

Assigning the 625 cm⁻¹ and other high frequency Raman bands of the Z₁ form of DNA using a normal modes analysis and computer graphics

G Vergoten*, P Lagant[†], W L Peticolas[‡], Y Moschetto[§], I Morize[¶], M C Vaney[¶] and J P Mornon[¶]

*Laboratoire de Génie Biologique et Médical, INSERM U279, Faculté de Pharmacie, Lille, France

[†]Institut de Recherches sur le Cancer, INSERM U124, Lille, France

[‡]Department of Chemistry, University of Oregon, Eugene, Oregon, USA

[§]Centre de Technologie Biomédicale, INSERM U279, Lille, France

[¶]Laboratoire de Minéralogie Cristallographie, Universités de Paris 6 et 7, Paris, France

The left-handed zigzag conformation of the ribose phosphate backbone of the double helix of Z-DNA is well established. A normal modes analysis has been performed using the crystal form of d(CpGpCpGpCpG); this is a half turn of double helical Z-DNA. The calculated 625 cm⁻¹ band, which has been shown to be specific for the Z conformation of DNA is discussed with reference to previous work. It would seem that the classical description of vibrational modes in terms of internal coordinates is inadequate for macromolecular systems. The great number of internal coordinates involved in one vibrational mode makes the assignment complex. A new method of interpreting the vibrational modes is described. It is based on the use of Cartesian displacement coordinates to describe the vibrational mode together with computer graphics for its direct visualization.

Keywords: graphic displays, Z₁-DNA, Raman bands, normal modes analysis, computer graphics

Received 21 April 1986

Accepted 13 May 1986

The hexanucleotide d(CpGpCpGpCpG) has been shown to adopt the Z conformation with 12 base pairs per turn of the left-handed double helix. The antiparallel strands have their deoxyguanosine residues in the C_{3'} endo Syn conformation while the deoxycytidine residues are in the C_{2'} endo Anti conformation¹. Wang *et al.*² have generated an idealized hexamer with a 6₅ axis and a 44.58 Å repeat (space group D_{6h}³, one molecule per unit cell). Atomic coordinates given in these two papers are used in the present calculation for the Z₁ form of DNA.

NORMAL MODES ANALYSIS

The asymmetric unit comprises 66 atoms (one CpG subunit and one water molecule) (see Figure 1). Six such asymmetric units are found in one unit cell of the crystal. Phase relations between the motion of the six units in the unit cell leading to the A₁, A₂, B₁, B₂, E₁ and E₂ symmetry species were obtained from the D_{6h} hexagonal symmetry group.

Optically active vibrations in the crystal were obtained according to the method of Shimanouchi *et al.*³. 1185 vibrational frequencies were thus calculated, among them three are of the rotatory lattice type. The remaining 1182 frequencies are assigned to internal modes of vibrations. The internal vibrations may be separated into local vibrations (relatively small number of atoms involved, high frequency, small amplitude) and overall vibrations (all atoms of the molecule concerned with the motion, low frequency, large amplitude).

The separation between these two types of internal vibrations may be fixed at 210 cm⁻¹ which corresponds to the value of the thermal energy at room temperature. The overall low frequency vibrations are directly related to the biological functions and properties⁴⁻⁶.

A general valence force-field was used. Redundancy symmetry coordinates were analytically determined for the ring molecules (deoxyribose, cytosine, guanine) and the pseudotetrahedral methylene and phosphate groups. They were then removed and the corresponding internal constraints were included in the force-field. The intramolecular force-field was constructed with starting values of the force constants given by several authors^{7,8}. In order to check the validity of the intramolecular force-field used, the method described by Peticolas *et al.* was employed for the G and C bases^{9,10}. It consists of using the force-field for calculating the resonance Raman intensities and comparing them with the observed ones.

