide bonds are included as well as that from the charged terminal groups  $^+\text{NH}_3$  and  $\text{CO}_2^-$ . In oligopeptides the latter contribution is dominant. The treatment is applicable to chains of any length comprising any sequence of amino acid residues. The results of Wyman and co-workers<sup>18</sup> on polyglycine homologs and of Kenner and co-workers<sup>19</sup> on diastereoisomeric oligomers of alanine are compared with the theoretical calculations.

## Theoretical Treatment

**The Unperturbed Chain.** A portion of the polypeptide chain is displayed in Figure 1. All structural parameters are assigned the values tabulated in a previous paper.<sup>1</sup> The charge of the  $^+\text{NH}_3$  group is assumed to be located at the center of the nitrogen atom and that of the  $\text{CO}_2^-$  group to be midway between the carboxyl oxygens. The dipole moment,  $\mathbf{m}_i$ , of the  $i$ th amide group is taken to be situated in the plane of this group and at an angle of  $56^\circ$  with the amide bond, *i.e.*, approximately parallel to the N-H bond.<sup>1</sup> Its magnitude is assigned the value  $|\mathbf{m}_i| = 3.7 \text{ D}$ .<sup>1</sup>

The molecular dipole moment  $\mathbf{y}$  for a chain of  $x + 1$  residues, numbered 0, 1, ...,  $x - 1$ ,  $x$ , is given by

$$\mathbf{y} = -e\mathbf{r} + \sum_{i=1}^x \mathbf{m}_i \quad (1)$$

where  $e$  is the electronic charge and  $\mathbf{r}$  is the vector from the  $^+\text{NH}_3$  charge to the  $\text{CO}_2^-$  charge. This vector is given by

$$\mathbf{r} = l_N + \sum_{i=1}^x l_{u,i} + l_C \quad (2)$$

(19) J. Beacham, V. T. Ivanov, G. W. Kenner, and R. C. Sheppard, *Chem. Commun.*, 386 (1965).

- (8) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2788 (1965).  
 (9) W. J. Leonard, R. L. Jernigan, and P. J. Flory, *J. Chem. Phys.*, **43**, 2256 (1965).  
 (10) T. W. Bates and W. H. Stockmayer, *ibid.*, **45**, 2321 (1966); T. W. Bates, *Trans. Faraday Soc.*, **63**, 1825 (1967).  
 (11) J. E. Mark and P. J. Flory, *J. Am. Chem. Soc.*, **88**, 3702 (1966).  
 (12) S. Lifson, *J. Chem. Phys.*, **30**, 964 (1959).  
 (13) K. Nagai, *ibid.*, **31**, 1169 (1959).  
 (14) C. A. J. Hoeve, *ibid.*, **32**, 888 (1960).  
 (15) M. V. Volkenstein, "Configurational Statistics of Polymeric Chains," (Engl. Transl.), Interscience Publishers Inc., New York, N. Y., 1963.  
 (16) P. J. Flory, *Proc. Natl. Acad. Sci. U. S. A.*, **51**, 1060 (1964); P. J. Flory and R. L. Jernigan, *J. Chem. Phys.*, **42**, 3509 (1965).  
 (17) T. M. Birshtein and O. B. Ptitsyn, "Conformation of Macromolecules," Interscience Publishers Inc., New York, N. Y., 1966.  
 (18) A general review of this work with pertinent literature references is given by J. T. Edsall and J. Wyman, "Biophysical Chemistry," Vol. I, Academic Press Inc., New York, N. Y., 1958, Chapter 6.

where  $l_{u,i}$  is the vector representing the  $i$ th "virtual bond," *i.e.*, the hypothetical bond drawn between the  $(i - 1)$ th and  $i$ th  $\alpha$ -carbon atoms (see Figure 1);<sup>1</sup>  $l_N$  is the vector for the chemical bond connecting the terminal nitrogen with the first  $\alpha$  carbon ( $i = 0$ );  $l_C$  is the vector joining  $C_{\alpha}$  with the assumed location of the negative charge of the carboxyl group. The scalar magnitudes of these bonds are  $l_N = 1.47 \text{ \AA}$ ,<sup>1</sup>  $l_u = 3.80 \text{ \AA}$  for  $i = 1, 2, \dots, x$ , and  $l_C = 2.10 \text{ \AA}$ .<sup>20</sup> It is important to observe that the vector  $\mathbf{r}$  (and its scalar magnitude  $r$ ) departs from the chain vector  $\mathbf{r}$ , previously defined<sup>1-4</sup> as the distance between terminal  $\alpha$ -carbon atoms. For the present purposes, it is necessary to include also the terminal bonds  $l_N$  and  $l_C$  specified in eq 2. The difference between the respective definitions of  $\mathbf{r}$  is trivial of course for very long chains.

