SUBCUR: Visualization of structural differences between DNA duplexes

Sanjay R. Sanghani,* Adrian H. Elcock, and Ian S. Haworth

Physical Chemistry Laboratory, Oxford University, South Parks Road, Oxford, UK

A computer program, SUBCUR, is described which permits analysis and rapid identification of geometrical differences and patterns of variance between two DNA duplexes. The program is compatible with the CURVES 3.1 package[1] and allows graphical visualization of the structural differences. Examples are provided which illustrate the applicability of the program in analyzing the different backbone conformations of two helices and the different curvatures of two helices.

Keywords: DNA, structure, visualization

INTRODUCTION

Small changes in the local structure of a DNA helix can have a significant effect upon ligand binding. For example, protein recognition of a particular DNA sequence will be influenced by changes in the helix fine structure. Many nucleases are thought to recognize patterns of charge density on the phosphate backbone, and this has stimulated the design of structurally modified oligonucleotides for use in the antisense and triplex anti-tumor strategies.[2] In addition, changes in sugar pucker have been cited as playing a role in the sequence selectivity of smaller ligands.[3] There is also interest in the micro heterogeneity of nucleic acid structure as a function of base sequence since certain sequences lead to natural curvature which may have an important role in the *in vivo* activity of DNA.

Torsional angle changes in the sugar phosphate backbone produce multiple conformations of B-, A-, and Z-DNA. For B-DNA, two distinct subclasses, B_I and B_{II} have been recognized. They differ in the rotation of two backbone torsion angles; B_I-DNA has ϵ and ζ (Figure 1) orientations of *trans* and *gauche* respectively whereas in B_{II}-DNA these are *gauche*$^-$ and *trans* respectively. The reversed conformation of ϵ and ζ causes a significant change in orientation of the phosphate with respect to the helix axis, which can have significant influence on ligand binding. The identification of such local structural features is of importance to the understanding of the DNA interaction with proteins and antitumor agents.

Computer modeling and molecular dynamics can provide invaluable information in this area. These approaches are often comparative in nature. For example, in an energy minimization, the behavior of two related DNA-binding ligands might be described in terms of their different binding energetics and differences in their induced structural changes to the DNA helix. Alternatively, in a molecular dynamics simulation, the time-dependent structural fluctuation of the helix might be of interest. However, as polynucleotides are systems of enormous structural complexity, the efficient comparison of two DNA duplexes is difficult and time consuming.

PROGRAM DESCRIPTION

SUBCUR facilitates rapid comparison of two DNA helices and visualization of the structural differences. The program is written in FORTRAN 77 and runs on VAX and UNIX platforms. The data input to SUBCUR are two outputs from the CURVES 3.1 program.[1] SUBCUR calculates the differences between these two files and generates a numerical output file in the CURVES format. The data is also output graphically (in postscript) to permit identification of the difference between the two structures. The specific parameter set (for example, backbone torsions, intrabase pair and interbase pair geometry, helical curvature) to be displayed is chosen by the user.

SUBCUR is viewed as complementary to and an extension of the CURVES algorithm and the *Dials-and-Windows* methodology for display of DNA conformation.[1] This latter approach provides an excellent framework within which the structure and dynamics of a DNA helix can be summarized. However, the amount of data included on each Dials-and-Windows plot is such that it becomes difficult to identify the specific differences in data for two conformations. The SUBCUR approach simplifies this procedure and hence aids in the interpretation of the Dials-and-Windows output. The following description and examples of SUBCUR output illustrate the complementarity of SUBCUR and CURVES/Dials-and-Windows.

