

CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures

Christophe Blanchet, Marco Pasi, Krystyna Zakrzewska and Richard Lavery*

Université Lyon 1, CNRS, UMR 5086, Bases Moléculaires et Structurales des Systèmes Infectieux, IBCP FR3302, 7 passage du Vercors, F-69367, France

Received February 14, 2011; Revised April 19, 2011; Accepted April 20, 2011

ABSTRACT

Curves+, a revised version of the *Curves* software for analyzing the conformation of nucleic acid structures, is now available as a web server. This version, which can be freely accessed at http://gbio-pbil.ibcp.fr/cgi/Curves_plus/, allows the user to upload a nucleic acid structure file, choose the nucleotides to be analyzed and after optionally setting a number of input variables, view the numerical and graphic results online or download files containing a set of helical, backbone and groove parameters that fully describe the structure. PDB format files are also provided for offline visualization of the helical axis and groove geometry.

INTRODUCTION

The rapidly growing database of high-resolution nucleic acid structures (1) has brought to light not only the broad range of structures that can be adopted by nucleic acids, but also the role of their conformational adaptability in a wide range of recognition processes. Recent years have also seen considerable progress in molecular simulations of nucleic acids. Such simulations are now capable of not only reproducing experimental conformations, but also of generating time trajectories that provide detailed dynamic information over increasingly long durations (2,3).

For both experimental and simulated structures, it is important to be able to analyze and compare conformations in detail. Analyzing the structure of nucleic acids poses some particular problems. Since the earliest days of fiber diffraction studies (4,5), it became useful to describe nucleic acids in terms of helical parameters (rise, twist, etc.). However, with the determination of high-resolution single crystal structures, it also became clear that globally helical nucleic acids could be irregular (6). This observation led Lavery and Sklenar to look for a way of analyzing

irregular structures without losing the notion of a helical axis. The resulting program, *Curves* (7,8), provided this information, along with a full set of helical parameters, a description of backbone geometry (single bond torsions and sugar puckers) and a preliminary analysis of helical grooves.

After accumulating the users comments on *Curves* over many years, we set about redesigning the algorithm to solve a number of problems, make the analysis faster and automate the analysis of molecular dynamics trajectories. The result was *Curves+* (9). *Curves+* uses a simpler and faster method for calculating helical parameters, resolves problems between so-called local and global sets of helical parameters, using a strategy similar to 3DNA (10) [also available as a webservice (11)], but maintains the calculation of a curvilinear helical axis and offers a more robust method of analyzing groove geometry. Both the axis and groove geometry can be viewed graphically within the webserver. It also takes into account changes in the conventions for describing nucleic acids adopted at a meeting in Tsukuba (12). *Curves+* is freely available as Fortran source code and is accompanied by user guides and a number of utility programs, including Canal, which can help in making a statistical summary of molecular dynamics trajectories analyzed using *Curves+* (see http://gbio-pbil.ibcp.fr/Curves_plus/Curves+.html).

However, it is also clear that not all users want to download and compile their own version of *Curves+* and so we now propose a web server version. This version has all of the options available in *Curves+* and provides numerical and graphical results online as well as in downloadable files. The only restriction is that it cannot be used for analyzing molecular dynamics trajectories. Such analyses would require more computing resources than we can provide and are also significantly slowed down by the transfer of very large files. For this type of analysis, it is necessary to download and compile your own version of *Curves+* (and the accompanying program Canal).

*To whom correspondence should be addressed. Tel: +33 4 72 72 26 37; Fax: +33 1 72 72 26 04; Email: richard.lavery@ibcp.fr

MATERIALS AND METHODS

The algorithms developed for Curves+ have been described in detail in our earlier publication (9). Here, we simply outline the main steps.

Reference frame

The Curves+ analysis begins by fixing a reference frame to each of the chosen bases. This frame follows the definitions set out in the Tsukuba convention (12). It is calculated starting from the atoms defining the glycosidic bond. In unmodified nucleic acids these are (N9-C1') for a purine and (N1-C1') for a pyrimidine and the vector perpendicular to the base plane defined by a vectorial product between the two bonds, $(N9-C1') \times (N9-C4)$ for purines and $(N1-C1') \times (N1-C2)$ for pyrimidines (see (9) for details). Since the bases can sometimes be deformed, Curves+ optionally replaces input base coordinates with standard base coordinates contained in a pre-defined library using least-squares fitting (we recommend keeping this option, see $fit = .t.$ in Table 1). This avoids deformations leading to unrepresentative frame orientations. Note that the three atoms required for building the reference frame of each base must be present, but other missing atoms will not block the analysis.

