

BPC 01330

## Conformational properties of a polyglycine chain with secondary and tertiary structures of lysozyme

C. Dolcet and M.P. Tarazona

*Departamento de Química Física, Universidad de Alcalá de Henares, Madrid, Spain*

Received 25 March 1988

Revised manuscript received 3 August 1988

Accepted 18 October 1988

Dipole moment; Lysozyme; Mean squared optical anisotropy; Molar Kerr constant; Polyglycine; Tertiary structure; Unperturbed dimension

Unperturbed dimension  $\langle r^2 \rangle_0$ , dipole moment  $\langle \mu^2 \rangle$ , mean squared optical anisotropy  $\langle \gamma^2 \rangle$  and molar Kerr  $\langle {}_mK \rangle$  constant of a polyglycine chain with the secondary and tertiary structures of lysozyme have been calculated and the results compared with polyglycine chains with the same number of repeat units but different conformations including  $\alpha$ -helix,  $\beta$ -sheet or random coil. Thus, the influence of secondary and tertiary structures can be investigated. The results obtained show that for  $\langle r^2 \rangle$  and  $\langle \gamma^2 \rangle$  this influence is at least of the same order of magnitude as that of the primary structure, and is much greater for  $\langle \mu^2 \rangle$  and  $\langle {}_mK \rangle$ .

### 1. Introduction

Recently, we have studied [1–3] the influence of polypeptide secondary structure on four conformational properties, namely unperturbed dimensions  $\langle r^2 \rangle_0$ , dipole moments  $\langle \mu^2 \rangle$ , mean squared optical anisotropy  $\langle \gamma^2 \rangle$  and molar Kerr constant  $\langle {}_mK \rangle$ . The simplest polypeptide, polyglycine, was chosen as the model for the study, and two rigid structures,  $\alpha$ -helix [4] and  $\beta$ -sheet [5], commonly found in proteins, together with random coil were selected for the different repeat units of the chain. The results indicated a marked dependence of those magnitudes on chain secondary structure. Specifically the molar Kerr constant was the most strongly influenced by the secondary structure, which is not surprising taking into account its known large variation when the primary structure of the polypeptides varies [6].

We thought it would be interesting to calculate the value of the above four magnitudes in a pro-

tein with well defined secondary and tertiary structures and to study the influence of these structures on the conformational magnitudes. As a model for the secondary and tertiary structures we chose the lysozyme molecule since it is a relatively small protein, its structure is perfectly determined [7] and it has  $\alpha$ -helix and  $\beta$ -sheet segments along the chain. The hen egg white lysozyme is a small enzyme composed of 129 amino acid residues. Its form is roughly ellipsoidal measuring about  $45 \times 30 \times 30$  Å and has a marked cleft on one side. The structure of this protein was determined by X-ray analysis with the multiple isomorphous replacement method [7]. Lysozyme is a globular protein and its secondary structure is formed by  $\alpha$ -helix and  $\beta$ -sheet segments together with long stretches of a rather irregular conformation. The  $\alpha$  and  $\beta$  segments are segregated, both in space and along the polypeptide chain. In order to simplify the model, and as a first approximation, instead of the primary structure of lysozyme we have used a polyglycine chain with the same number of repeat units as lysozyme and with different secondary and tertiary structures. Thus, the values of the

Correspondence address: M.P. Tarazona, Departamento de Química Física, Universidad de Alcalá de Henares, Madrid, Spain.

conformational magnitudes of interest had been calculated for a polyglycine chain of 129 repeat units in four different cases: (1) with the actual secondary and tertiary structures of lysozyme, (2) in  $\alpha$ -helix conformation, (3) in  $\beta$ -sheet conformation and (4) in a random coil conformation. From the comparison of the results information of the influence of these structures can be obtained. The use of a polyglycine chain with the secondary and tertiary structures of lysozyme instead of the actual primary structure does not permit to obtain actual values of these magnitudes for the lysozyme. However, since the influence of primary structure of lysozyme in these properties had already been reported [6,8–10], comparison between the respective influences of the structures has been possible.

## 2. Methods

Values of the four conformational properties studied for a polyglycine chain of the same number of repeat units as lysozyme, in a well defined conformation, may be calculated by Flory's method of matrix multiplication [11,12] as follows.

The square of the end-to-end distance may be evaluated by serial multiplication of generator matrices

$$r^2 = \prod_{i=1}^x G_i \quad (1)$$

where  $G_i$  represents a  $5 \times 5$  generator matrix defined elsewhere [11,13].