It will be expedient to replace eq 1 for the molecular dipolar moment by the following sum of group moments  $\mathbf{p}_i$

$$\mathbf{u} = \sum_{i=0}^{x+1} \mathbf{p}_i \quad (3)$$

where

$$\begin{aligned} \mathbf{p}_0 &= -el_N \\ \mathbf{p}_i &= -el_u + \mathbf{m}_i \quad 1 \leq i \leq x \\ \mathbf{p}_{x+1} &= -el_C \end{aligned} \quad (4)$$

Equations 3 and 4 require, of course, that all vectors be expressed in the same coordinate system. The resultant moment obtained by transformation of these vectors to a common basis depends of course on the configuration of the chain.

Pursuant to the execution of the required transformations, we define a right-handed coordinate system  $X_i Y_i Z_i$  for each virtual bond. The axis  $X_i$  is taken in the direction of virtual bond  $i$ ;  $Y_i$  lies in the plane of the  $i$ th peptide unit, its direction being chosen to form an acute angle with virtual bond  $i - 1$ ;  $Z_i$  is taken in the direction perpendicular to this plane as required to complete a right-handed coordinate system. The moment  $\mathbf{p}_i$ , expressed in matrix form as a column vector in this reference frame is

$$\mathbf{p}_i = -e \begin{bmatrix} l_{u,i} \\ 0 \\ 0 \end{bmatrix} + \mathbf{m}_i \quad 1 \leq i \leq x \quad (5)$$

The terminal bonds are similarly written

$$\mathbf{p}_0 = -e \begin{bmatrix} l_N \\ 0 \\ 0 \end{bmatrix} \quad (6)$$

$$\mathbf{p}_{x+1} = -e \begin{bmatrix} l_C \\ 0 \\ 0 \end{bmatrix} \quad (7)$$

Let  $\hat{\mathbf{T}}_i$  denote the transformation from the coordinate system  $X_{i+1} Y_{i+1} Z_{i+1}$  of virtual bond  $i + 1$  to that of  $i$ . This transformation can be compounded from a series of three transformations through coordinate systems  $(xyz)$  similarly defined for the skeletal bonds adjoining the  $\alpha$ -carbon atom.<sup>1</sup> Thus,  $l_{u,i+1}$  will first be transformed from its representation in reference frame  $X_{i+1} Y_{i+1} Z_{i+1}$  to the coordinate system for the  $C_i^{\alpha}-C_i$

skeletal bond of residue  $i$ . Next it will be transformed to the coordinate system for skeletal bond  $N_i-C_i^{\alpha}$ , and finally to  $X_i Y_i Z_i$ . Each of these transformations involves a rotation about the initial  $z$  (or  $Z$ ) axis equal to the angle  $\theta$  between the respective  $x$  (or  $X$ ) axes, and a further rotation,  $\pi - \varphi$ , about the  $x$  (or  $X$ ) axis of the final reference frame.<sup>1</sup> They are therefore of the form

$$\mathbf{T}(\theta, \varphi) = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ \sin \theta \cos \varphi & -\cos \theta \cos \varphi & \sin \varphi \\ \sin \theta \sin \varphi & -\cos \theta \sin \varphi & -\cos \varphi \end{bmatrix} \quad (8)$$

The three transformations are  $\mathbf{T}(-\eta, \pi + \psi)$ ,  $\mathbf{T}(\theta^{\alpha}, \varphi)$ , and  $\mathbf{T}(\xi, \pi)$ , respectively, where  $\theta^{\alpha}$  (*ca.*  $70^\circ$ ) is the supplement of the skeletal bond angle at the  $\alpha$  carbon, and  $\varphi$  and  $\psi$  are the skeletal bond rotations defined above and in Figure 1. The structural parameters  $\eta$  and  $\xi$  are defined also in Figure 1. The transformation from the virtual bond coordinate system  $X_{i+1} Y_{i+1} Z_{i+1}$  to that of  $X_i Y_i Z_i$  is therefore

$$\hat{\mathbf{T}}_i = \mathbf{T}(\xi, \pi) \mathbf{T}(\theta^{\alpha}, \varphi_i) \mathbf{T}(-\eta, \pi + \psi_i) \quad (9)$$

where  $0 < i < x$ . The parameters  $\xi$ ,  $\eta$ , and  $\theta^{\alpha}$  are fixed;  $\hat{\mathbf{T}}_i$  is to be regarded therefore as a function of the rotation angles  $\varphi_i$  and  $\psi_i$  (see Figure 1). Transformations required for terminal bonds are given by

$$\hat{\mathbf{T}}_0 = \mathbf{T}(\theta^{\alpha}, 0) \mathbf{T}(-\eta, \pi + \psi_0) \quad (10)$$

and

$$\hat{\mathbf{T}}_x = \mathbf{T}(\xi, \pi) \mathbf{T}(\theta^{\alpha}, \varphi_x) \quad (11)$$

as may readily be confirmed. The molecular dipole moment  $\mathbf{u}$  expressed in the coordinate system for the zeroth skeletal bond  $+N-C_0^{\alpha}$  is

$$\mathbf{u} = \mathbf{p}_0 + \sum_{j=1}^{x+1} \prod_{i=0}^{j-1} \hat{\mathbf{T}}_i \mathbf{p}_j \quad (12)$$

