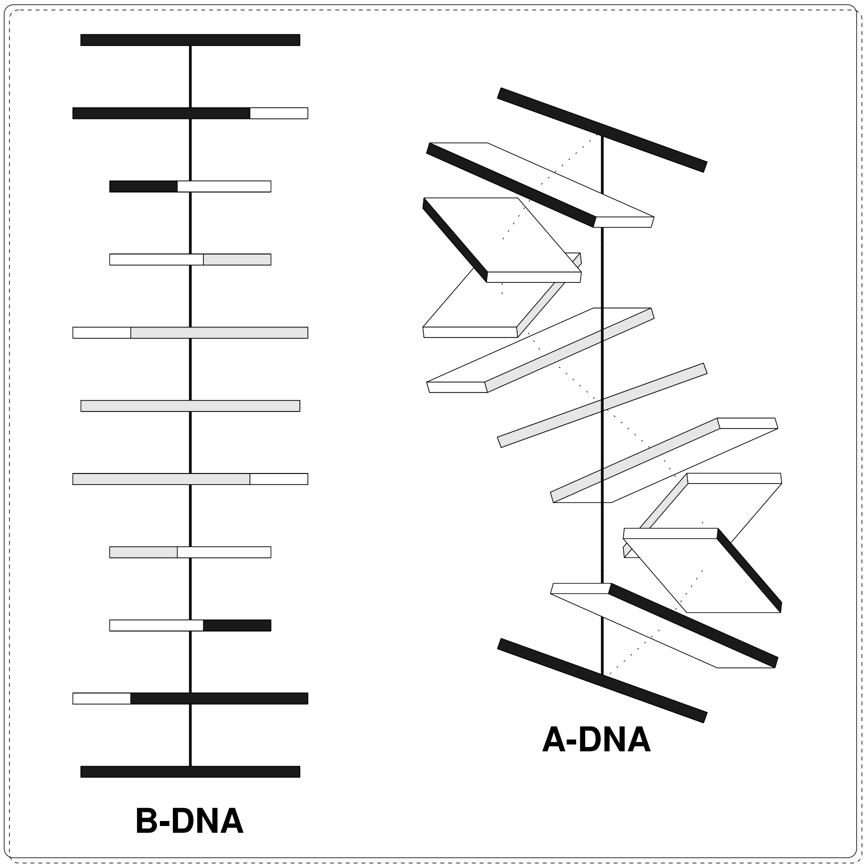
3DNA (v1.5) — A 3-Dimensional Nucleic Acid

Structure Analysis and Rebuilding Software Package



by

Xiang-Jun Lu

xiangjun@rutchem.rutgers.edu

Wilma K. Olson Laboratory

Department of Chemistry, Rutgers University

610 Taylor Road, Piscataway, NJ 08854

November 12, 2002

Contents

1 Introduction

This manual is far from complete, but should get you started.
I would greatly appreciate your help in making it suit your needs better.

Our recent comparison studies (Lu & Olson, 1999; Lu *et al.*, 1999) on seven popular analyzing programs for nucleic acid structures demonstrated that the choice of reference frame rather the mathematical formulation has the greater effect on the calculated parameters. Broadly speaking, there are two classes of reference frames. One is defined by the purine (R) C8 atom and pyrimidine (Y) C6 atom, and the mean base-pair plane as in `FreeHelix/NewHelix` (Dickerson, 1998). The other uses the average of the two constituent base reference frames in a base-pair as in `Curves` (Lavery &

Sklenar, 1989). While each has its own advantages and limitations, the base-centered one was recommended at a workshop held on January 13-14, 1999 at the AIST-NIBHT Structural Biology Center in Tsukuba, Japan (Olson & *et al.*, 2000).

3DNA is a comprehensive software package for the analysis and rebuilding of nucleic acid structures based on the new recommended reference frame. Mathematically, it follows the `CEHS` definition (El Hassan & Calladine, 1995) as implemented in `SCHNAaP/SCHNArP` (Lu *et al.*, 1997a; Lu *et al.*, 1997b) with the global helical parameters replaced by a new set of *local* ones. This matrix based algorithm is rigorous and reversible in that it allows for exact reconstruction of the structure from a set of parameters which can be deduced from the structure. The parameters such defined also have simple geometrical meaning thus the distance between base-pair centers is $\sqrt{\mathrm{Shift}^2 + \mathrm{Slide}^2 + \mathrm{Rise}^2}$, the bending angle between two base-pairs is $\sqrt{\mathrm{Roll}^2 + \mathrm{Tilt}^2}$, and the angle between base-pair normal and local helical axis is $\sqrt{\mathrm{Inclination}^2 + \mathrm{Tip}^2}$.

3DNA is a completely new written software package conforming strictly to the `ANSI/ISO C`. It is much more robust and efficient than `SCHNAaP/SCHNArP` and has new and improved features. Its analysis part can automatically classify a dinucleotide step as A-, B-, or TA-like (Guzikevich-Guerstein & Shakked, 1996) based on the positioning of phosphorus atoms with respect to the middle frame. This can be very useful for pinpointing conformational transitions in ligand-bound DNA, especially in a short fragment which is often neglected by other analyzing programs (Lu *et al.*, 2000). High quality "standardized" base-stacking diagrams can be generated by a utility program following the analysis. Authentic `CEHS/SCHNAaP` parameters are also available, which can be used for comparisons with the 3DNA recommended ones to see how a difference in reference frames can affect the calculated parameters, especially for Rise in heavily deformed DNA structures. Also note that `FreeHelix/NewHelix` parameters are very similar to the `CEHS/SCHNAaP` ones due to their similarity in reference frames. For completeness, local parameters using the seven methods (Lu & Olson, 1999; Lu *et al.*, 1999) based on the new recommended reference frame are also calculated. Obviously, they are all quite similar to 3DNA recommended ones.