The square of the dipole moment of the chain  $\mu^2$  may be evaluated by an expression similar to eq. 1 but with substitution of the dipole moment vector of unit  $i$   $\mu_i$  [9] for the length of the same unit in the generator matrices  $G_i$ .

Square optical anisotropy  $\gamma^2$  is defined as described [14–16]

$$\gamma^2 = 3/2 \text{ Trace } \langle \hat{\alpha} \hat{\alpha} \rangle \quad (2)$$

where  $\hat{\alpha}$  is the anisotropic part of the polarizability tensor  $\alpha$  of the molecule. Computation of  $\gamma^2$  for the chain, which requires prior knowledge of

the contributions  $\hat{\alpha}$  of each repeat unit [10] may be accomplished as

$$\gamma^2 = 3/2 \prod_{i=1}^x P_i \quad (3)$$

where  $P_i$  values are  $11 \times 11$  generator matrices defined elsewhere [11,13,17].

The last conformational magnitude studied, the Kerr constant of a given molecule, governs the birefringence produced by an electric field in a sample of independent and uncorrelated molecules and can be calculated as described [18,19]

$${}_m K = \frac{2\pi N}{15kT} \left[ \frac{\mu^T \hat{\alpha} \mu}{kT} + \frac{\epsilon - 1}{n^2 - 1} \text{Trace} \langle \hat{\alpha} \hat{\alpha} \rangle \right] \quad (4)$$

where  $N$  is Avogadro's number;  $k$  is the Boltzmann constant;  $T$  is the absolute temperature;  $\epsilon$  is the static dielectric constant of the medium;  $n$  its refractive index; and  $\mu$  is the permanent dipole moment. The term  $\text{Trace} \langle \hat{\alpha} \hat{\alpha} \rangle$  can be calculated using the  $P_i$  generator matrices of eq. 3. The term  $\mu^T \hat{\alpha} \mu$  can be calculated as

$$\mu^T \hat{\alpha} \mu = \prod_{i=1}^x Q_i \quad (5)$$

where  $Q_i$  values are the  $26 \times 26$  generator matrices defined elsewhere [11].

The values of the conformational magnitudes for the polyglycine chain were calculated assuming the four different cases described in the Introduction. In the first case the actual secondary and tertiary structures of lysozyme were taken into account by means of the dihedral angles  $\phi_i$  and  $\psi_i$  for each repeat unit given by Blake et al. [7]. These angles together with the geometrical parameters of table 1 permit the corresponding transformation matrices to be obtained. In the second case, the calculations were carried out for a polyglycine chain in an  $\alpha$ -helix conformation, and values of  $124.8^\circ$  and  $134.8^\circ$  for rotations  $\phi_i$  and  $\psi_i$  respectively [1] were used to generate the helix. Third, the  $\beta$ -sheet was assumed to be the conformation of the chain, thus angles of  $\phi_i = 65.05^\circ$  and  $\psi_i = 289.1^\circ$  were used to generate a parallel  $\beta$ -sheet [2]. In the fourth and last case, instead of a fixed conformation as in the above cases, a random coil conformation was supposed; therefore, eqs. 1, 3

Table 1

## Geometrical parameters

Type of bond	Length (nm)	Type of angle	Angle <sup>a</sup>
N-C <sup>α</sup>	0.145 ± 0.002	C <sup>α</sup> CN	115°
C <sup>α</sup> -C	0.153 ± 0.001	OCN	124°
C-N	0.1325 ± 0.0005	C <sup>α</sup> CO	120°
C-O	0.123 ± 0.001	CNC <sup>α</sup>	121°
N-H	0.100 ± 0.004	CNH	124°
C <sup>α</sup> -H	0.100 ± 0.001	HNC <sup>α</sup>	115°
		NC <sup>α</sup> C	110°
		η	22°
		ξ	13°

<sup>a</sup> Standard deviation equal to ± 3 for all angles.

and 5 had to be averaged over all the conformations of the chain. For polyglycine these averages may be carried out for each single repeat unit [20], and they can therefore be obtained by replacing the generator  $G_i$ ,  $P_i$  and  $Q_i$  with their averages  $\langle G_i \rangle$ ,  $\langle P_i \rangle$ ,  $\langle Q_i \rangle$  over skeletal bonds  $\phi_i$  and  $\psi_i$  (fig. 1).