An example of the SUBCUR output for backbone torsional parameters is shown in Figure 2. The plot shows the difference in a given parameter of strand 1 of helix 1 and strand 1 of helix 2 on the left of the plot, while similar

*Present address: Institut de Biologie Physico-Chimique, 13 Rue Pierre et Marie Curie, 75008 Paris, France.
Address reprint requests to Dr. Haworth at his current address: John Stauffer Pharmaceutical Sciences Center, University of Southern California, 1985 Zonal Ave., Los Angeles, CA 90033, USA.
Received 19 August 1992; revised 30 October 1992; accepted 3 November 1992

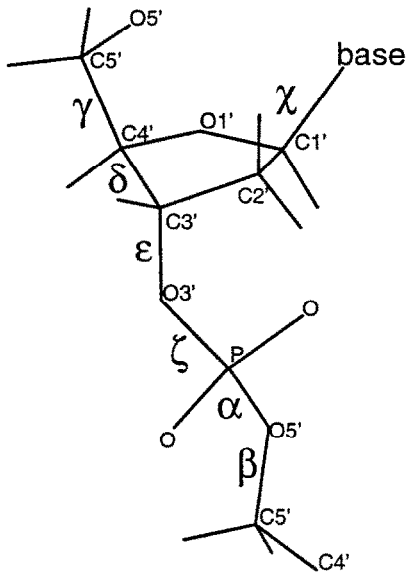

Figure 1. Definition of torsion angles in the sugar/phosphate backbone of a nucleotide unit.

information for strand 2 of the two helices is given on the right. The user-requested parameters are listed and the base sequence of each strand (5′ to 3′ for strand 1, 3′ to 5′ for strand 2) is exhibited for helices 1 and 2. In the example given two different conformations of the same helix are contrasted and hence only one sequence is given.

The difference between corresponding torsion angles is represented by a line (above the horizontal indicating a particular torsion to be more positive in helix 1 than in helix 2). The *scaling value* for the six backbone torsion parameters and the glycosidic torsional angle (Figure 1) is shown at the base of the plot. Hence the largest change among all these torsion angles (ζ of C7) is 71.77° and all other changes are displayed relative to this value. There is facility within the program to scale to user-defined values.

The scaling value at the top of the plot gives the maximum difference for the three sugar parameters (c1c2—the torsional angle O1′-C1′-C2′-C3′, c2c3, and the phase angle). All the other data for these parameters are scaled to this value. For a more immediate view of the changes in sugar pucker, the broad conformational range (C3′-endo, O1′-exo, etc.) is contrasted. The categorization is taken directly from the CURVES output, which divides the sugar pseudorotation cycle[4] into ten segments of 36°, with, for example, a C3′-endo pucker assigned to any phase angle between 0° and 36°. In the SUBCUR plot, a dot then represents a situation in which the sugar pucker remains unchanged from helix 1 to helix 2. Circles of increasing radii denote increasingly larger deviations of the pucker of a sugar ring in helix 2 from the pucker of the corresponding sugar in helix 1. Hence in Figure 2, the smallest circles indicate a change in sugar conformation such that the ring has moved from a one conformational space segment to a neighbouring segment. The larger circle for C1 indicates a much larger conformational change in the pseudorotation cycle.

A second example of a SUBCUR plot, contrasting the degree of curvature of two helices, is shown in Figure 3. Once again, the data is initially generated using a CURVES analysis. Following Ravishanker et al.,[1] the helix curvature is obtained by subtracting the 'end to end distance' (the through-space distance) of the global helical axis computed by CURVES from the *path length* (the through-axis distance) and expressing this value as a percentage of the path length (the percent shortening). Using data from a standard CURVES output, this can be done over the entire length of the helix or for any given part of the helix. The reader is referred to the original CURVES references[1] for a more detailed discussion of these parameters.

The SUBCUR plot in Figure 3 shows the difference in curvature for two dodecamer helices, the strand 1 base sequences of which are given in a 5′ to 3′ direction on the left of the plot (with helix 1 on the extreme left). The 1,12 column gives the difference in curvature (the difference in percent shortening) over the whole helix (from base pair 1 to base pair 12). The 1,11 column provides a measure of the different curvature over the two possible 11 base pair segments (that starting from the C1 base pair and that starting at the G2 base pair). Similarly, the 1,10 column shows similar data for the three possible 10 base pair segments and so on. Difference data for path lengths, end-to-end distances, etc., can also be output in this format, hence the inclusion of a 1,2