Intra-base pair parameters

Once the base reference frames are constructed, it is possible to obtain the helical parameters for all base pairs. This is achieved by calculating the rigid body transformation that maps one base reference frame into the other. In standard double helices, we take into account pseudodyad symmetry by rotating the reference frame of the second strand 180° around the pseudodyad axis that points into the grooves (with reverse Watson-Crick pairs, this transformation is about the base normal). The rigid body transformation consists of a translation and a rotation. Making a half-translation and a half-rotation creates a reference frame for the base pair. The intrabase pair rotational parameters can now be calculated as the three components of

the rotation in the coordinates of the mid-frame (buckle with respect to the vector pointing into the major groove, propeller with respect to the vector pointing along the long axis of the base pair and opening with respect to the vector normal to the base pair). The three translations are calculated in the same way as the components of the translation in the mid-frame (shear, stretch and stagger, respectively, using the same vectors as for the rotational parameters).

Inter-base pair parameters

Following the same scheme, and starting from the base pair reference frames, we now calculate the rigid-body transformation between two successive frames (corresponding to a base pair step) and generate a new mid-frame. Calculating the components of the rigid-body rotation and translation in this new mid-frame leads to the three interbase pair rotational parameters (tilt, roll and twist) and the three translational parameters (shift, slide, rise).

Helical axis

The helical axis is calculated from the screw axes which link successive base pair reference frames. Each screw transformation defines a vector U passing through a point p . This point can be moved to either base pair reference frame by sliding it along U until the vector from p to the origin of the base pair frame is perpendicular to U . For all but the terminal base pairs, this procedure leads to two U vectors at each base pair that are averaged and then smoothed, using a polynomial weighting function taking into account up to four axes on either side, weighting each axis i with $w_i = 1 - i^2/5^2$. This yields the final (potentially curvilinear) helical axis defined by vectors U_H and point P_H at the level of each base pair. Note that, if successive base pairs are not helically symmetric (for example, as in Z-DNA, which has a dinucleotide repeat symmetry), you can set the Curves+ symmetry repeat option ($isym$) to any

Table 1. Curves+ user options and their default values

Option	Default	Meaning
$isym$	1	Make an analysis assuming a symmetry repeat every $isym$ nucleotides (e.g. for Z-DNA $isym = 2$ is necessary, for poly(dCG) _n it might be useful). $isym$ only influences the helical axis calculation
$wback$	2.9 Å	Groove width is measured between spline curves running through the phosphorus atoms. This width is reduced by $2 \times wback$ which represents the average radius of the phosphodiester backbone
$wbase$	3.5 Å	Groove depth is measured as a distance from a groove width vector down to the long axis of the base pairs. $wbase$ is subtracted from this distance to account for the average half-width of a base pair.
$back$	P	Sets the atom/s used for defining the backbone splines. If one atom name is given it applies to all strands. If different atoms are used for different strands, separate the atom names with / (e.g. P/C5*)
fit	true	Fits a standard base geometry to the input base coordinates using a least squares algorithm. This is recommended for most analyses and is particularly important for low resolution data, or for snapshots from molecular dynamics trajectories
$circ$	false	Set to true if analyzing a circular nucleic acid
$line$	false	Set to true to impose a linear helical axis
$refo$	false	Set to true to use the old base reference frame of Curves (only for comparison purposes).
$test$	false	If $test$ is true, extra output is provided. This mainly includes the quality of the rms fit of standard bases to each input base (if $fit = true$). It also results in listing the points (P_H) and vectors (U_H) defining helical axis
$zaxe$	false	If true, the helical axis is not calculated, but is assumed to be the Cartesian Z-axis. This is occasionally useful for structures built using modeling software, where the axis is known from the construction technique.