The rebuilding part of 3DNA can be used for generating sequence-dependent atomic structures, without or with the sugar-phosphate backbone, suitable as starting point for molecular mechanics calculations and molecular dynamics simulations. It is also very convenient to generate publication quality Calladine-Drew style schematic representations of DNA, either in PS, XFIG[1], or Raster3D (Merritt & Bacon, 1997)[2] format. With XFIG, the picture can be edited and annotated, combined with others, or exported to image formats such as GIF, JPEG and TIFF. With Raster3D, the Calladine-Drew style block representation of bases can be combined with the backbone and protein in either atomic or schematic presentations.

Fifty five different types of fiber DNA and RNA structures based on the work of Chandrasekaran &

[1]http://www.xfig.org/

[2]http://www.bmsc.washington.edu/raster3d/raster3d.html

Arnott (1989) can also be conveniently generated. The atom naming and ordering conventions of the NDB are strictly followed which make it easy for direct comparison with X-ray crystal structures or theoretical models.

2 Version history

To be completed.

3 Installation

The latest version of 3DNA, in binary form for Linux, SGI and Windows, is available from the following URL: `http://rutchem.rutgers.edu/~xiangjun/3DNA/`. Download the version that fits your system and then do the following:

- `gunzip Linux_X3DNA_v1.5.tar.gz`, you get file `Linux_X3DNA_v1.5.tar`
- `tar xvf Linux_X3DNA_v1.5.tar`, you get directory `X3DNA` and four subdirectories underneath: `BASEPARS`, `FIBER`, `bin`, and `Examples`, the contents of which are detailed below.
- If `X3DNA` is not directly under your home directory, set environment variable `X3DNA` as follows:

```
in csh | tcsh:  setenv X3DNA Your_Directory_Containing_X3DNA
  e.g.,  setenv X3DNA /usr1/xiangjun/X3DNA
in bash | sh: export X3DNA=Your_Directory_Containing_X3DNA
  e.g.,  export X3DNA=/usr1/xiangjun/X3DNA
in Windows (suppose 3DNA is installed under C:\X3DNA):
  set X3DNA=C:\X3DNA
  Or do the followings (globally):
     Start
         Settings
             Control Panel
                 System
                     Advanced
                         Environment Variables
                             at "User variables" section
                                Variable: X3DNA
                                Variable Value: C:\X3DNA
```

Then append to your command search path the 3DNA `bin` subdirectory:

```
in csh | tcsh:  set path = ($path $X3DNA/bin)
in bash | sh: export PATH=$PATH:$X3DNA/bin
in Windows:
  Start
      Settings
          Control Panel
              System
                  Advanced
                      Environment Variables
                          at "System variables" section
                             click on Path and "Edit" it
                              by appending ";C:\X3DNA\bin" to it.
```

Another option for Windows users is to install `CygWin` from `http://www.cygwin.com/`.

3.1 BASEPARS

This directory contains standard residue geometry files and other parameters controlling various aspects of 3DNA. They are all in text format, so users can view the structures with `rasmol` or edit the parameters as they see fit.

- `Atomic_?.pdb (? = A, C, G, T or U)` are the default standard residue geometries used by 3DNA for analyzing and rebuilding full atomic nucleic acid structures in PDB format.

  Under its subdirectory `ATOMIC`, there are four sets of standard residue geometry files:

  `ADNA_std?.pdb`, `BDNA_std?.pdb`, `NDB96_std?.pdb` and `RNA_std?.pdb`.

  The base geometries by Clowney *et al.* (1996) are used (downloaded from the NDB archive: `http://ndbserver.rutgers.edu/NDB/archives/index.html`). The `NDB96` set includes base and C1′ atoms, without sugar-phosphate backbone. `ADNA` set uses C3′-endo sugar-backbone conformation as defined by Leslie *et al.* (1980) fiber studies, but with a χ torsion angle of $-157°$, average of high resolution single crystal X-ray oligonucleotide structures. `RNA` set is the same as `ADNA` except for an additional O2′ atom for each residue. `BDNA` set is similarly defined as `ADNA` except for a C3′-endo sugar fiber sugar-backbone conformation and the $-108°$ χ torsion angle.

  The `NDB96` set is the default. To use another data set, simply overwrite the corresponding `Atomic_?.pdb` in `BASEPARS` or copy them to your current working directory. A utility program, `cp_std` (see below), can do this automatically for you. You can also use other residue geometries with the help of the utility program `std_base`.

Note the standard set contains only the five common residues, A, C, G, T and U. Residue I can be got by deleting the N2 atom from G. Their modified counterparts, +A, +C, +G, +I, +T, +U which are changed to lower case by 3DNA, can be approximated by using their normal forms. For +C, for example, 3DNA requires file `Atomic_c.pdb`, which can be simply a copy of `Atomic_C.pdb`.

- `Block_BP.alc` defines the default base-pair rectangular block. It has a size of 10 Å (long) by 4.5 Å (wide) by 0.5 Å (thick) and is in ALCHEMY format. It is used for drawing the Calladine-Drew style schematic presentation of DNA structures. `Block_R.alc` is for the purine base (R) ($4.5 \times 4.5 \times 0.5$). `Block_Y.alc` is for the pyrimidine base (Y) ($3.0 \times 4.5 \times 0.5$).

  Under its subdirectory `BLOCK`, there six block geometry files. `Block_M.alc` has half the size of `BLOCK_BP.alc`, and can be used if you would like the two blocks consisting a base-pair to be of the same size. `Block_Ms.alc` is slightly smaller in length than `BLOCK_M.alc` to avoid possible overlaps in a compressed base-pair (*i.e.* with negative Stretch). Furthermore, the blocks do not necessarily to be rectangular, as shown in `Block_R_nr.alc`.

- `Pxyz.dat` contains the *xyz* coordinates of phosphorus atoms with regard to the middle dinucleotide reference frame. Four sets were defined, corresponding respectively to average values in high resolution A- and B-DNA crystal structures, their intermediate and TA-DNA. New set can be added following the format. This file is used by `rebuild` for generating DNA structures with only base and phosphorus atoms. `PxyzH.dat` is the same as above except the coordinates are given in terms of the middle helical frame.