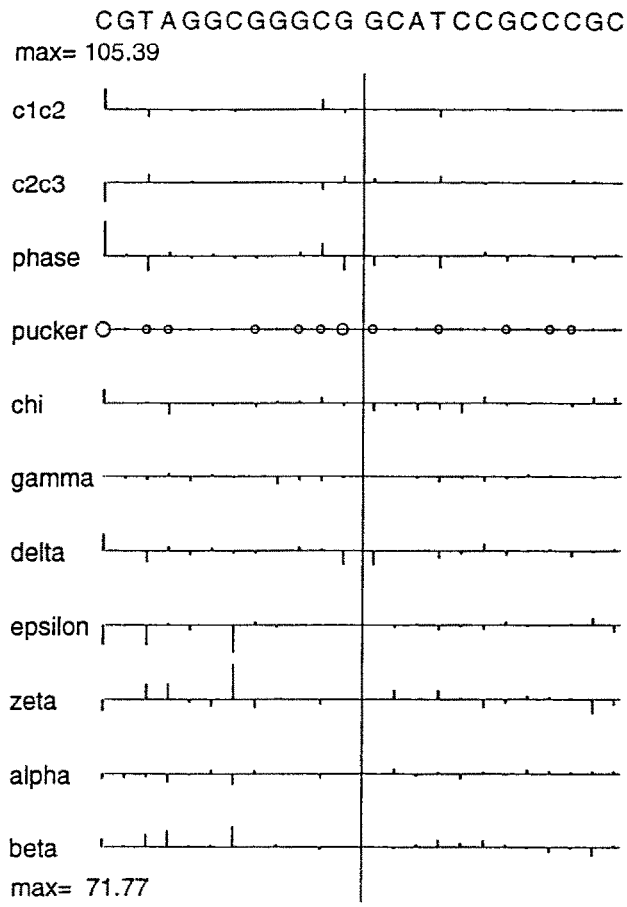

Figure 2. SUBCUR plot of the difference between two structures taken from a molecular dynamics trajectory of d(CGTAGGCGGGCG).d(CGCCCGCCTACG). Backbone parameters are plotted from γ to β in a 5′ to 3′ direction for a given base.

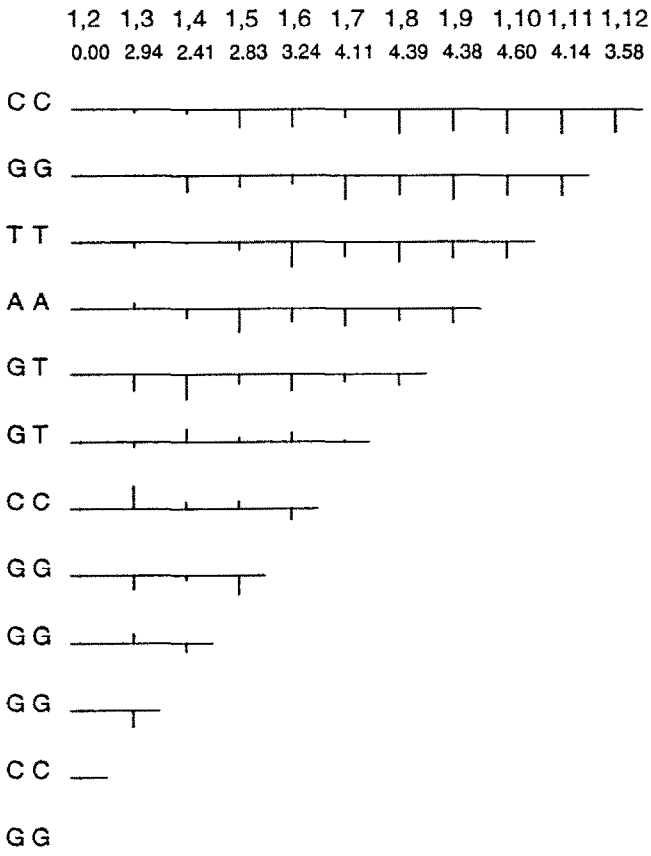

Figure 3. SUBCUR plot of the difference in curvature between conformations of the B-DNA dodecamers d(CGTAGGCGGGCG).d(CGCCCGCCTACG) and d(CGTATTCGGGCG).d(CGCCCGAATACG).

column. In the specific case of curvature difference all 1,2 values are necessarily zero. Data in each column are plotted relative to the largest difference in that column, the value of which is given at the head of the column. A line below the horizontal indicates the specific segment of helix 2 has greater curvature than the equivalent segment of helix 1.