As is customary, the absolute values of the magnitudes for the whole chain were transformed into values the repeat unit defined as  $C_x = \langle r^2 \rangle_0 / x l^2$ ,  $D_x = \langle \mu^2 \rangle / x \mu_0^2$ ,  $G_x = \langle \gamma^2 \rangle / x$  and  $K_x = \langle {}_m K \rangle / x$  where  $x$  is the polymerization degree,  $l$  the length of virtual bond and  $\mu_0$  the dipole moment of the repeat unit.  $C_x$  and  $D_x$  are dimensionless while  $G_x$  and  $K_x$  are given in the units  $10^{-6} \text{ nm}^6$  and  $10^{-27} \text{ m}^5 \text{ V}^{-2} \text{ mol}^{-1}$ , respectively. In these expressions the brackets meaning aver-

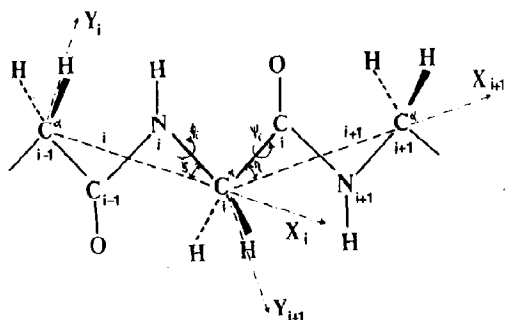


Fig. 1. A segment of a polyglycine chain in its planar all *trans* conformation. Virtual bonds connecting consecutive  $\alpha$ -carbons are shown by dashed lines. Coordinate systems are represented by dot-dashed lines.  $z$  axes complete right handed frames.

ages only have relevance for the random coil conformation; for the other three conformations that are well defined there is only a unique value, so the brackets are superfluous.

### 3. Results and discussion

Figs. 2–5 represent the calculated values of the four magnitudes  $C_x$ ,  $D_x$ ,  $G_x$  and  $K_x$  respectively versus the polymerization degree  $x$  for a poly-

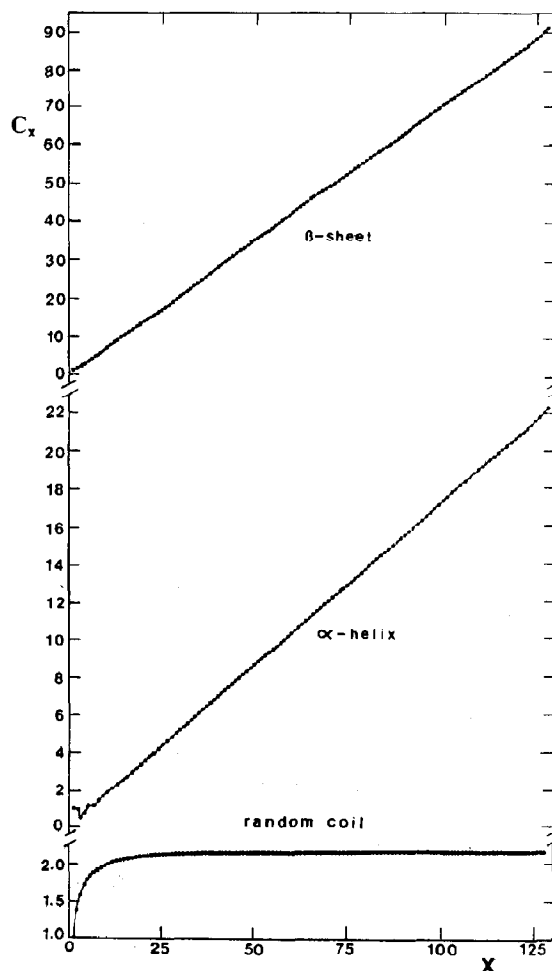
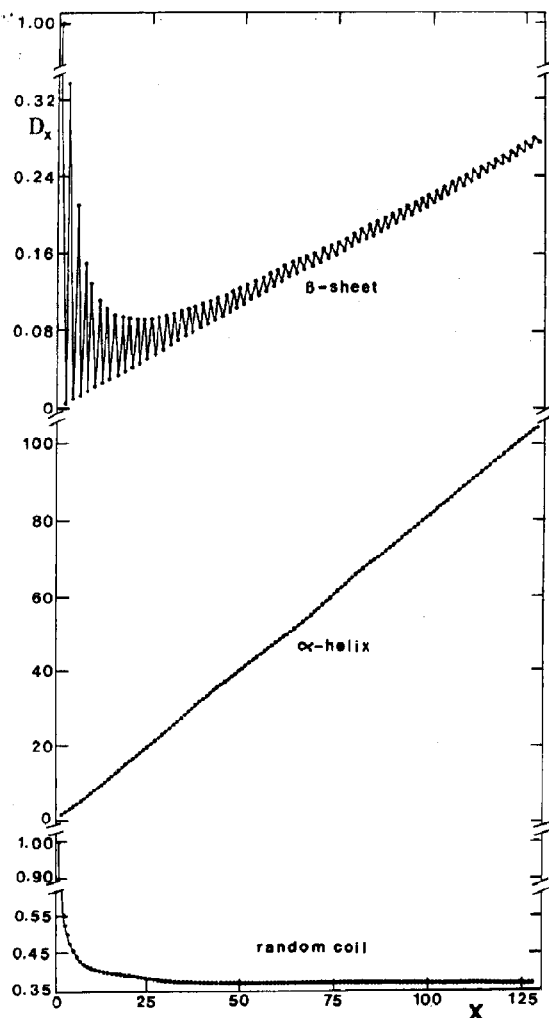
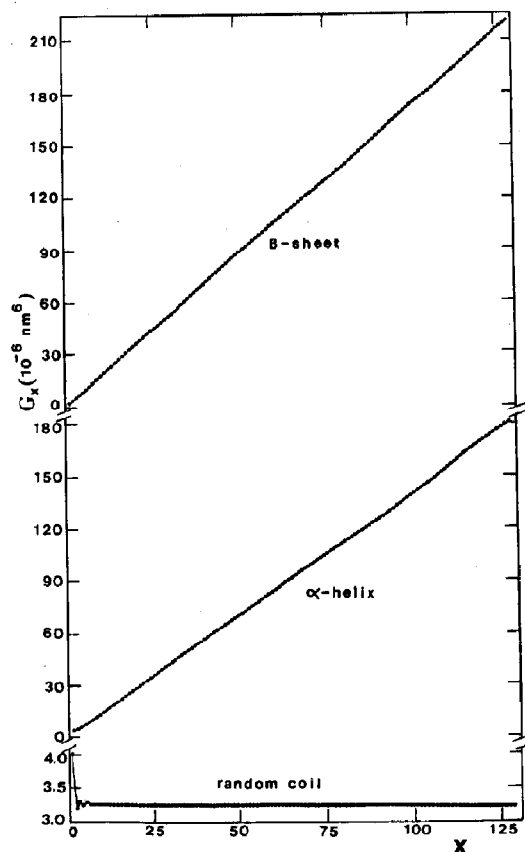