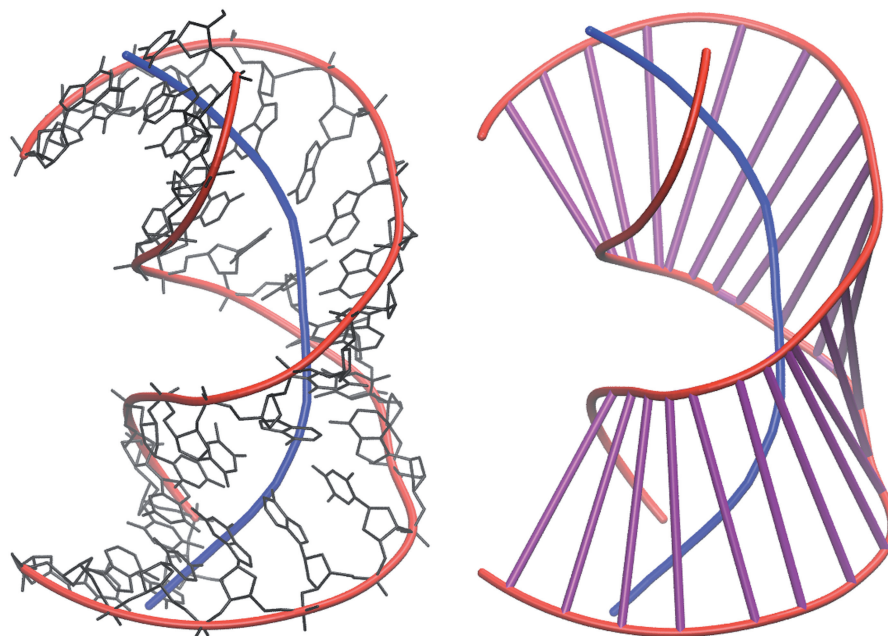


Figure 1. Graphical output of a Curves+ analysis for DNA complexed with the TATA-box binding protein (PDB code 1CDW). The DNA is shown in black, the calculated helical axis in blue, the backbone splines in red and the minor groove width vectors in purple. For clarity, the DNA is only shown in the left-hand image and the groove vectors only in the right-hand image. The impact of protein binding, opening up the minor groove and bending the DNA, is clearly visible.

positive integer in order to derive the helical axes from appropriate combinations of base pair frames.

Base pair-axis parameters

These parameters are calculated in a similar way to the intrabase pair parameters, using the rigid-body transformation between the base pair reference frame and a reference frame built on the local helical axis using the point P_H , the vector U_H and a pseudodyad vector, perpendicular to U_H , defined using the equivalent vector of the base pair reference frame. Following the definition of P_H , there are only two non-zero translational base pair-axis parameters (X displacement and Y displacement) and two rotational parameters (inclination and tip).

Helical rise and twist

The helical axis reference frames created for calculating base pair-axis parameters also serve to define the so-called helical rise and twist. These parameters are obtained as the translation and rotation corresponding to the screw transformation between two successive helical axis reference frames. Helical rise and twist (denoted H-Ris and H-Twi in the Curves+ output) describe the translation and rotation between successive base pairs with respect to the helical axis rather than with respect to a base pair step considered in isolation (which is the case with the standard interbase pair parameters, including rise and twist). In B-like double helices, these two sets of parameters will be very similar. However, when the base pairs are inclined with respect to the helical axis (as in A-DNA) they can differ considerably. Using a typical A-DNA (PDB code 1d13) as an example, the average rise and

twist are 3.26 Å and 31.7°, respectively, whereas the average H-Ris and H-Twi are 2.72 Å and 33.1°, respectively (9). Both sets of values can be useful, depending on whether you are interested in a single base pair step, or in the overall helical geometry.

Backbone parameters

Curves+ calculates all the single bond torsions. For each nucleotide, following the 5'→3' direction yields α (O3'-P-O5'-C5'), β (P-O5'-C5'-C4'), γ (O5'-C5'-C4'-C3'), δ (C5'-C4'-C3'-O3'), ϵ (C4'-C3'-O3'-P) and ζ (C3'-O3'-P-O5'), plus the glycosidic angle χ (O4'-C1'-N1-C2 for pyrimidines and O4'-C1'-N9-C4 for purines). The sugar ring is described using the pseudorotational parameters phase and amplitude that are based on all five ring torsions (13) [see also (9)]. Each sugar is also described with its conventional pucker name (C2'-endo, C3'-exo, etc.).