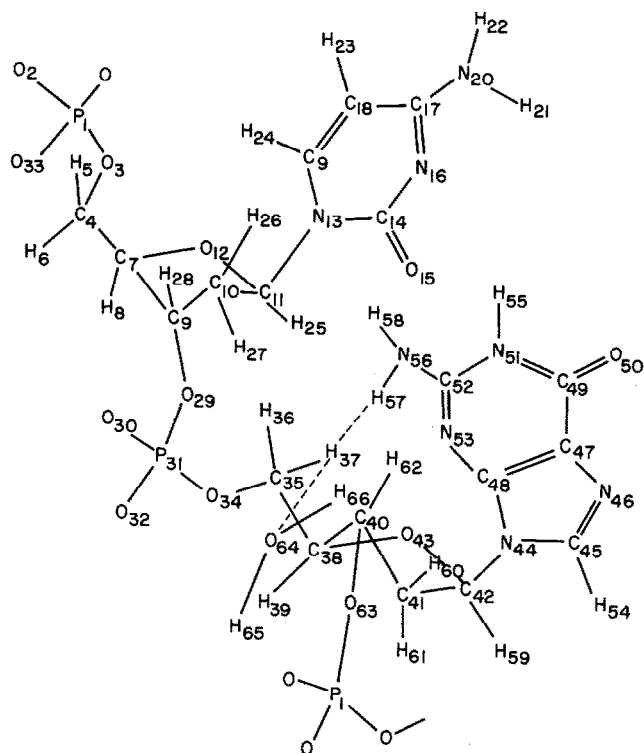


Figure 1. 66 atom model used to generate half a turn of the Z conformation of DNA

The intermolecular force-field is composed of several types of interactions. Nonbonded atom-atom interactions are expressed in terms of a Buckingham's type function:

$$V(r) = A \exp(-Br) - Cr^{-6}$$

where r is the distance between the nonbonded atoms under consideration; the parameters A , B and C are characteristic of one type of atom-atom contact. There are 913 contacts that have been defined in one unit cell. The hydrogen bonds of the base pairs were described with a Urey-Bradley potential function. Dipole-dipole interactions for the C=O and the phosphate groups were also added. A total of 285 force constants were thus defined and used to refine the original general valence force field. The optically active normal vibrations were determined using the CVOA computer program¹¹ which was run on the IBM 4381 group 2 computer (of the Ecole Polytechnique, Montréal, Canada). The calculation was carried out using 18 h CPU time and 7.8 Mbyte core memory. In this paper the Raman 625 cm⁻¹ mode of vibration which is characteristic of the Z conformation will be described. The complete results will be given in a subsequent paper.

625 cm⁻¹ AND OTHER RAMAN BANDS OF Z DNA

According to previous work, the 625 cm⁻¹ band observed in the Raman spectrum of the Z hexamer has been found to be highly sensitive to the Z conformation^{12,13}. The corresponding frequency obtained in the present calculation is 617.5 cm⁻¹; this corresponds to a normal mode of vibration highly localized at the guanosine residue shown in Figure 1. The potential energy distribution expressed in terms of the internal coordinates is complex. It is made up of the following main contributions:

- 12% (C₄₀-C₃₈-O₄₃) bending
- 11% (O₆₃-C₄₀-C₄₁) bending

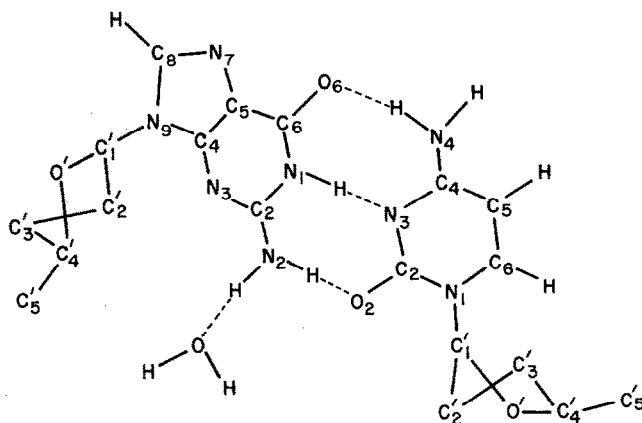


Figure 2. Usual numbering of the GC base pairing together with the corresponding deoxyribose moieties

- 7% (N₄₄-C₄₅-N₄₆) bending
- 7% (C₄₁-C₄₂=N₄₄) bending
- 7% (C₄₁-C₄₂) stretching
- 5% (N₄₄-C₄₂-O₄₃) bending
- 4% (N₄₄-C₄₂) in-plane bending

This description corresponds quite well to the prediction of the participation of glycosidic bond vibrations, presented by Nishimura *et al.*¹³ with an appreciable vibrational coupling between the guanine and ribose.

Figure 2 gives the usual numbering of atoms involved in the base pairing and shows also the C_{3'} endo and C_{2'} endo conformation of the sugar moiety of the deoxyguanosine and deoxycytidine respectively.

Colour Plate 1 shows the extreme atomic displacements during the vibration calculated at 617.5 cm⁻¹. It can be clearly seen that the corresponding motion only concerns the deoxyguanosine moiety. The deoxyribose ring undergoes a bending motion and the guanine residue is subject to a large amount of in-plane bending of the N₁H.