Fig. 2. Plot of  $C_x$  versus polymerization degree  $x$  for a polyglycine chain with the same number of repeat units of lysozyme in the  $\beta$ -sheet,  $\alpha$ -helix and random coil conformations. It is important to notice the different scales used in the vertical axis.

Fig. 3. Plot of  $D_x$  versus  $x$  (see legend of fig. 2).

glycine chain with the same number of residues as lysozyme and with all their units in one of the studied  $\beta$ -sheet,  $\alpha$ -helix or random coil conformations. It is important to notice the different scales used in the vertical axis and the great variation for  $K_x$  which make us represent its natural logarithm as  $\ln K_x$ . For chains with the  $\alpha$ -helix and  $\beta$ -sheet rigid conformations, the four magnitudes increase [1-3] with increasing  $x$  whereas for the random coil conformation the asymptotic limit is reached at low  $x$  values [6,8-10]. For the  $\beta$ -sheet conformation the values of  $D_x$  (fig. 3) and  $K_x$  (fig. 5)

oscillate strongly from one polymerization degree to the next one for low  $x$  values although the oscillations diminish at higher  $x$  values. In the case of the  $D_x$  the explanation is straightforward. If one chooses a coordinate system with the  $X$  axis parallel to the propagation direction of the sheet (i.e., joining the odd  $C^\alpha$  of the chain), the components of the dipole moment vector corresponding to a repetitive unit of the chain placed in an odd position are  $\mu_0 = (0.045 \mu_0, 0.029 \mu_0, -0.99 \mu_0)$  and  $\mu_0 = (0.045 \mu_0, -0.029 \mu_0, 0.99 \mu_0)$  for an even repetitive unit. Thus, for an even number of repeat units the components  $\mu_y$  and  $\mu_z$  cancel each other out and only  $\mu_x$  will contribute to the total dipole moment. Whereas for an odd number of repeat units besides the  $\mu_x$  components, the  $\mu_y$  and  $\mu_z$  components for one repeat unit will contribute to the total value of  $\mu$ . When

Fig. 4. Plot of  $G_x$  versus  $x$  (see legend of fig. 2).