Groove geometry

In order to define the geometry of the grooves, Curves+ calculates a cubic spline curve for each nucleotide strand, by default passing through the phosphorus atoms P_i and tangential to the vectors $P_{i-1}-P_{i+1}$ (9). If desired, the input parameter 'back' allows the user to choose an alternative atom for the spline definition and, optionally, different atoms for different strands. Each spline curve is divided into a number of equal intervals. Starting from vectors joining equivalent points along these curves, we rotate the vectors in either direction looking for vectors of minimal length. In a normal double helix, there will be a minimum close to the starting vector in both directions

corresponding to the major and minor grooves. Groove depths at each base pair level are then defined using the distance from the center point of the minimal length vector to the mid-point of the long axis of the base pair, defined as a vector from a purine C8 atom to a pyrimidine C6 atom. To make groove widths and depths more relevant for understanding the space actually available within a helical groove, both measures are reduced to account for the average radius of the backbones and the average half-widths of the base pairs (see the user options `wback` and `wbase`). Note that groove widths and depths can only be defined where minimal length vectors joining the backbone splines can be found. This is not possible towards the ends of a helical nucleic acid fragment and may also become impossible elsewhere if the backbones are strongly deformed.

Special cases

Most of the above descriptions have assumed that we are analyzing a double stranded, antiparallel helical structure, with standard nucleic acid chemistry and standard Watson–Crick base pairs. However, none of these characteristics is required. The program can deal with structures having between one and four strands, with any orientations. Non-standard chemistries can be taken into account by defining new single bond torsions to analyze (using the backbone library, see below), or new atoms to use in constructing the base reference frame and, optionally, for the least-squares fitting of a standard base (using the base library, see below). It is also possible to allow for abasic sites, or for bases out of helical stack that could perturb the helical axis calculation. Curves+ can also treat strands of different lengths, gaps, and single-stranded ends. Note that because of this flexibility, and because the calculated helical axis depends on which nucleotides are analyzed, the user must specify the nucleotides within each strand to be treated and organize the strand input to indicate base pairing.

WEB SERVER

Input data

The input data to the Curves+ web server consists of three parts: the user options; the base and backbone libraries; the structure to analyze and information on the nucleotides to include in the analysis.

User options. The user can modify the Curves+ analysis in a number of ways to suit particular cases. All options have default values that are appropriate for most standard analyses. The meaning of all the options and their default values are given in Table 1.

Base library. The base library defines which atoms are to be used to calculate the base reference frame for each type of base and also provides the standard ring geometry of each base. If the option `fit = true`, these standard geometries will be least-squares fitted (14) to each base in the input structure. This avoids deformations of the bases leading to unrealistic helical parameters. Deformations

often occur with snapshots drawn from molecular dynamics trajectories and, occasionally, with low-resolution experimental structures. The standard base library provides data for adenine (A), guanine (G), thymine (T), cytosine (C), inosine (I), uracil (U), pseudouracil (P) and Y-base (Y) (the latter two being useful in tRNAs). If it is necessary to analyze other bases, the user can modify the base library to include them.

Backbone library. The backbone library defines which backbone torsions and ring puckers are to be calculated. It first defines the backbone torsions: α (O3'-P-O5'-C5'); β (P-O5'-C5'-C4'); γ (O5'-C5'-C4'-C3'); δ (C5'-C4'-C3'-O3'); ϵ (C4'-C3'-O3'-P); ζ (C3'-O3'-P-O5'). For the glycosidic angle, χ , the atoms depend on whether a purine ($\chi = \text{O4'-C1'-N9-C4}$) or a pyrimidine ($\chi = \text{O4'-C1'-N1-C2}$) is being analyzed. Lastly, the standard ribose/deoxyribose sugar ring is indicated by the ring atoms C1'-C2'-C3'-C4'-O4'. These atoms enable the ring torsion angles to be calculated. These torsions are not listed directly, but are used to determine each ring pucker in terms of phase angle and amplitude (13). If users work with chemically modified nucleic acids, it is possible to define the new torsions and rings to be analyzed by modifying the standard library.

The input format for both libraries is explained in the header to the standard files (`standard_b.lib` for the bases and `standard_s.lib` for the backbone). These files can be downloaded from http://gbio-pbil.ibcp.fr/Curves_plus/Downloads.html.

Structure, strand direction and nucleotide numbers. The structure to be analyzed must be read in standard PDB format (15). Note that any amino acids, water molecules or hetero atoms are automatically eliminated from the PDB file at input. Also, if any atoms appear as multiple copies within a single nucleotide, the first copy will be used (if this is not desired the file should be edited before analysis). Also, since backbone atom names use either of two conventions (i.e. with a single quote, C5', or with a star, C5*), Curves+ converts all quotes to stars on input. Remember this if you choose new atom names to define the backbones via the parameter 'back'). For nucleotides, both upper and lower case names are recognized, but only the first letter of the unit name is used (the exception is a first letter 'D' commonly indicating deoxynucleotides, which will be automatically removed).