The mean-square dipole moment for the chain as a whole is

$$\begin{aligned} \langle \mu^2 \rangle_0 &= \langle \mathbf{u}^T \mathbf{u} \rangle_0 = \sum_{i=0}^{x+1} p_i^2 + \\ &2 \sum_{j=1}^{x+1} \sum_{i=0}^{j-1} \mathbf{p}_i^T \langle \hat{\mathbf{T}}_i \hat{\mathbf{T}}_{i+1} \dots \hat{\mathbf{T}}_{j-1} \rangle \mathbf{p}_j \end{aligned} \quad (13)$$

where the superscript  $\mathbf{T}$  denotes the transpose, and angle brackets signify the statistical mechanical average of the quantity enclosed therein. Subscripts zero denote the unperturbed chain, *i.e.*, the chain free of constraints or perturbations such as may be imposed by excluded volume interactions, and in particular by the Coulombic attraction between terminal charges. Because rotations  $\varphi$  and  $\psi$  within a given residue are sensibly independent of those in adjacent residues,<sup>1,2</sup> the average product of  $\hat{\mathbf{T}}$  matrices may be replaced by the product of average  $\hat{\mathbf{T}}$  matrices.<sup>1-3</sup> Equation 13 consequently simplifies to

$$\langle \mu^2 \rangle_0 = \sum_{i=0}^{x+1} p_i^2 + 2 \sum_{0 \leq h < j \leq x+1} \mathbf{p}_h^T \langle \prod_{i=h}^{j-1} \hat{\mathbf{T}}_i \rangle \mathbf{p}_j \quad (14)$$

The matrix  $\langle \hat{\mathbf{T}}_i \rangle$  is the matrix  $\hat{\mathbf{T}}_i$  averaged over the variable angles  $\varphi_i$  and  $\psi_i$ ; that is, for  $0 < i < x$

(20) L. Pauling, "The Nature of the Chemical Bond," Cornell University Press, Ithaca, N. Y., 1960, p 275.

$$\langle \hat{\mathbf{T}}_i \rangle = \frac{\int_0^{2\pi} \int_0^{2\pi} \hat{\mathbf{T}}_i \exp[-V(\varphi_i, \psi_i)/kT] d\varphi_i d\psi_i}{\int_0^{2\pi} \int_0^{2\pi} \exp[-V(\varphi_i, \psi_i)/kT] d\varphi_i d\psi_i} \quad (15)$$

the energy associated with the conformation  $\varphi_i, \psi_i$  for residue  $i$  being represented by  $V(\varphi_i, \psi_i)$ . Averages of terminal transformation matrices are

$$\langle \hat{\mathbf{T}}_0 \rangle = \mathbf{T}(\theta^\alpha, 0) \langle \mathbf{T}(-\eta, \pi + \psi_0) \rangle \quad (16)$$

and

$$\langle \hat{\mathbf{T}}_x \rangle = \mathbf{T}(\xi, \pi) \langle \mathbf{T}(\theta^\alpha, \varphi_x) \rangle \quad (17)$$

where the quantities within angle brackets on the right-hand side of eq 16 and 17 represent averages analogous to eq 15. The matrix  $\langle \hat{\mathbf{T}}_0 \rangle$  is averaged over a single angle  $\psi_0$ , subject to the potential  $V(\psi_0)$ ; the matrix  $\langle \hat{\mathbf{T}}_x \rangle$  is averaged over  $\varphi_x$  and  $\psi_x$ , subject to the potential  $V(\varphi_x, \psi_x)$ . (Although  $\hat{\mathbf{T}}_x$  is a function only of  $\varphi_x$ , apart from the aforementioned structural parameters, the potential affecting the rotation  $\varphi_x$  about the  $\text{N}_x\text{-C}_x^\alpha$  bond depends on the rotation  $\psi_x$  about the  $\text{C}_x^\alpha\text{-C}_x$  bond. Hence, the matrix  $\hat{\mathbf{T}}_x$  must be averaged over both  $\varphi_x$  and  $\psi_x$ , subject to the potential  $V(\varphi_x, \psi_x)$ .) The formulation of these energies will be taken up later.

The evaluation of eq 14 may be performed by taking the serial product of matrices  $\mathcal{G}_i$  in accordance with methods developed previously.<sup>16</sup> The result is

$$\langle \mu^2 \rangle_0 = 2[1 \ 0 \ 0 \ 0 \ 0] \prod_{i=0}^{x+1} \mathcal{G}_i \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (18)$$

where

$$\mathcal{G}_i = \begin{bmatrix} 1 & \mathbf{p}^T \langle \hat{\mathbf{T}} \rangle & p^2/2 \\ \mathbf{0} & \langle \hat{\mathbf{T}} \rangle & \mathbf{p} \\ 0 & \mathbf{0} & 1 \end{bmatrix}_i \quad (19)$$

with  $i = 0, 1, \dots, x, x + 1$ . The 0's which occur in the  $5 \times 5$  matrix  $\mathcal{G}_i$  are null matrices of orders required to conform; the serial index  $i$  appended to  $\mathcal{G}_i$  is understood to apply to  $\mathbf{p}$ ,  $p^2$ ,  $\mathbf{p}^T$ , and  $\langle \hat{\mathbf{T}} \rangle$ . The matrix  $\langle \hat{\mathbf{T}}_{x+1} \rangle$  is left undefined; only elements of the last column of  $\mathcal{G}_{x+1}$  enter the product in eq 18.