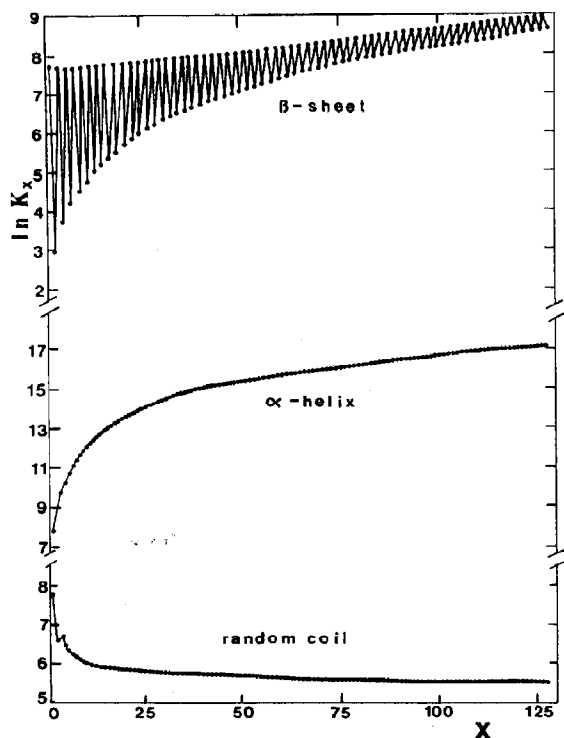


Fig. 5. Plot of  $\ln K_x$  versus  $x$  (see legend of fig. 2).

$x$  increases the component  $\mu_x$  grows bigger and the oscillations diminish. The values obtained for  $K_x$  are similar to those for  $D_x$  due to the importance of the term  $\mu^T \alpha \mu$  in the values calculated for this magnitude.

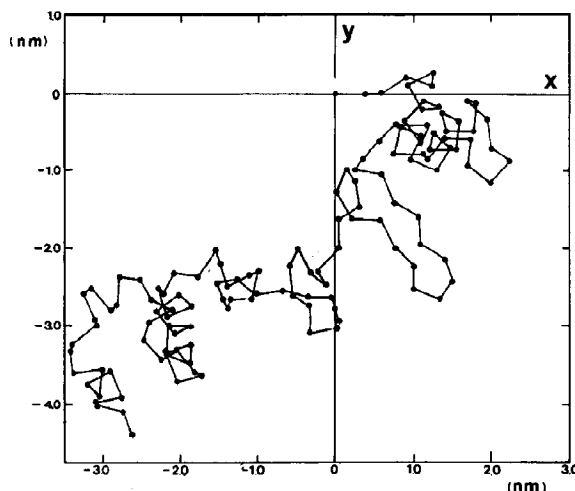


Fig. 6. Projection upon the plane  $xy$  of the coordinates of the successive  $\alpha$ -carbons of a polyglycine chain with the secondary and tertiary structure of lysozyme. The coordinate system corresponds to the first virtual bond of the chain.

The next step in the calculation has been to study a polyglycine chain with the lysozyme secondary and tertiary structure. The geometrical form of the chain can be represented by the end-to-end vector  $r$  for each one of their repeat units, i.e., the coordinates of consecutive  $\alpha$ -carbons in the coordinate system of the first virtual bond. Figs. 6 and 7 represent the projections  $xy$  and  $xz$  respectively of the vector for the successive  $\alpha$ -carbons. These figures show the irregular form of

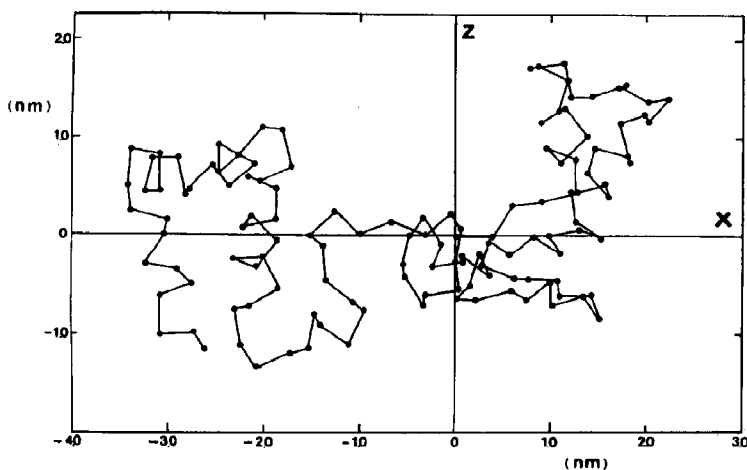


Fig. 7. Projection upon the plane  $xz$  (see legend of fig. 6).