In addition to this band, vibrational modes at 697 cm⁻¹ (30.5% N₅₁-C₅₂=N₅₃ bending, 12% C₅₂-N₅₁ stretching, 17% C₄₉=O₅₀ in-plane bending and 9.5% N₄₄-C₄₈ stretching), 684 cm⁻¹ (71.5% N₅₁-H₅₅ out-of-plane bending and 14% C₅₂-N₅₆ out-of-plane bending) and 664 cm⁻¹ (28% C₄₂-N₄₄ stretching, 12% C₄₁-C₄₂-N₄₄ bending, 10% N₄₄-C₄₅=N₄₆ bending and 9.6% O₄₃-C₄₂ stretching) are assigned to the guanine for the first two modes and to the guanosine for the latter one. According to the work of Thaman¹⁰ a band is observed at 675 cm⁻¹ in the Raman spectrum of the Z hexamer. It must be noted that the calculations carried out by the authors also give two predicted bands, one at 674 cm⁻¹ (19% C₁₀-C₁₁ stretching, 13.6% C₁₀-C₁₁-N₁₃ bending, 10% C₁₈-H₂₃ out-of-plane bending and 9% C₇-O₁₂ stretching) and the other at 668 cm⁻¹ (26.7% C₁₁-O₁₂ stretching, 9% C₁₀-C₁₁-N₁₃ bending and 9% C₁₈-H₂₃ out-of-plane bending) due to the cytidine moiety (see Figure 1).

A guanine band has also been observed at 1320 cm⁻¹ for the high salt form of poly (dGdC).poly (dGdC). A band has been calculated by the authors at 1320 cm⁻¹ which is described as 41.5% C₄₁-H₆₀-H₆₁ scissor; 10% N₅₆-H₅₇-H₅₈ rocking; 9% C₄₅-H₅₄ in-plane bending and 7% C₅₂-N₅₆ in-plane bending. Another band is calculated at 1327 cm⁻¹ which can also be assigned to the guanine moiety (35% C₄₁-H₆₀-H₆₁ scissor, 20.5% C₄₅-H₅₄ in-plane bending, 9.8% N₅₆-H₅₇-H₅₈ rocking and 8% C₅₂-N₅₆ in-plane bending).

It is clear that two further types of experiments and calculations are needed in order to confirm some of the earlier results. First, it seems very important to observe and calculate the resonance Raman intensities^{9,10}. Second, specific isotopic substitutions (deuteration) would be of great help in assigning the observed Raman band and then refining the calculation.

It is also very clear that, in the case of such complicated molecular systems, the description of modes in terms of the potential energy distribution among the internal coordinates does not give a comprehensive view of the corresponding vibrational motion. In order to overcome this difficulty, a new technique is proposed here which involves the direct visualization of the motion with the help of computer graphics.

VISUALIZATION OF MOLECULAR VIBRATIONS USING COMPUTER GRAPHICS

The dimensions of the Cartesian coordinate column vector **X** and the internal coordinate column vector **R** of the asymmetric unit are 198 and 244 respectively. The normal coordinates column vector **Q** of **A**₁ symmetry is 202.

The following matrix relations exist between these different vectors:

$$\mathbf{R} = \mathbf{B}\mathbf{X} = \mathbf{L}\mathbf{Q}$$

where **B** is a rectangular matrix since the size of **R** is greater than that of **X**.

The inverse transformation **X** = **AR** is thus needed with **BA** = **E** where **E** is the identity matrix.

Under these conditions the inverse kinetic energy matrix **G** is defined as:

$$\mathbf{B}\mathbf{M}^{-1}\mathbf{\tilde{B}} = \mathbf{G} \quad (1)$$

where **M**⁻¹ is a diagonal matrix of the order of 198 whose elements are the inverse of the atomic masses and **B**[~] the transposed matrix of **B**.

Multiplying equation (1) by **G**⁻¹ leads to:

$$\mathbf{B}\mathbf{M}^{-1}\mathbf{\tilde{B}}\mathbf{G}^{-1} = \mathbf{G}\mathbf{G}^{-1} = \mathbf{E} = \mathbf{B}\mathbf{A} \quad (2)$$

From equation (1):

$$\mathbf{A} = \mathbf{M}^{-1}\mathbf{\tilde{B}}\mathbf{G}^{-1} \quad (3)$$

Since **X** = **L**_x**Q** = **ALQ** = (**M**⁻¹ **B****G**⁻¹**L**)**Q**, the **L**_x matrix is thus defined:

$$\mathbf{L}_x = \mathbf{M}^{-1}\mathbf{\tilde{B}}\mathbf{G}^{-1}\mathbf{L} \quad (4)$$