The foregoing scheme departs in one respect from previous formulations: the  $\mathcal{G}$  matrices have been amended to include diagonal terms of the first sum in eq 14. These terms have previously been treated separately. The last element,  $p^2/2$ , in the first row of  $\mathcal{G}$  takes account of them.

The matrix  $\mathcal{G}_i$  depends exclusively on the nature of the  $i$ th amino acid residue. Calculation of  $\langle \mu^2 \rangle_0$  for a polypeptide chain of any length or amino acid sequence is readily accomplished by taking the serial product of  $\mathcal{G}$  matrices with the aid of a computer.

**Perturbation of the Configuration by Coulombic Interaction between Terminal Charges.** Electrostatic attraction between the terminal charges may be expected to decrease somewhat their distance of separation  $r$ , and hence to lower the mean-square molecular dipole moment  $\langle \mu^2 \rangle$ . If this effect were large, it might alter the contribution from the amide group moments,  $\mathbf{m}_i$ , as well as that arising from the separation of the terminal

charges. We shall find the perturbation by the Coulombic attraction between chain termini to be small, though not negligible in general. Moreover, the contribution to  $\langle \mu^2 \rangle$  from the terminal charges dominates overwhelmingly that from the  $\mathbf{m}_i$  for chains of moderate length. Hence, it will be legitimate to treat only the perturbation of  $e^2 \langle r^2 \rangle_0 \cong \langle \mu^2 \rangle_0$  by the terminal charge interaction.

Let  $W_0(\mathbf{r})$  represent the probability density distribution of end-to-end vectors  $\mathbf{r}$  for unperturbed chains of given structure; that is,  $W_0(\mathbf{r})d\mathbf{r}$  is the probability that the chain vector is  $\mathbf{r}$  within the volume element  $d\mathbf{r}$ . Specifically, the chain thus represented is subject only to the conformational potentials  $V$  cited above for the units comprising the chain. These potentials are of short range in the sense that each depends only on pairs of consecutive bond rotation angles; for the cases considered, this dependence involves pairs  $(\varphi_i, \psi_i)$  only, as pointed out above. The distribution function can be expressed quite generally as<sup>21</sup>

$$W_0(\mathbf{r}) = (3/2\pi \langle r^2 \rangle_0)^{3/2} \exp[-3r^2/2 \langle r^2 \rangle_0] \times [(1 + 15h_4 + \dots) - (30h_4/\langle r^2 \rangle_0)r^2 + \dots] \quad (20)$$

where  $h_4 = -(1/2^3)(1 - 3\langle r^4 \rangle_0/5\langle r^2 \rangle_0^2)$ . Higher terms included explicitly depend on the fourth moment,  $\langle r^4 \rangle_0$ , which could be evaluated if desired.<sup>16</sup> For present purposes the first term of the series will suffice, *i.e.*, approximation by the Gaussian function

$$W_0(\mathbf{r}) = (3/2\pi \langle r^2 \rangle_0)^{3/2} \exp(-3r^2/2 \langle r^2 \rangle_0) \quad (21)$$

will prove adequate.

The perturbing potential is of the form

$$V' = -e^2/\epsilon r \quad (22)$$

where  $\epsilon$  is the effective dielectric constant in the region between the two charges. Hence, the probability that the magnitude of  $\mathbf{r}$  is in the range  $r$  to  $r + dr$  is

$$4\pi r^2 W(\mathbf{r})d\mathbf{r} = \text{constant} \times r^2 W_0(\mathbf{r}) \exp(B/r) dr \quad (23)$$

where  $B = e^2/\epsilon kT = 1.67 \times 10^5/\epsilon T$  in angstrom units. The mean-square dipole moment is given accordingly by

$$\langle \mu^2 \rangle = \frac{e^2 \int_a^\infty r^4 W_0(\mathbf{r}) \exp(B/r) dr}{\int_a^\infty r^2 W_0(\mathbf{r}) \exp(B/r) dr} \quad (24)$$

where  $a$  is an arbitrary lower limit of a few angstroms which is introduced to avoid the singularity at  $r = 0$ . The integrals may be evaluated numerically.

In aqueous media where  $\epsilon$  is large,  $B/r$  is less than unity over most of the significant range  $a < r < 2\langle r^2 \rangle_0^{1/2}$ . Hence, series expansion of  $\exp(B/r)$  in eq 24 to second order in  $B/r$  followed by analytical integration from a lower limit  $a = 0$  may be useful. Substitution of eq 21 for  $W_0(\mathbf{r})$  followed by integration from  $r = 0$  to  $\infty$ , and expression of the result in a series up to second order in  $B/\langle r^2 \rangle_0^{1/2}$ , yields

$$\langle \mu^2 \rangle = \langle \mu^2 \rangle_0 [1 - (2/3\pi)^{1/2} B/\langle r^2 \rangle_0^{1/2} - (1 - 2/\pi) B^2/\langle r^2 \rangle_0] \quad (25)$$

Terms in higher powers of  $B/\langle r^2 \rangle_0^{1/2}$  cannot be developed because of divergence of integrals over higher

(21) K. Nagai, *J. Chem. Phys.*, **38**, 924 (1963).

terms in the series expansion of  $\exp(B/r)$  in eq 24. In fact, the result given by eq 25 is exact only in the limit  $B/\langle r^2 \rangle_0^{1/2} = 0$ . If  $B/\langle r^2 \rangle_0^{1/2}$  is small, however, eq 25 may afford a satisfactory approximation for replacing numerical integration of eq 24.