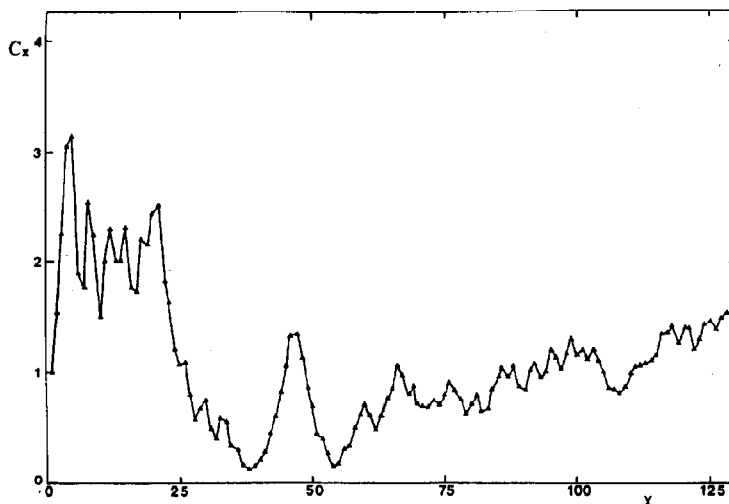


Fig. 8. Plot of  $C_x$  versus polymerization degree  $x$  for a polyglycine chain with the secondary and tertiary structures of lysozyme.

the chain, which can fit in a parallelepipedon of an approximate volume of  $80 \text{ nm}^3$  whereas if the whole chain had the most extended conformation,  $\beta$ -sheet, its length would be  $41.5 \text{ nm}$ .

The results obtained for the four conformational properties studied,  $C_x$ ,  $D_x$ ,  $G_x$ , and  $K_x$ , are plotted versus the polymerization degrees in figs. 8–11 respectively. Although the chain has a fixed and well defined conformation, its propagation direction varies frequently, and so the values for the magnitudes oscillate between broad limits.

The calculated values for the complete chain ( $x = 128$ ) are presented in the 3rd row in table 2. In this table we have included the results obtained if all the units of the chain have the same conformation, random coil (fourth row),  $\alpha$ -helix (fifth row) or  $\beta$ -sheet (sixth row). The first two rows also include the results obtained by Saiz et al. [6,8–10] for the influence of the lysozyme primary structure in these magnitudes. The first row shows the values of the considered properties for the lysozyme molecule with its actual primary structure

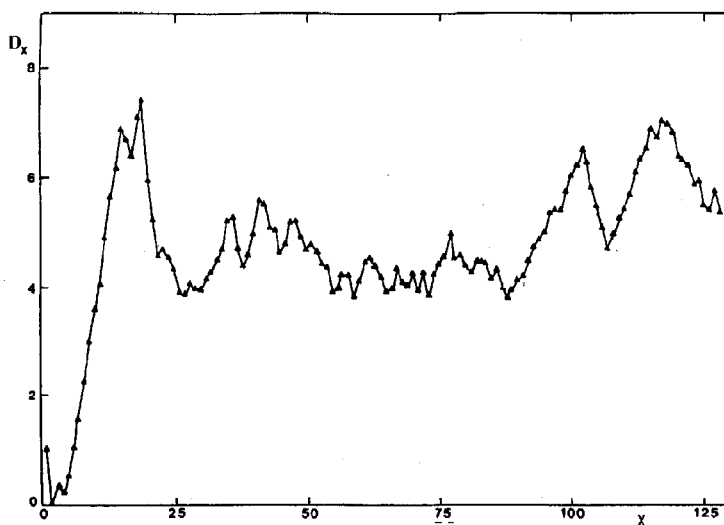
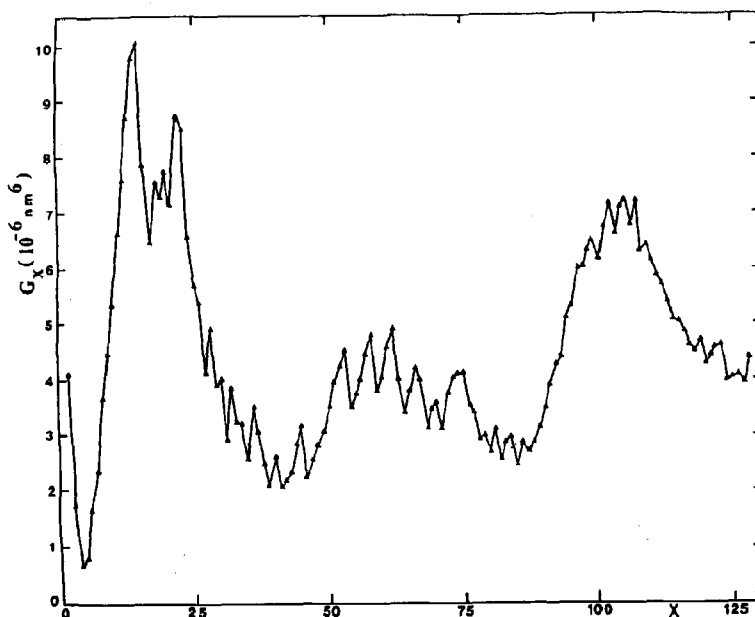