If the secular equation **GFL** = **LA** is solved, a new definition of the **L**_x matrix can be obtained which is expressed in terms of the force constants matrix **F**:

$$\mathbf{L}_x = \mathbf{M}^{-1}\mathbf{\tilde{B}}\mathbf{F}\mathbf{L}\mathbf{A}^{-1} \quad (5)$$

The motion of atom *i* during the normal vibration *a* is given in terms of time dependent Cartesian coordinates *x*_i^a, *y*_i^a, *z*_i^a whose expressions are:

$$\begin{aligned} x_i^a &= x_i^0 + D(\mathbf{L}_x)_{xi}^a \sin(2\pi c\bar{\nu}_a t) \\ y_i^a &= y_i^0 + D(\mathbf{L}_x)_{yi}^a \sin(2\pi c\bar{\nu}_a t) \\ z_i^a &= z_i^0 + D(\mathbf{L}_x)_{zi}^a \sin(2\pi c\bar{\nu}_a t) \end{aligned}$$

where *x*_i⁰, *y*_i⁰, *z*_i⁰ are the Cartesian coordinates of atom *i* at their equilibrium position; $\bar{\nu}_a$ is the wavenumber of the normal vibration *a*; (**L**_x)_{xi}^a (**L**_x)_{yi}^a (**L**_x)_{zi}^a are the elements

of the **L**_x matrix previously defined; $D = \frac{6.8465\sqrt{T}}{\bar{\nu}_a}$ is

a coefficient used to make all physical quantities dimensionally homogenous, *T* being the absolute temperature.

The equations derived above can be obtained if the total energy of the system is assumed to be *kT*, the Angstrom is the unit of length, and the amu is the mass unit.

The motion is obtained by giving the quantity 2π*C* $\bar{\nu}$ to the values -90°, -80°... +80°, +90°.

A 16 mm film was prepared using the Manosk software and a PS300 Evans and Sutherland graphic system linked to a DPS6 CII Honeywell Bull computer¹⁴. Three high frequency (815, 787 and 612.5 cm⁻¹) and three low frequency (2.1, 2.7 and 3.8 cm⁻¹) bands were visualized. It is clear that further results are needed in order to obtain a better force-field to describe the lowest frequencies which are directly related to the conformational adaptability of the molecule.

CONCLUSION

For the case of complex biological molecules (nucleic acids, proteins, membranes) the description of the vibrational modes in terms of internal (or symmetry) coordinates is meaningless since the corresponding modes appear to be coupled to a great extent.

A method has been proposed to avoid this difficulty which involves the use of normal modes analysis and computer graphics. The technique is particularly useful at low frequencies where all the atoms are involved. For the case of the regular left-handed double helix of Z-DNA, the use of computer graphics makes possible the distinction, among the overall vibrations, between lateral 'breathing' bending and twisting (transverse acoustic vibrations) and longitudinal 'breathing' (longitudinal acoustic vibrations).

REFERENCES

- 1 Wang, A H J et al. *Nature* Vol 282 (1979) pp 680-686
- 2 Wang, A H J et al. *Science* Vol 211 (1981) pp 171-176
- 3 Shimanouchi, T et al. *J. Chem. Phys.* Vol 35(5) (1961) pp 1597-1612
- 4 Remington, S et al. *J. Mol. Biol.* Vol 158 (1982) pp 11-152
- 5 Huber, R and Bennett, Jr, W S *Biopolymers* Vol 22 (1983) pp 261-279
- 6 Lesk, A M and Chothia, C J. *Mol. Biol.* Vol 174 (1984) pp 175-191
- 7 Baret, J F et al. *J. Raman Spectros.* Vol 8(6) (1979) pp 291-304
- 8 Majoube, M *Biopolymers* Vol 24(6) (1985) pp 1075-1087
- 9 Peticolas, W L et al. *J. Chem. Phys.* Vol 73(9) (1981) pp 4185-4191
- 10 Peticolas, W L *Ber. Bunsenges Phys. Chem.* Vol 85 (1981) pp 481-484
- 11 Takeuchi, H *PhD Thesis* University of Tokyo (1975)
- 12 Thaman, T J et al. *Nucleic Acids Res.* Vol 9 (20) (1981) pp 5443-5447
- 13 Nishimura, Y et al. *Nucleic Acids Res.* Vol 11(5) (1983) pp 1579-1588
- 14 Vaney, M C et al. 'Manosk, a graphics program designed for macromolecular modelling' *J. Mol. Graph.* Vol 3 No 3 (September 1986) Abstract No 23 p 123