In the approximation that  $\langle \mu^2 \rangle_0 = e^2 \langle r^2 \rangle_0$ , which will be shown to hold very well for the lower peptides, eq 25 may be written

$$\langle \mu^2 \rangle / \langle \mu^2 \rangle_0 = 1 - (2/3\pi)^{1/2} B e / \langle \mu^2 \rangle_0^{1/2} \quad (25')$$

to terms of first order. This expression may be expected to hold only for very small perturbations. Its actual range can be assessed by comparison with results of numerical integration, which should be valid over a much greater range of  $B/\langle r^2 \rangle_0^{1/2}$ .

### Numerical Calculations and Comparison with Experimental Results

**Averaged Transformation Matrices.** Conformational energies  $V(\varphi_i, \psi_i)$  for glycol and for D- and L-alanyl residues situated within a polypeptide chain comprising these and similar residues have been calculated previously<sup>1,2</sup> from empirical estimates of the bond torsional strain and of the interaction energies between all pairs of nonbonded atoms whose distance of separation depends exclusively on the rotation angles  $\varphi$  and  $\psi$  for that residue.<sup>1,2</sup> The nonbonded interactions taken into account comprised van der Waals repulsions, London attractions, and electrostatic interactions<sup>1,2</sup> between sets of partial charges chosen to represent the electrical asymmetry of the amide group.<sup>2</sup> The averaged matrices,  $\langle \hat{T}_i \rangle$ , calculated according to eq 15 from these conformational energies are<sup>2,22</sup>

$$\langle \hat{T}_{i,\text{gly}} \rangle = \begin{bmatrix} 0.36 & -0.077 & 0.0 \\ -0.092 & -0.37 & 0.0 \\ 0.0 & 0.0 & -0.12 \end{bmatrix} \quad 0 < i < x \quad (26)$$

$$\langle \hat{T}_{i,\text{L-ala}} \rangle = \begin{bmatrix} 0.51 & 0.20 & 0.59 \\ -0.046 & -0.61 & 0.21 \\ 0.65 & -0.23 & -0.30 \end{bmatrix} \quad 0 < i < x \quad (27)$$

$$\langle \hat{T}_{i,\text{D-ala}} \rangle = \begin{bmatrix} 0.51 & 0.20 & -0.59 \\ -0.046 & -0.61 & -0.21 \\ -0.65 & 0.23 & -0.30 \end{bmatrix} \quad 0 < i < x \quad (28)$$

Calculation of  $\langle \hat{T}_{0,\text{gly}} \rangle$  according to eq 16 required evaluation of the potential  $V(\psi_0)$  associated with rotation about the  $C_0^\alpha-C_0$  bond, whereby the position of the  $^+\text{NH}_3$  group is varied with respect to the first peptide unit of the chain. For this purpose nonbonded interactions of  $^+\text{NH}_3$  were replaced by hard-sphere potentials, the minimum permitted distance of approach between any pair of atoms or groups being taken as the sum of their van der Waals radii.<sup>23,24</sup> The ion-dipole interaction between  $^+\text{NH}_3$  and the adjacent amide group

(22) A temperature of 37° was assumed in the calculation of these matrices.<sup>2</sup> Inasmuch as their variations with temperature are inconsequential over a small temperature range, they may be employed to calculate dipole moments at temperatures differing somewhat from 37°.

(23) A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).

(24) A "small" radius of 1.55 Å was assumed for  $^+\text{NH}_3$ , this being the van der Waals radius of nitrogen.<sup>23</sup> Increasing this radius by ca. 0.2 Å virtually excludes all values of  $\psi_0$ .

was included in  $V(\psi_0)$  as estimated from the array of charges chosen by Brant, *et al.*,<sup>2</sup> to represent the electrical asymmetry of the amide group. To calculate  $\langle \hat{T}_{0,\text{L-ala}} \rangle$  and  $\langle \hat{T}_{0,\text{D-ala}} \rangle$ , the nonbonded interactions of the  $-\text{CH}_3$  substituent at the  $\alpha$  carbon must also be taken into account. These interactions were calculated with the potential functions previously reported.<sup>2</sup> The nonbonded interactions of the hydrogen atom attached to  $C_0^\alpha$  were also included in the calculation of  $V(\psi_0)$  for  $\langle \hat{T}_{0,\text{L-ala}} \rangle$  and  $\langle \hat{T}_{0,\text{D-ala}} \rangle$ . These interactions make an inconsequential contribution to  $V(\psi_0)$ , however. Each  $\langle \hat{T}_0 \rangle$  matrix was numerically computed by summing the integrals in eq 16 in 10° increments of  $\psi_0$ . The assumed temperature in each calculation was that of the relevant experimental measurements (30.5 and 25° for alanyl and glycol peptides, respectively). These zeroth transformation matrices are