Fig. 9. Plot of  $D_x$  versus  $x$  (see legend of fig. 8).

Fig. 10. Plot of  $G_x$  versus  $x$  (see legend of fig. 8).

but in random coil conformation. The second row shows the results obtained if the different amino acid residues were ordered randomly in the chain. As can be easily seen from even a cursory inspection of the first two rows, only the Kerr constant is sensitive to this change in the primary structure.

Comparison of the values obtained for polyglycine chains (third and sixth rows) indicates that changing from the lysozyme secondary and tertiary structure to a random coil conformation only slightly affects  $C_x$  and  $G_x$  but does strongly modify the dipole moment and the Kerr constant.

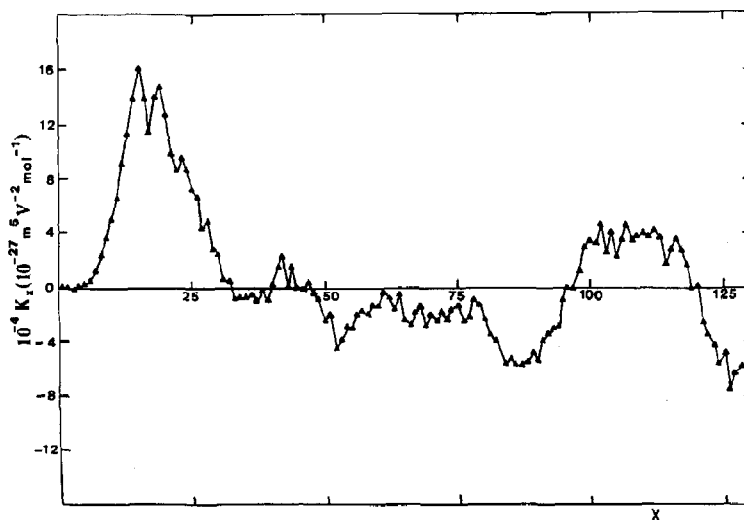
Fig. 11. Plot of  $K_x$  versus  $x$  (see legend of fig. 8).

Table 2

Values of the conformational magnitudes calculated for the different simplified models of lysozyme molecule

Chemical composition	Amino acid sequence	Secondary and tertiary structure	$C_x$	$\langle \mu^2 \rangle / x^a$ ( $D^2$ )	$G_x$ ( $10^{-6} \text{ nm}^6$ )	$K_x$ ( $10^{-27} \text{ m}^5 \text{ V}^{-2} \text{ mol}^{-1}$ )
Lysozyme	Actual	Random coil	5.5	9.0	9.2	$-3.41 \times 10^2$
Lysozyme	Random	Random coil	5.9	8.8	9.6	80
Polyglycine	—	Lysozyme	1.5	92.6	4.4	$-5.85 \times 10^4$
Polyglycine	—	Random coil	2.2	6.3	3.3	$2.71 \times 10^2$
Polyglycine	—	$\alpha$ -helix	22.2	$1.79 \times 10^3$	179.2	$3.00 \times 10^7$
Polyglycine	—	$\beta$ -sheet	91.3	4.7	220.5	$5.46 \times 10^3$

<sup>a</sup> We use the ratio  $\langle \mu^2 \rangle / x$  instead of the dipole ratio  $D_x = \langle \mu^2 \rangle / x \mu_0^2$  in order to facilitate the comparison between polyglycine chains and those chains with the actual chemical structure of lysozyme since in the last case the contributions  $\mu_0$  of the residues are different. (The brackets meaning averages only apply to random coil conformations.)