- `fig_image.par` contains parameters controlling the style of generated XFIG files, which can be edited by users to suit their liking. Similarly, `ps_image.par` holds parameters defining the drawing style for postscript images. Finally, `raster3d.par` and `my_header.r3d` are for Raster3D input.

- `misc_3dna.par` contains various parameters mainly for `analyze` and the utility program `find_pair`.

- `baselist.dat` contains a comprehensive list of currently known base residues and their standard counterpart. It makes analysis of unusual DNA and RNA structures straightforward.

- `trans_pep.pdb & trans_pep.alc` are trans peptide unit used for drawing peptide block in protein structures, in the same way as basepair blocks.

3.2 FIBER

This directory contains repeating unit for each type of the 55 fibers DNA and RNA structures. The original data as provided by Chandrasekaran & Arnott (1989) is given in subdirectory `Data`. Directories `Str01–Str55` store the "clean-up" version of each repeating unit in a format suitable for building the structure with utility program `fiber`.

3.3 bin

This directory contains executables of the 3DNA package. Most of which are utility programs with some in short Perl script. Detailed usage of each program is described in Section 4.

3.4 Examples

Four subdirectories are included to illustrate the various functionalities of the 3DNA package. You are strongly recommended to study these examples carefully in order to use 3DNA more effectively.

- `Analysis_Rebuild` contains the analysis/rebuilding results of four structures: `adh026` (A-DNA), `bdl084` (B-DNA), `pde0128` and `pd0001` (DNA-protein complexes). The `*.pdb` data files were downloaded from the `NDB`. The `*.inp` files are the corresponding input to the analysis routine (`analyze`) and `*.out` are the output containing various structural parameters.

  Input file `multi_str.inp` illustrates how to analyze multiple structure from one input file.

  `README` contains detailed information on how to run the `analyze/rebuild` programs to generate results in this directory, and the RMS deviations between 3DNA rebuilt structures and the experimental ones. For base atoms, the RMS is virtually zero, and with backbone atoms it less than 0.85Å even for the 146bp nucleosomal DNA.

- `Calladine_Drew` illustrates how to generate the two sets of DNA schematic pictures made popular by Calladine & Drew (1997) (Figures 1 and 2). The `README` file provides every detail. Note that the plots in these two figures are on the same scale and in the same orientation.

- `NMR` gives an example on how to analyze multiple NMR structures from the PDB. Follow the `README` file there.

- `Stacking` shows the procedures for generating "standardized" base-stacking diagrams (Figure 3). Check `README` file there.

- `Triplex` gives examples on how to analyze triplex and parallel duplex structures. Follow `README` file there for details.

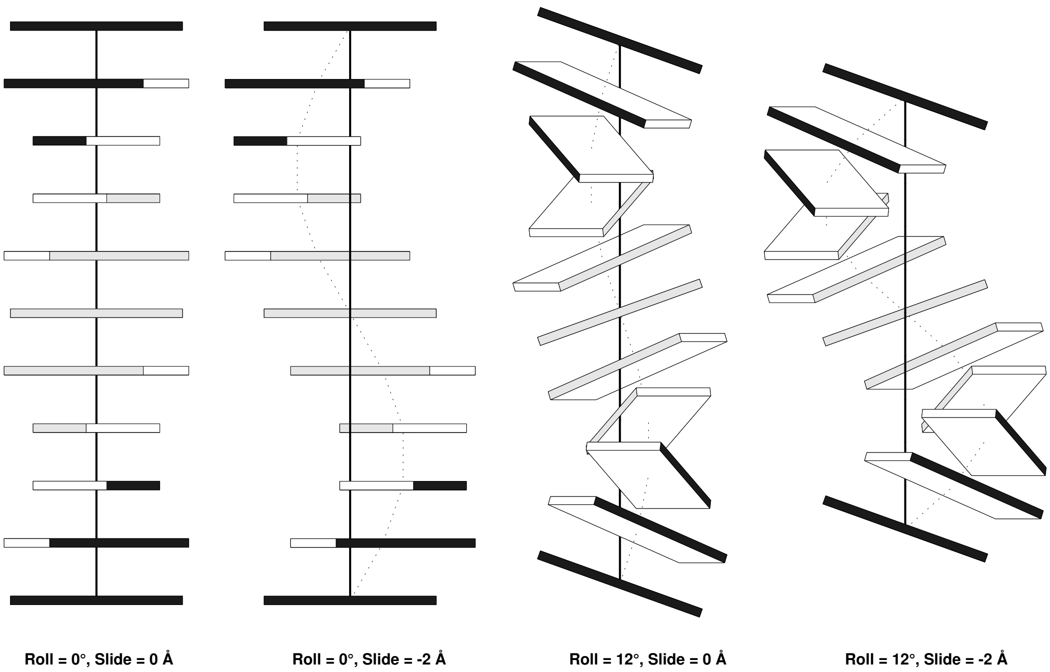


Figure 1: One complete *helical* turn of DNA having twist of 36°, showing the effects of introducing uniform roll and slide at each step (Calladine & Drew, 1997).