$$\langle \hat{T}_{0,\text{gly}} \rangle = \begin{bmatrix} 0.36 & -0.027 & 0.0 \\ 0.86 & -0.39 & 0.0 \\ 0.0 & 0.0 & -0.12 \end{bmatrix} \quad (29)$$

$$\langle \hat{T}_{0,\text{L-ala}} \rangle = \begin{bmatrix} 0.36 & -0.015 & 0.93 \\ 0.85 & -0.40 & -0.34 \\ 0.37 & 0.92 & -0.13 \end{bmatrix} \quad (30)$$

The matrix  $\langle \hat{T}_{0,\text{D-ala}} \rangle$  is obtained from eq 30 by reversing the signs of the 13, 23, 31, and 32 elements of  $\langle \hat{T}_{0,\text{L-ala}} \rangle$ .

The averaged matrices,  $\langle \hat{T}_x \rangle$ , were computed according to eq 17. The potential,  $V(\varphi_x, \psi_x)$ , was calculated from nonbonded interactions<sup>2</sup> between all atoms separated by distances which are dependent upon one or both of the rotation angles  $\varphi_x$  and  $\psi_x$ , and upon these two rotation angles only. The electrostatic interaction of the anionic charge of  $\text{CO}_2^-$  with the partial charges<sup>2</sup> of the adjacent amide group was included in  $V(\varphi_x, \psi_x)$ . Additionally, the torsional strain energy for rotations about the  $\text{N}_x-C_x^\alpha$  bond was taken into account.<sup>2</sup> (Lack of relevant experimental data precludes estimation of this energy for rotations about the  $\text{C}_x^\alpha-C_x\text{O}_2^-$  bond. The energy of bond torsional strain is a relatively unimportant determinant of the configuration of polypeptide chains,<sup>2</sup> however.) The matrices were computed by summing the integrals in eq 17 in 30° increments of  $\varphi_x$  and of  $\psi_x$ . Temperatures of 30.5 and 25° were assumed for the calculation of  $\langle \hat{T}_x \rangle$  for alanyl residues and for glycol residues, respectively. The resulting matrices are

$$\langle \hat{T}_{x,\text{gly}} \rangle = \begin{bmatrix} 0.53 & 0.84 & 0.0 \\ 0.76 & -0.52 & 0.0 \\ 0.0 & 0.0 & -0.92 \end{bmatrix} \quad (31)$$

$$\langle \hat{T}_{x,\text{L-ala}} \rangle = \begin{bmatrix} 0.52 & 0.85 & 0.077 \\ 0.72 & -0.50 & 0.33 \\ 0.32 & -0.12 & -0.87 \end{bmatrix} \quad (32)$$

The matrix  $\langle \hat{T}_{x,\text{D-ala}} \rangle$  is obtained from  $\langle \hat{T}_{x,\text{L-ala}} \rangle$  by reversing the signs of the 13, 23, 31, and 32 elements of  $\langle \hat{T}_{x,\text{L-ala}} \rangle$ .

**Dipole Moments of Diastereomeric Alanine Oligomers.** We have calculated the dipole moments for the 14 diastereoisomers of di-, tri-, and tetraalanine peptides. These calculations have been carried out using eq 18 and 19 with  $\langle \hat{T} \rangle$  given by eq 27, 28, 30, and 32 (with appropriate sign inversions for D residues;

see above), and with the  $\mathbf{p}_i$  furnished by eq 5-7. Values of  $\mathbf{m}_i$ ,  $l_N$ ,  $l_u$ , and  $l_C$  given above were adopted for these calculations. Root-mean-square dipole moments  $\langle \mu^2 \rangle_0^{1/2}$  thus obtained are given in the third column of Table I. The fourth column gives the dipole moments  $e\langle r^2 \rangle_0^{1/2}$  arising solely from the separation of the terminal charges, the  $\mathbf{m}_i$  being set equal to zero.

Values of the dipole moments calculated by including the  $\mathbf{m}_i$  are consistently larger than those found with the  $\mathbf{m}_i = 0$ , but the difference is very small. The latter feature doubtless is a consequence of the orientations of the group dipole moments approximately transverse to the chain axes, with their directions alternating from one unit to the next in the planar form of the chain (see Figure 1). The moment  $\mathbf{m}_i$  makes an angle of  $100^\circ$  with  $l_{u,i}$ . The projection of  $\mathbf{m}_i$  on  $l_{u,i}$ , though small, is therefore in the direction of  $-l_{u,i}$ . Hence, it augments the dipole  $-er$  arising from the charges at the chain termini.

The effect of Coulombic attraction between the terminal charges of the polyaniline chains has been left out of account. For these oligomers of alanine, in which steric constraints are severe<sup>1,2</sup> and dipolar interactions between amide groups confer a preference for extended conformations,<sup>1,2</sup> approximation of the distribution  $W(\mathbf{r})$  by the Gaussian function is dubious. The foregoing perturbation treatment carried out on this basis is therefore questionable.