Curves+ can treat nucleic acids with up to four strands. For each strand, the user must indicate the strand direction (5' \rightarrow 3' or 3' \rightarrow 5') and the numbers of the nucleotides to be analyzed. Since nucleotides in different strands often have the same unit numbers, the input to Curves+ does not use PDB unit numbers, but rather indicates nucleotides by the order in which they appear in the PDB file, i.e. 1 = 1st unit, 2 = 2nd unit and so on. Note that since proteins are automatically eliminated at input, the nucleic acid will generally be at the beginning of the file and so working out the numbering is simplified. If molecules other than proteins precede the nucleic acid, be careful to count these subunits when working out your numbering. Note that you only need to input numbers

for the nucleotides you want to analyze. It is never necessary to edit the PDB file to remove any other nucleotides.

Nucleotide numbers can be input as a simple series of integers (e.g. 1 2 3 . . .) or you can save time using a colon to indicate a consecutive series (e.g. 1:10 implies all numbers from 1 to 10). These two types of input can be combined. There are also some special cases to consider. Zero can be used to indicate missing nucleotides. This enables treatment of duplex nucleic acids with single-stranded ends or gaps. Nucleotide numbers can also be negative, meaning the corresponding base should not be used in calculating the helical axis. This is useful when a base is flipped out of helical stack or inverted because of unusual base pairing. Including such bases in the helical axis calculations would lead to strong perturbations. Note that the nucleotide order within each strand should be so chosen that paired bases occur in identical positions. We can illustrate this using the structure of a B-Z junction with single-stranded ends (each two nucleotides in length) and two flipped-out bases in the middle (PDB code 2ACJ). The full structure can be analyzed using the following input (note the use of zeros to ensure the paired bases fall at the same positions):

```
Strand 1 (5'→3'): 1:17 0 0
Strand 2 (3'→5'): 0 0 34:18
```

This will give correct intrabase pair, inter base pair (excepting H-Twi and H-Ris) and backbone parameters, but an incorrect helical axis and axis-bp parameters. A correct helical axis (and groove) analysis requires eliminating the unstacked bases as follows:

```
Strand 1 (5' →3'): 3:10 12:17
Strand 2 (3'→5'): 34:27 25:20
```

Once the strand data has been provided by the user, pressing the 'submit' button starts the calculation and moves the user to the results page where the output can be viewed or saved in files. Note that the server can be tested with pre-loaded data (a simple B-DNA). For this option, click on 'test page' at the top of the input form and then press 'submit'. Other examples, including more complicated cases, are available for download from http://gbiopbil.ibcp.fr/Curves_plus/Downloads.html.

Output data

Once the calculation is completed, you will see a new page displaying the names of the files output from the Curves+ analysis. The helical and backbone parameters are given in the STR.lis file, where STR is the name of the input structure file. Other files are in PDB format and can be used to make images of the results in your own graphic programs. The helical axis is given in the STR_X.pdb file and the backbone and groove geometry in the STR_b.pdb file. Note that if the calculation fails for some reason, either or both of the graphic output files may not be generated. In this case, check the STR.lis file to find the source of the error.

Visualization. Before discussing the content of the output files, we note that you can immediately visualize the results of the analysis using the Jmol applet integrated into the

web server. To do this, click on 'axis and groove analysis in 3D'. This will open a new page and display the input molecule along with the helical axis, the backbone splines and the vectors illustrating the groove widths (discussed below). This image can be rotated, zoomed and translated using the mouse. The page also provides several options for modifying the image. The user can choose to display or hide the nucleic acid, the helical axis, the backbone splines or the groove width vectors, individually for each groove. The nucleic acid display can also be modified from a line view to either CPK or surface representations. The image can also be reoriented so the vector joining the ends of the helical axis is vertical, set back to its original orientation or made to spin slowly. The background color of the image can be modified (by clicking on the small colored rectangles below the image).

Output files. The STR.lis output can be viewed in the browser (by a left-click on its name, note some web browsers may add a .txt extension to this file by default) or saved in a file for future reference (by a right-click on its name). Following a summary of the input data and user options, this file contains the output parameters in five sections: (A) base pair-axis parameters and also the incremental bend in the helical axis at each base level, (B) intra-base pair parameters, (C) inter-base pair parameters, (D) backbone parameters and sugar puckers and (E) groove parameters [For an illustration of the helical parameters, see http://gbio-pbil.ibcp.fr/Curves_plus/Helical_parameters.html or the supplementary data to our earlier publication (9)]. Sections (A–C) also list the average helical parameters for the entire fragment analyzed and section (A) gives the overall axis bend. Note that all translational parameters are given in angstroms and all rotational parameters are in degrees.