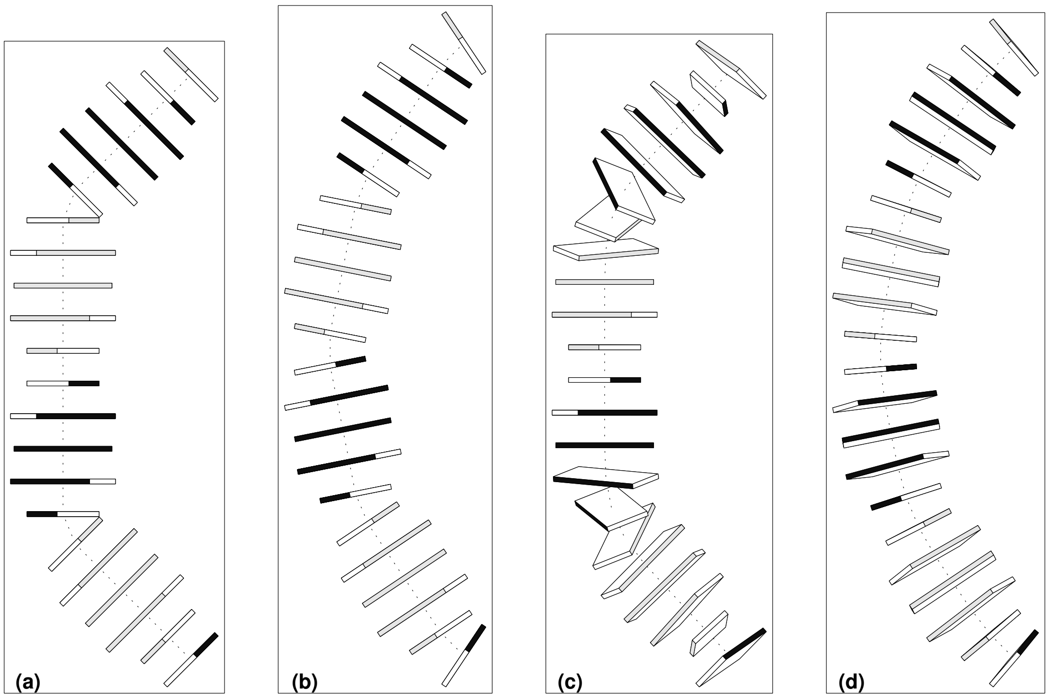


Figure 2: Two complete *helical* turns of DNA, with a curvature of 45° per turn, or 4.5° per step on average. Such tight curvature may be achieved, in principle by any of the distributions of roll angle shown in parts (a) to (d) (Calladine & Drew, 1997).

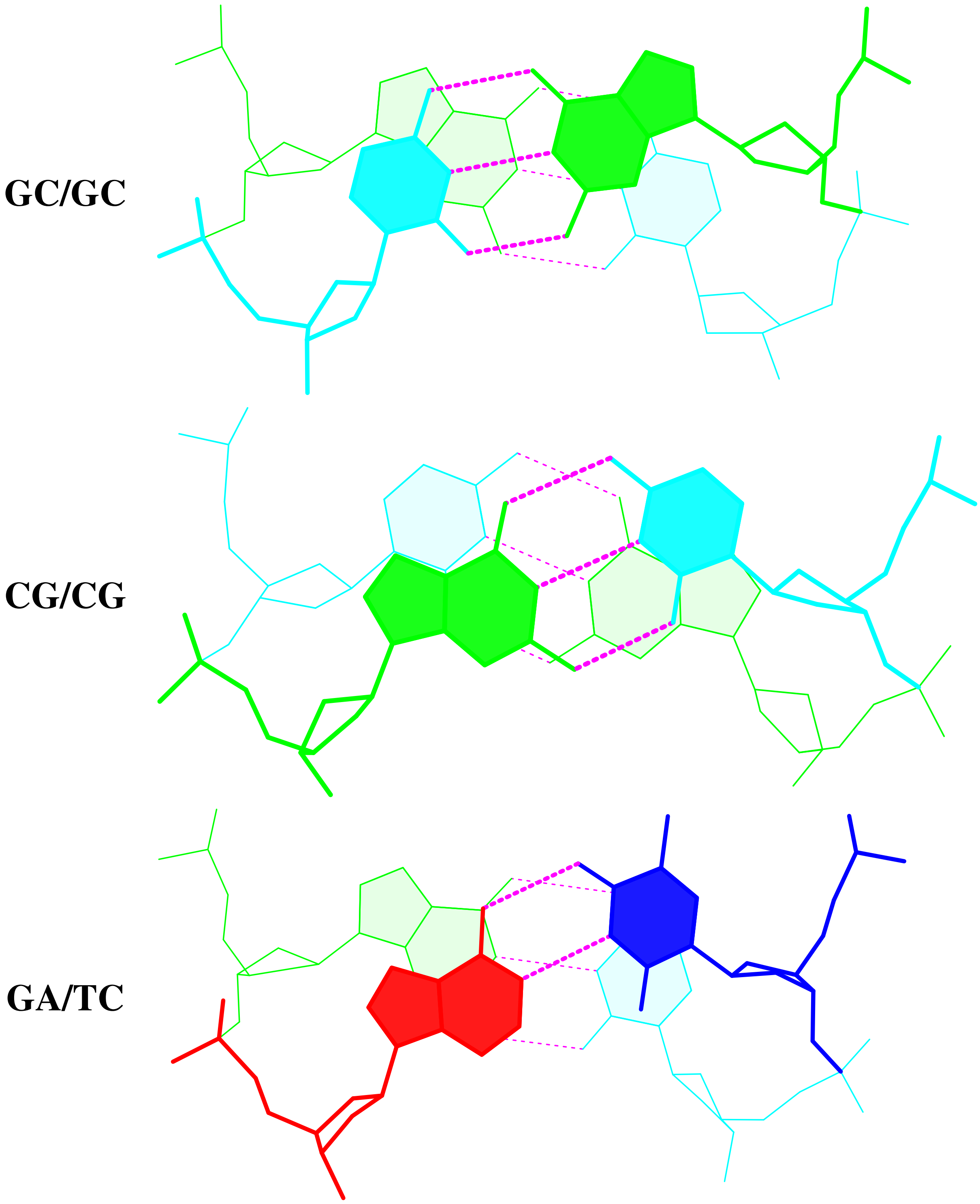


Figure 3: “Standardized” base stacking diagrams of three consecutive dimer steps of the 1.4 Å B-DNA structure d(C**GCGA**ATTCGCG)$_2$ (Shui *et al.*, 1998) (`BDL084`).

4 How to run 3DNA

Running 3DNA is simple although it might take a while to use it more effectively. For each of the program, there is a simple on-line help illustrating its usage.