Dielectric increments of all of the diastereoisomers of di-, tri-, and tetraalanine have been measured in aqueous solution by Beacham, Ivanov, Kenner, and Sheppard.<sup>19</sup> We have calculated mean-square dipole moments of the dissolved peptides from these data by the procedure described by Edsall and Wyman.<sup>25</sup> The small contributions from inductive effects, which may be estimated from the optical polarizabilities, were ignored in these calculations.<sup>18</sup> The peptide molar volume required for the calculation of the dipole moment from the dielectric increment was obtained from the measured molar volume of the glycine peptide of the same degree of polymerization  $x$ .<sup>26</sup> To this volume was added 16.3 ml for each residue, this being the volume attributable to the  $\text{CH}_2$  group<sup>27</sup> by which the alanyl residue differs from glycol. Accurate molar volumes of the residues are not required. Values of  $\langle \mu^2 \rangle^{1/2}$  calculated in this manner from the dielectric increments reported by Beacham, *et al.*,<sup>19</sup> are given in the second column of Table I.

Dipole moments obtained from dielectric increments in aqueous solution are subject to error in their absolute magnitudes. Relative magnitudes for the various diastereoisomeric homologs should be significant, however. Uncertainties in the method notwithstanding, observed and calculated moments for the dipetides agree very well. Calculated values for the tripeptides exceed those observed by 10 to 15%. The discrepancy increases for the tetrapeptides. The direction of the deviation and its increase with chain length are consistent with the hypothesis that Coulombic attraction between terminal groups is responsible for the deviation. Whereas those conformations of the dipeptide of

alanine which bring the terminal groups closer together are precluded by steric overlaps and by dipolar interactions between the amide groups, the more copious array of configurations accessible to higher homologs affords opportunity for the Coulombic interaction between charged terminal groups to bias the average extension  $\langle r^2 \rangle$ , and hence  $\langle \mu^2 \rangle$ . Quantitative account of the depression of  $\langle \mu^2 \rangle$  by the Coulombic attraction between charges has not been attempted for these peptides for the reasons stated above. The magnitude of the effect for the tetramer may be shown, however, to be commensurate with predictions from eq 25, the effective dielectric constant  $\epsilon$  being assigned its value for water.

The diastereoisomeric peptides for given degrees of polymerization (1, 2, and 3) are listed in Table I in

Table I. Dipole Moments of Oligomers of Alanine at  $30.5^\circ$  in Debyes

Peptide	Exptl <sup>a</sup>	Calcd (eq 18)		Relative values	
		incl $\mathbf{m}_i$	$\mathbf{m}_i = 0$	Exptl <sup>a</sup>	Calcd (eq 18) incl $\mathbf{m}_i$
LL	28.8	29.3	28.9	1.00	1.00
LD	26.8	27.8	27.4	0.93	0.95
LLL	40.1	44.6	43.7	1.0	1.0
LLD	37.2	43.1	42.1	0.93	0.97
DLL	35.5	40.6	39.7	0.89	0.91
DLD	33.9	38.2	37.3	0.85	0.86
LLLL	48.0	58.2	56.8	1.00	1.00
DLLL	45.6	53.6	52.2	0.95	0.92
LLLD	44.5	56.2	54.9	0.93	0.97
DDLL	43.8	53.2	52.0	0.91	0.91
DLLD	40.7	52.2	50.8	0.85	0.90
LLDL	40.7	50.3	49.1	0.85	0.86
LDLL	38.8	45.9	44.8	0.81	0.79
LDLD	36.8	43.0	41.8	0.77	0.74

<sup>a</sup> From the dielectric increments determined in aqueous solutions by Beacham, Ivanov, Kenner, and Sheppard.<sup>19</sup>

order of decreasing experimental dipole moment. The same order is reproduced in the calculated values with one exception: the order of LLLD and DLLL according to the calculations is the reverse of that reported to have been observed. Otherwise, the effect of stereosequence on the dipole moment is well reproduced by the calculations. Inasmuch as the dipole moment depends overwhelmingly on  $\langle r^2 \rangle^{1/2}$ , this conclusion reduces to confirmation of the configurational-statistical calculations of the chain dimensions.<sup>1-3</sup>

It will be observed that the lowest dipole moments for the di-, tri-, and tetrapeptides occur for the alternating LD, DLD, and LDLD species. Inasmuch as  $\langle \mu^2 \rangle_0$  is very nearly proportional to  $\langle r^2 \rangle_0$ , these results confirm those of Miller, *et al.*,<sup>3</sup> who found by calculation that strictly alternating D,L copolymers of alanine should exhibit the lowest dimensions of any stereochemical composition or sequence.

The relative decrease in averaged distance between chain termini, caused by their mutual Coulombic attraction, should be approximately the same for peptides of the same degree of polymerization  $x$ . Experimental and calculated relative values of  $\langle \mu^2 \rangle$  for a given  $x$  are presented in columns 5 and 6, respectively,

(25) Reference 18, pp 367-371.