APPLICATION

As illustrative examples, the SUBCUR plots in Figures 2 and 3 use data taken from a molecular dynamics trajectory of the B-DNA duplexes d(CGTAGGCGGGCG).d(CGCCCGCCTACG) and d(CGTATTCGGGCG).d(CGCCCGAATACG). The former sequence contains the Sp1 promoter recognition site (5′-AGGCGGG) found, for example, in the human monamine oxidase gene.[5] The latter sequence is part of a non-Sp1-binding mutant. The molecular dynamics simulation was conducted using the AMBER 4.0 force field[6] with the DNA solvated in a box of 2500 TIP3P water molecules. A period of 80 ps of dynamics was recorded under periodic boundary conditions and using the SHAKE algorithm for bond length constraint. A detailed description and analysis of the simulation will be given elsewhere.[7]

Figure 2 shows the difference in the backbone and sugar parameters of two structures taken from the molecular dynamics trajectory of the Sp1 binding sequence. The first helix is an average of structures generated between 38 ps and 45 ps while the second helix is a similar average taken between 55 ps and 62 ps. The data suggest that a B_I-to-B_{II} conformational transition has occurred at the second CG step (C7–G8) in strand 1. The most significant differences in backbone torsion parameters between the two structures occur at ϵ and ζ of C7. ζ is 71.77° more negative in helix 2 than in helix 1. ϵ has correspondingly moved to a more positive orientation, exactly as expected for a B_I-to-B_{II} conformational shift. Hence SUBCUR has successfully located a specific structural transition in the trajectory.

Figure 3 contrasts the curvature of the average structures of the wild type and mutant sequences taken over the second half of the molecular dynamics trajectory (from 40 ps to 80 ps). The 1,12 data shows the mutant conformation is bent significantly more than the wild type helix. It is also apparent, however, that the increased curvature in the mutant is not uniform over the entire helix. Several 1,3 segments of the wild type sequence exhibit greater curvature. This is particularly true of the G8–G10 region where the localized curvature of the wild type helix might be of importance in protein recognition.[7]

REFERENCES

1 Lavery, R. and Sklenar, H. *CURVES 3.1 1990: Helical Analysis of Irregular Nucleic Acids.* Institut de Biologie Physico-Chimique, Paris, France; Ravishanker, G., Swaminathan, S., Beveridge, D.L., Lavery, R., and Sklenar, H. *J. Biomol. Struct. Dyn.* 1989, **6,** 669–699; Lavery, R. and Sklenar, H. *J. Biomol. Struct. Dyn.* 1989, **6,** 655–667; Lavery, R. and Sklenar, H. *J. Biomol. Struct. Dyn.* 1988, **6,** 63–91

2 Uhlmann, E. and Peyman, A. *Chem. Rev.* 1990, **90,** 544–584

3 Haworth, I.S., Elcock, A.H., Rodger, A., and Richards, W.G. *J. Biomol. Struct. Dyn.* 1991, **9,** 553–569

4 Saenger, W. *Principles of Nucleic Acid Structure.* Springer-Verlag, New York, 1983

5 Zhu, Q-S., Grimsby, J., Chen, K., and Shih, J.C. *J. Neuroscience* 1992, **12,** 4437–4446

6 Pearlman, D.A., Case, D.A., Caldwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P., and Kollman, P.A., 1991, AMBER 4.0, University of California, San Francisco

7 Haworth, I.S., unpublished results