1. Analyzing part

   - `analyze [inpfile1 inpfile2 ...]`
     Sample input files are given in directory `Examples/Analysis`.
   - `cehs [inpfile1 inpfile2 ...]`
     `cehs` gives the original `CEHS/SCHNAaP` parameters, to which `FreeHelix/NewHelix` ones should be quite similar. `cehs` is provide for completeness and comparison purpose.
   - `find_pair [options] pdbfile inpfile`
     `find_pair` is used to generate an input file for `analyze/cehs`, starting directly from a PDB file. It can also generate input files for the popular nucleic acid analysis program `Curves` with option `-c`.
   - `manalyze [-cehs] inpfile`
     A `Perl` script for analyzing multiple structures
   - `nmr_strs [-cehs] inpfile n1 n2`
     A `Perl` script for analyzing multiple NMR structures

2. Rebuilding part

   - `rebuild [options] [-negx] inpfile outfile`
     For rebuilding DNA structures of either atomic model in PDB format or schematic representation in ALCHEMY format.
   - `regular_dna [options] outfile`
     A utility program to generate input file for `rebuild` for the construction of regular DNA structures.
   - `frame_mol [options] -n1[,n2] reffile inpfile outfile`
     A utility program for attaching local helical frames and set the orientation of a structure.
   - `rotate_mol rotfile inpfile outfile`
     A utility program for adjusting the orientation of a structure by rotations in `MolScript` style. In RasMol, a structure can be rotated etc, but the corresponding new coordinates can *not* be saved. `rotate_mol` is provided simply for this purpose: Rotate your structure in RasMol to the orientation you like, then `write molscript rotfile`.
   - `alc2img [options] [-s=factor] inpfile outfile`
     Generating schematic representation of DNA from a ALCHEMY file.

- `pdb2img [options] [-s=factor] inpfile outfile`
  Generating schematic block representation of nucleic acid bases or protein peptide bonds directly from a PDB file.

3. Fiber models

- `fiber options pdbfile`
  The options is `-num`, where `num` is between 1 and 38. `-a|-b|-c|-d|-z` can be used for A-DNA, B-DNA, C-DNA, D-DNA and Z-DNA models.

4. Stacking diagrams

- `stack2img [options] [-s=factor] inpfile outfile`
  Program to generate "standardized" base-stacking diagram and atomic representations of nucleic acid structures with filled base rings and hydrogen bonds.
- `mstack2img ['options'] stackfile n1 n2 prefix`
  A `Perl` utility calling `stack2img` to generate stacking diagrams for a set of dinucleotide structures.

5. Set standard bases

- `std_base [options] [angle1 distance angle2] inpfile outfile`
  A utility program for setting standard base geometries used by `analyze` and `rebuild`.

6. Other utilities

- `r3d_atom`
  A utility program for generating scenes of atomic CPK, ball-and-stick, filled-base-ring, hydrogen bonds between bases, colored by atomic type, base residue or in gray, for input to Raster3D. Some of its functionalities are similar to `rods, balls` and `ring3d` of the Raster3D distribution, but with a more general approach for locating base rings.
- `cp_std prefix`
  A `Perl` script for changing standard base geometry files.
- `ex_str -num [-nmr] pdbfile`
  A utility program for extracting a specific structure unit or an NMR structure from a multiple-structure PDB file.
- `get_part [options] inpfile outfile`
  Extracting protein, nucleic acid or other components from a PDB file.
- `o1p_o2p inpfile [outfile]`
  Checking the designation of O1P and O2P atoms in a PDB file. Useful for RMS fittings.

- `block_atom`
  A simple Perl script for generating block plus atomic representation of a nucleic acid structure in ALCHEMY format, which can be displayed using RasMol. For example, with `block_atom bdl084` (assuming a PDB data file `bdl084.pdb` exists), one get a file `bdl084.alc`, which can be display with `rasmol -alchemy -noconnect bdl084.alc`. Note the undocumented option `-noconnect` is needed to avoid RasMol recalculating bond connections.

- `dcmnfile`
  A simple Perl script for clean-up. Various programs in 3DNA generate auxiliary files with fixed names which can be deleted with this utility program.

5 A tutorial introduction

This section shows some detailed examples on how to use 3DNA properly and effectively. Only some of the main features are covered for beginners to get started. As mentioned above, a serious user should go through the `Examples` directory provided with 3DNA and play with it.

The universal molecular graphics program `rasmol` (version 2.6+, but not 2.7.1) is indispensable. Also, `gv (ghostview)`, `xfig` (version 3.2+) and `Raster3D` (version 2.5+) are highly recommended. Most likely they might be already available on your machine, since they are all extremely valuable freewares.

5.1 Base-pair geometry parameters

Figures 4 and 5 illustrate the definition of various parameters describing the base-pair geometry in a nucleic acid structure. These two figures were themselves generated with the help of 3DNA.

5.2 Analysis of the Tc3 transposon DNA

The crystal structure of Tc3 transposase in complex with transposon DNA was solved by van Pouderoyan *et al.* (1997). This transposon DNA structure is unusual in that it contains an A/B junction: the G+C rich end adopts the A-DNA conformation while the A+T rich part has the B-DNA conformation. As a consequence, this DNA is non-linear, *i.e.*, curved.

The NDB reference code for this structure is `pde0128`. Download the structure and display it with `rasmol`, you will find that protein and DNA are together. To analyze the helical DNA structure, the program `(analyze)` needs to know which base-residue (sequential number in the PDB file) pairs with which. This information (in file `pde0128.inp`) can be obtained with the utility program `find_pair` as follows:

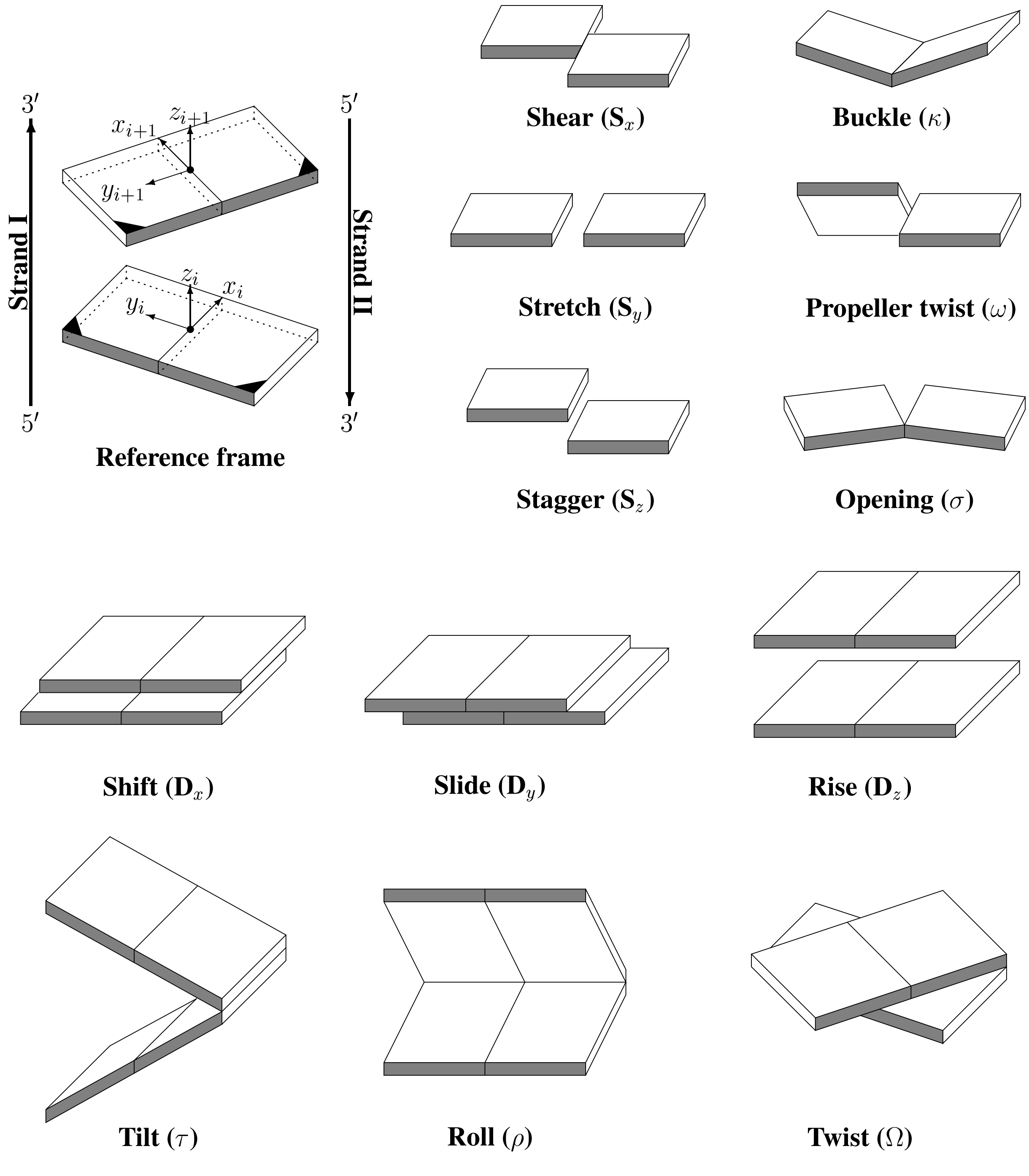


Figure 4: Pictorial definitions of parameters that relate complementary base pairs and sequential base-pair steps. The base-pair reference frame is constructed such that the *x*-axis points away from the (shaded) minor groove edge. Images illustrate positive values of the designated parameters (Dickerson *et al.*, 1989).