(26) Reference 18, p 372.

(27) J. T. Edsall in "Proteins, Amino Acids, and Peptides," E. J. Cohn and J. T. Edsall, Ed., Reinhold Publishing Corp., New York, N. Y., 1943, p 157.

of Table I. Theory and experiments are in quantitative accord for these relative values of  $\langle \mu^2 \rangle$ , except for the one discrepancy noted above.

**Glycine Oligomers.** Dipole moments of the glycine peptides  $^+\text{NH}_3\text{CH}_2\text{CO}[\text{NHCH}_2\text{CO}]_{x-1}\text{NHCH}_2\text{CO}_2^-$  with  $x + 1 = 2$  to 7 determined in aqueous solutions<sup>18</sup> are shown by the points plotted in Figure 2. Unperturbed moments  $\langle \mu^2 \rangle_0^{1/2}$  calculated according to eq 18, 19, 26, 29, and 31 are represented by curve 1 which connects points calculated at integral values of  $x$ , the amide group moments  $m_i$  being included. Curve 2 represents  $\langle \mu^2 \rangle_0^{1/2}$  calculated with the  $m_i = 0$ . The dashed line labeled 3 connects values of  $\langle \mu^2 \rangle^{1/2}$  for  $x + 1 = 6$  and 7 calculated for the chain subject to perturbation by the Coulombic attractions between terminal charges, the effective dielectric constant  $\epsilon$  being equated to its value, 78.5, for water. This curve has been obtained by correcting curve 2 by numerical integration of eq 24 with  $a = 3 \text{ \AA}$ . Almost the same result is given by curve 4, which was obtained similarly using the approximate eq 25'. The approximation afforded by eq 25 is much less satisfactory, however; values thus calculated fall considerably below the experimental points.

The unperturbed moments calculated for the glycine dimer and trimer are in good agreement with the experimental values determined by Wyman and co-workers.<sup>18</sup> Those for the higher homologs fall below the curve for  $\langle \mu^2 \rangle_0^{1/2}$ , the departure increasing with chain length. The perturbation due to Coulombic attraction between terminal charges, as calculated by numerical integration of eq 24, gives a good account of this departure. We have not pursued this calculation to lower homologs because of limitations of the method cited above. It is to be observed, however, that the greater *diversity* of configurations available to the glycylic residue renders its polymers better suited to the perturbation treatment than alanine peptides of the same chain length  $x$ . Whereas the alanyl residue is restricted to two principal minima in its conformational energy map over the rotation angles  $\varphi$  and  $\psi$ , one of

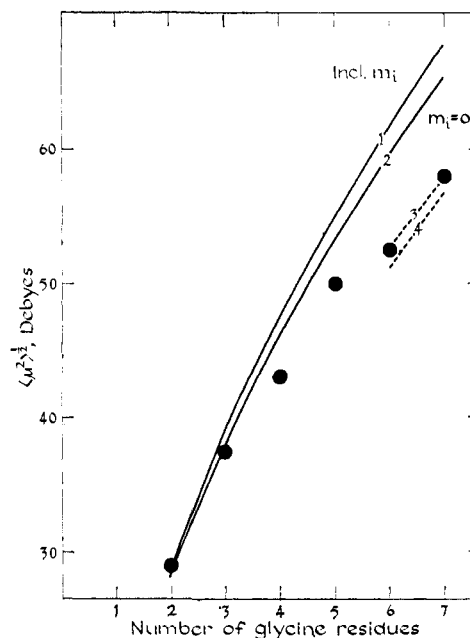


Figure 2. Root-mean-square dipole moments of glycine homologs determined in aqueous solutions<sup>18</sup> (points) and calculated as described in the text (curves). Unperturbed moments  $\langle \mu^2 \rangle_0^{1/2}$  calculated according to eq 18 are given by curve 1, the amide group moments  $m_i$  being included. Curve 2 represents  $\langle \mu^2 \rangle_0^{1/2}$  with  $m_i = 0$ . The dashed lines labeled 3 and 4 connect values of  $\langle \mu^2 \rangle^{1/2}$  for  $x + 1 = 6$  and 7 calculated for the chain subject to Coulombic attraction between terminal charges. Curve 3 corrects curve 2 by numerical integration of eq 24 with  $\epsilon = 78.5$  and  $a = 3 \text{ \AA}$ , and curve 4 corrects curve 2 using the approximate eq 25' with  $\epsilon = 78.5$ . (The dipole moment for  $x + 1 = 7$  was determined in a 5.1 *M* aqueous urea solution,<sup>18</sup> this solvent having a dielectric constant of 91.4.<sup>18</sup> Adoption of a value of  $\epsilon = 91.4$  raises the calculated point for  $x + 1 = 7$  on both curves 3 and 4 by <3%.)

these being favored over the other,<sup>2</sup> the symmetry of the glycylic residue renders  $-\varphi, -\psi$  equivalent to  $\varphi, \psi$ . Consequently, four minima occur.<sup>2</sup> The choices thus accessible to the glycylic residue permit a greater variation in  $r$ .