However, the four magnitudes vary drastically if the three fixed structures (lysozyme,  $\alpha$ -helix or  $\beta$ -sheet) are considered.

It is worthwhile to point out the extraordinary sensitivity of the Kerr constant to any change in the secondary structure. Thus their values oscillate in a range of  $10^{11}$  whereas the range for the following magnitude, the dipole moment, is  $10^3$ , which although great is not comparable to that of the Kerr constant. The Kerr constant is then a magnitude which can be very useful to study the variations in the structure of polypeptides. Apart from the Kerr constant, the other three magnitudes, especially dimensions and dipole moment, can complement each other in the study of the structure since, as can be easily seen in table 2, their values increase or decrease with different form with distinct structures.

Comparison between the first and fourth rows indicates that the kind of amino acid residues that form the chain (actual residues or polyglycine residues) is not as significant as one might think given the enormous change in the chemical composition of the chain that takes place when the chemical structure of lysozyme is substituted by that of a polyglycine chain. Thus the differences between the values for the four magnitudes for a lysozyme molecule and for a polyglycine chain, both in random coil conformation (first and fourth rows respectively) are of the same order of magnitude or even smaller, for dipole moment and Kerr constant, than the differences between two polyglycine chains of which one is in random coil

conformation and the other has the secondary and tertiary structure of lysozyme (fourth and third rows respectively). Therefore this work indicates the strong influence that the secondary and tertiary structures of lysozyme have in the calculated values of the four conformational magnitudes studied.

### Acknowledgment

We are grateful to Ms. Carol F. Warren (I.C.E. of U.A.H.) for her editorial help.

### References

- 1 C. Dolcet, M.P. Tarazona and E. Saiz, *An. Quim.* 80 (1984) 476.
- 2 C. Dolcet and M.P. Tarazona, *Int. J. Peptide Protein Res.* 30 (1987) 548.
- 3 C. Dolcet and M.P. Tarazona, *Biophys. Chem.* 29 (1988) 253.
- 4 L. Pauling, R.B. Corey and H.R. Branson, *Proc. Natl. Acad. Sci. U.S.A.* 37 (1951) 205.
- 5 L. Pauling and R.B. Corey, *Proc. Natl. Acad. Sci. U.S.A.* 39 (1953) 253.
- 6 A. Lopez-Piñero, M.P. Tarazona and E. Saiz, *Polym. Bull.* 10 (1983) 373.
- 7 C.C.F. Blake, G.A. Mair, A.C.T. North, D.C. Phillips and V.R. Sarma, *Proc. R. Soc. (London)* B167 (1967) 365.
- 8 A. Lopez-Piñero, F. Mendicuti and E. Saiz, *An. Quim.* 72 (1981) 323.
- 9 A. Lopez-Piñero and E. Saiz, *Int. J. Biol. Macromol.* 5 (1983) 37.
- 10 A. Lopez-Piñero, M.P. Tarazona and E. Saiz, *J. Chim. Phys.* 80 (1983) 529.



- 11 P.J. Flory, *Statistical mechanics of chain molecules* (Interscience, New York, 1969).
- 12 P.J. Flory, *Macromolecules* 7 (1974) 381.
- 13 P.J. Flory and Y. Abe, *J. Chem. Phys.* 54 (1971) 1351.
- 14 R.T. Ingwall and P.J. Flory, *Biopolymers* 11 (1972) 1527.
- 15 G.D. Patterson and P.J. Flory, *J. Chem. Soc. Faraday Trans. II* 68 (1972) 1098; G.D. Patterson and P.J. Flory, *J. Chem. Soc. Faraday Trans. II* 68 (1972) 1111.
- 16 I.L. Fabelinskii, *Molecular scattering by light* (Plenum Press, New York, 1968).
- 17 A.E. Tonelli, Y. Abe and P.J. Flory, *Macromolecules* 3 (1970) 303.
- 18 A.D. Buckingham and J.A. Pople, *Phys. Soc.* 68A (1955) 905.
- 19 U.W. Suter and P.J. Flory, *J. Chem. Soc. Faraday Trans. II* 73 (1977) 1521.
- 20 R.T. Ingwall, E.A. Czurylo and P.J. Flory, *Biopolymers* 12 (1973) 1137.