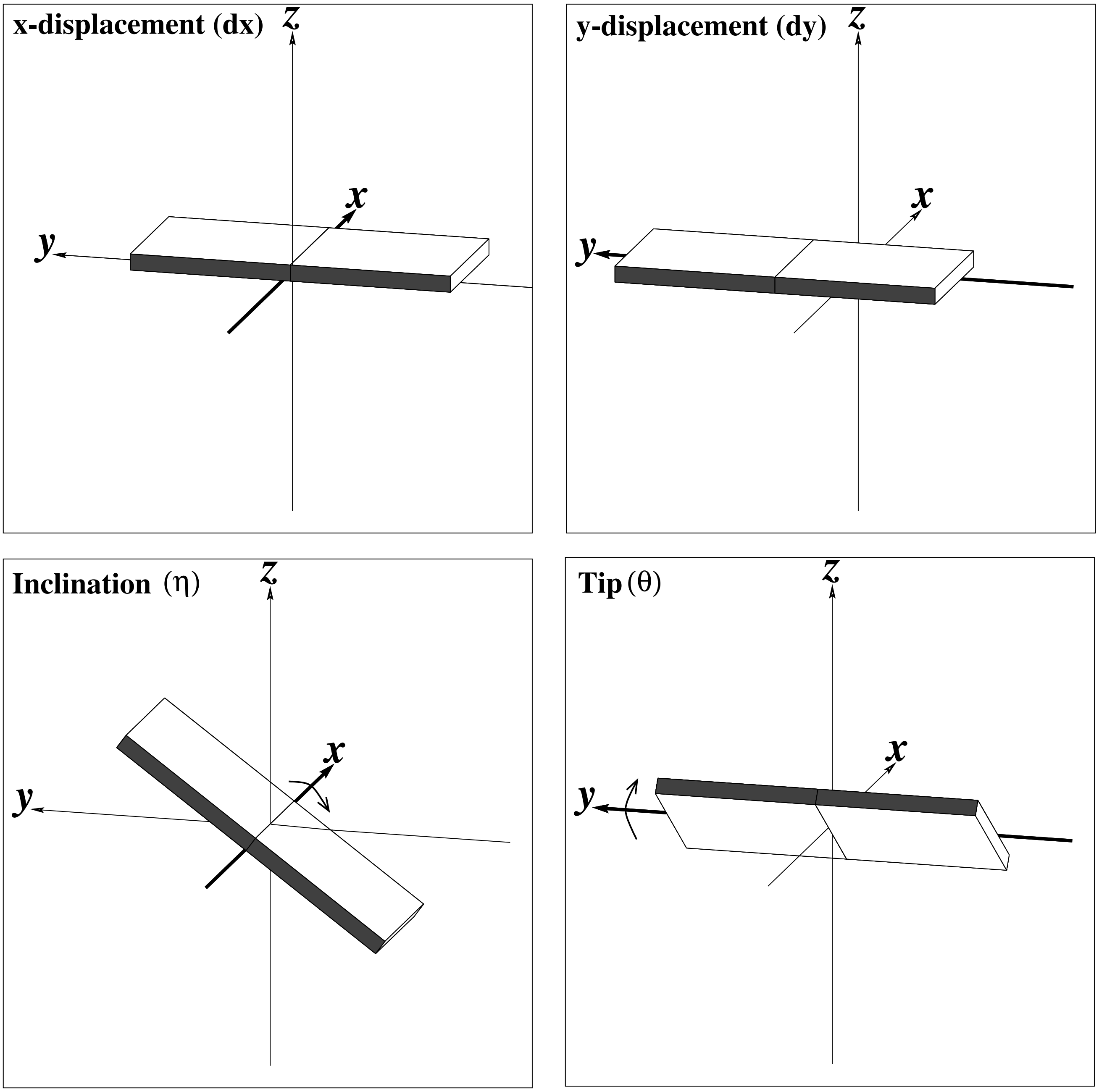


Figure 5: Helical parameters. Pictorial definitions of parameters that relate base pairs to its helical frame.

```
find_pair pde0128.pdb pde0128.inp
```

where `pde0128.pdb` is the data file in PDB format downloaded from the NDB and `pdeo128.inp` is the file required by `analyze`.

```
analyze pde0128.inp
```

The above two steps can be combined as follows:

```
find_pair pde0128.pdb stdout | analyze
```

Among several other output files, the one named `pde0128.out` (see page 20) gives a detailed listing of the DNA structural parameters. Each part is briefly summarized as follows:

- `RMS deviation` ··· gives the rms deviation between each experimental base and the standard reference. Normally the value is less than 0.05 Å. For each base, the corresponding PDB residue information is provided. A base-pair is classified as either Watson-Crick or non-Watson-Crick (denoted with a star).

- `H-bonding information` gives atom-list and their length in Å of all possible H-bonds in each of the base-pairs.

- `Overlap area` in Å^2 between polygons defined by atoms on successive bases. Polygons projected in the mean plane of the designed base-pair step.

  Values in parentheses measure the overlap of base ring atoms only. Those outside parentheses include exocyclic atoms on the ring.

- `Origin (Ox, Oy, Oz)` ··· gives the origins and mean base-pair normal vectors in the coordinate system of the given structure. Dickerson *et al.* has found it very useful by drawing the `Ny` vs. `Nx` normal plot to demonstrate the curvature of a DNA molecule.

- `Local base-pair parameters`, *i.e.*, Shear, Stretch, Stagger, Buckle, Propeller, and Opening (Figure 4). *Generally speaking*, Propeller always has a value of about $-10°$ in A- and B-DNA (Calladine & Drew, 1997), while Buckle can be either positive or negative of up to 20°. The other four show much less variations.

- `Local base-pair step parameters` gives the familiar Shift, Slide, Rise, Tilt, Roll, and Twist (Figure 4). Slide-Roll-Twist show more variations than the other three. Slide is the best single parameter among the six in discriminating A- from B-DNA.

- `Local base-pair helical parameters`, *i.e.*. X-displacement, Y-displacement, Helical rise, Inclination, Tip, and Helical twist, (Figure 5) are also commonly used. For canonical A-DNA, for example, the X-displacement characterizes the "hole" in the top view (Figure 6), and the Inclination means that the bases are non-perpendicular to the helical axis in the side view (see cover image).

- `lambda` $\cdots$ gives the virtual angle between C1'-YN1 or C1'-RN9 glycosidic bonds and the base-pair C1'-C1' line. For mismatched base-pairs, the two λ angles are normally quite different.

- `Classification of each dinucleotide step` $\cdots$ is based on the Zp and ZpH values. It applies to right-handed A-, B- and TA-DNA with Watson-Crick base-pair steps. This classification should be checked against the step/helical parameters and the backbone torsion angles etc for final assignment. Generally speaking, this classification can be taken as a good starting point for pinpointing possible structural transitions in large DNA structures. For `pde0128`, this section says that the transposon DNA is not simply a deformed B-DNA, but one with an A/B junction.

- `Minor and major groove widths` are calculated based on the method proposed by El Hassan & Calladine (1998). It uses simple cross-strand P–P distances with and without an angle correction, and assign the groove widths to each dinucleotide step. Please note that the sum of vdW radii (5.8 Å) of the two phosphate groups is **not** subtracted from the values given here.

- `Structure classification` gives the overall assignment of the structure as right-handed (A-, B-, TA-DNA), left-handed Z-DNA, and other two topologically possible forms.

- `Global linear helical axis` is defined by equivalent C1' and RN9/YN1 atom pairs as in `SCHNAaP`. It gives the unit vector along the helical axis and the two end points which can be used to locate the axis by adding these two points as `ATOM/HETATM` records in your PDB file and draw a line bewteen them. While such a linear-fitting makes no much sense for a strongly curved DNA structure, the deviation from regular linear helix is a good measure of the overall structural deformation.

- `Main chain and chi torsion angles.`

- `Sugar conformational parameters` include five internal torsional angles, and amplitude and phase angle of pseudorotation of the sugar ring.

- `Same strand` P–P and C1'–C1' virtual bond distances have been used in the literature to discriminate A- and B-DNA conformations, although it is now clear that they are not as good as Zp and Slide.

- `Helix radius` gives the radial displacement of P, O4', and C1' atoms in local helix frame of each dimer.
- `Position (Px, Py, Pz)` and local helical axis vector (Hx, Hy, Hz) for each dinucleotide step.