The groove parameters in section (E) require some further explanations. Groove geometry is measured for any nucleic acid with more than one strand. WIJ and DIJ indicate the widths and depths measured between strands I and J. In a canonical right-handed duplex, with the first strand oriented 5'→3', W12 and D12 correspond to the minor groove and W21 and D21 to the major groove. If you are in any doubt as to which groove WIJ/DIJ refers to, see the 3D visualization of the Curves+ results discussed above. Groove widths are measured as minimal distances between the backbone spline curves passing through the phosphorus atoms. Values are given at base pair levels and halfway between these levels. As mentioned above, some values may be absent if no minimal distances can be found.

The two other output files provided for the user, STR_X.pdb, which contains the calculated helical axis, and STR_b.pdb, which contains the backbone splines and the groove width vectors, can be downloaded for creating your own images in a molecular graphics programs such as VMD (16) or Chimera (17).

CONCLUSION

The Curves+ web server provides a simple way of analyzing the conformation of a wide range of nucleic

acid structures. After reading a PDB file, the user indicates the nucleotides to be treated and optionally sets a number of input parameters. Numerical and graphic results can be downloaded, but it is also possible to view the calculated axis and groove geometry within the web server.

ACKNOWLEDGEMENTS

The authors thank M. Moakher (Ecole National d'Ingénieurs, Tunis), D. Petkeviciute and J.H. Maddocks (Institut de Mathématiques B, Swiss Federal Institute of Technology, Lausanne) for their contributions to the development of the Curves+ algorithms.

FUNDING

Funding for open access charge: CNRS; ANR Blanc project ALADDIN (R.L. and K.Z.); IBISA project GRISBI (C.B.).

Conflict of interest statement. None declared.

REFERENCES

- Berman, H.M., Westbrook, J., Feng, Z., Iype, L., Schneider, B. and Zardecki, C. (2002) The nucleic acid database. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 889–898.
- Orozco, M., Noy, A. and Pérez, A. (2008) Recent advances in the study of nucleic acid flexibility by molecular dynamics. *Curr. Opin. Struct. Biol.*, **18**, 185–193.
- Mackerell, A.D. and Nilsson, L. (2008) Molecular dynamics simulations of nucleic acid-protein complexes. *Curr. Opin. Struct. Biol.*, **18**, 194–199.
- Rosenberg, J.M., Seeman, N.C., Kim, J.J., Suddath, F.L., Nicholas, H.B. and Rich, A. (1973) Double helix at atomic resolution. *Nature*, **243**, 150–154.
- Rosenberg, J.M., Seeman, N.C., Day, R.O. and Rich, A. (1976) RNA double helices generated from crystal structures of double helical dinucleoside phosphates. *Biochem. Biophys. Res. Commun.*, **69**, 979–987.
- Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980) Crystal structure analysis of a complete turn of B-DNA. *Nature*, **287**, 755–758.
- Lavery, R. and Sklenar, H. (1989) Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.*, **6**, 655–667.
- Lavery, R. and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.
- Lavery, R., Moakher, M., Maddocks, J.H., Petkeviciute, D. and Zakrzewska, K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.*, **37**, 5917–5929.
- Lu, X.J. and Olson, W.K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.*, **3**, 1213–1227.
- Zheng, G., Lu, X.J. and Olson, W.K. (2009) Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res.*, **37**, W240–W246.
- Olson, W.K., Bansal, M., Burley, S.K., Dickerson, R.E., Gerstein, M., Harvey, S.C., Heinemann, U., Lu, X.J., Neidle, S. *et al.* (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, **313**, 229–237.
- Westhof, E. and Sundaralingam, M. (1983) A method for the analysis of puckering disorder in five-membered rings: The relative mobilities of furanose and proline rings and their effects on polynucleotide and polypeptide backbone flexibility. *J. Am. Chem. Soc.*, **105**, 970–976.
- McLachlan, A.D. (1979) Gene duplications in the structural evolution of chymotrypsin. *J. Mol. Biol.*, **128**, 49–79.
- Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L. and Jain, S. (2002) The protein data bank. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 899–907.
- Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.