The output file `bp_step.par` contains information for a rigorous reconstruction of the base-pair geometry. It will be overwritten each time when another analysis is performed unless it is renamed.

```
  19 base-pairs
   0  ***local base-pair & step parameters***
        Shear  Stretch  Stagger Buckle Prop-Tw Opening   Shift  Slide    Rise    Tilt    Roll   Twist
G-C     -0.18   -0.27    0.58    8.57   -9.77   -8.26    0.00   0.00    0.00    0.00    0.00    0.00
G-C      0.16   -0.26    0.13   -5.72  -10.04   -3.52    0.19  -1.13    3.65    5.58    3.88   39.73
G-C     -0.10   -0.25    0.42   -1.54   -7.63   -0.67    0.43  -1.67    2.94   -3.14   10.32   29.99
G-C      0.02   -0.33    0.08   -3.56   -6.03    1.74    0.17  -2.24    3.13    1.31    6.72   31.31
G-C     -0.21   -0.24   -0.13   -3.78   -0.69    3.13    0.08  -1.92    3.33    3.67    8.74   27.62
G-C      0.04   -0.22   -0.35   -8.83  -11.08   -3.03    0.50  -1.91    3.41    4.23    5.60   31.56
G-C     -0.32   -0.40   -0.38  -16.06  -17.27   -3.53    0.18  -2.14    3.43    4.68   10.29   27.30
T-A      0.43   -0.40   -0.05  -11.97  -14.80   12.03   -0.07  -0.42    3.15   -0.69    5.91   35.56
C-G      0.52   -0.43   -0.34  -10.52    4.23    3.87    0.25   0.76    3.46    3.31    0.88   37.83
C-G      0.27   -0.33    0.52  -15.53   -4.19   -2.52   -1.73   0.69    3.77   -7.18   -1.10   31.39
T-A     -0.18   -0.23   -0.13   -9.98   -5.91   -1.56    0.35  -0.32    3.19    5.36    6.16   26.98
A-T     -0.57   -0.49    0.47    0.67   -7.27   -0.77    0.41   1.89    3.06   -6.05   -2.74   45.60
T-A     -0.34   -0.35    0.33   -3.99  -14.34   -5.83   -0.34  -0.18    3.44    1.73   -3.42   31.69
A-T     -0.69   -0.02   -0.04   -2.98   -0.45   12.91    0.29   2.18    3.37   -2.75   -6.21   49.54
G-C     -0.51   -0.07   -0.61    3.47   -9.27   12.05    0.33   1.29    3.25   -0.15   15.17   23.41
A-T     -0.66   -0.15   -0.12   12.32  -13.66    3.93   -0.68   0.30    2.96   -4.96    0.68   35.87
A-T     -0.41   -0.38    0.73   21.10  -23.76   -1.56   -0.06  -0.03    2.97   -6.65   -0.95   34.68
C-G      0.40   -0.55    0.33   -5.21  -10.24   -3.28    0.37  -1.03    3.93    4.89    1.28   38.27
T-A      0.56   -0.49    0.46  -21.17   -8.34   -4.72   -0.19  -1.07    3.61    1.91    4.78   37.16
```

Let's rebuild an atomic structure with the following command:

```
rebuild -atomic bp_step.par tc3_base.pdb
```

Depending on your setting of the standard base geometry (*i.e.*. `Atomic_?.pdb` files), you will get a structure with either only base atoms (the default) or with an approximate sugar-phosphate backbone attached. Use `rasmol` to have a look.

Your can also rebuild a schematic Calladine-Drew style picture with the following commands:

`rebuild bp_step.par tc3_bp1.alc` [for one block per *base-pair*]

`rebuild -block2 bp_step.par tc3_bp2.alc` [for one block per base]

You need to use `rasmol` with the `-alchemy` command-line option to display files `tc3_bp1.alc` and `tc3_bp2.alc` since they are in `ALCHEMY` format. *e.g.*,

```
rasmol -alchemy tc3_bp2.alc
```

With this image, the Buckle and Propeller deformations are immediately obvious.

5.3 Fiber models

The 55 types of fiber nucleic acid models by Chandrasekaran & Arnott (1989) can be easily generated with the program `fiber`. Use `fiber -m` to get a list of all structures. Here I will use calf thymus A-DNA (number 1 in the list) as an example to illustrate its usage.

```
fiber -a fiber_A.pdb
```

Here `-a` is the same as `-1` and `fiber_A.pdb` is the output file name. Your will then be asked to input your base sequence. It could either be from a data file (complete sequence) or from keyboard (enter only the repeating sequence, which is the default). Type `enter` means the default for input from keyboard. You are prompted for repeating unit with a default for polyA. Type `atcg` (either case is Okay and uncommon bases will be ignored) for a mixed A-T-C-G repeating sequence. Finally you are asked for the number of repeats (default is 10). Type 5 so you get $(ATCG)_5$, *i.e.*, 20 base-pairs. Display it with `rasmol fiber_A.pdb`.

High quality postscript picture (Figure 6) of `fiber_A.pdb` can be generated with the utility `stack2img` as follows:

```
stack2img -cao fiber_A.pdb fiber_A.ps
```

The option `-cao` means color-coded (`c`), atomic-model (`a`) with filled base-rings (`o`). If you add the option `-f`, you will get an image in XFIG format which you can easily edit. This is actually what these *atomic* options are intended for.

5.4 Input for **`render`** in Raster3D etc

The Calladine-Drew style base-pair representations are themselves quite useful as shown in Figures 1 and 2. They are, however, even more helpful when combined with atomic and schematic representations of ligands and proteins etc. This can be achieved with the `render` program of Raster3D as follows:

Use RasMol to find the view you want, and `write molscript molfile` at the command window. Then use the utility program `rotate_mol` to get a new PDB file with coordinates corresponding to your chosen view (RasMol does not write back coordinates in a new view). The new PDB coordinates are the *common reference* for scenes generated by different programs (3DNA, MolScript, Raster3D etc) to be properly `render`ed by Raster3D. Specifically, avoid any further coordinates transformation by MolScript or Raster3D.

You might use the utility `get_part` to divide a structure into nucleic acid part (default or -n), protein part (-p) and others (-o) to be used with different programs. For example, use 3DNA utilities `r3d_atom` and `pdb2img/stack2img` for nucleic acids, and MolScript for the protein part. You might need to delete the header part from MolScript with `del_ms -n`, a simple Perl script comes with

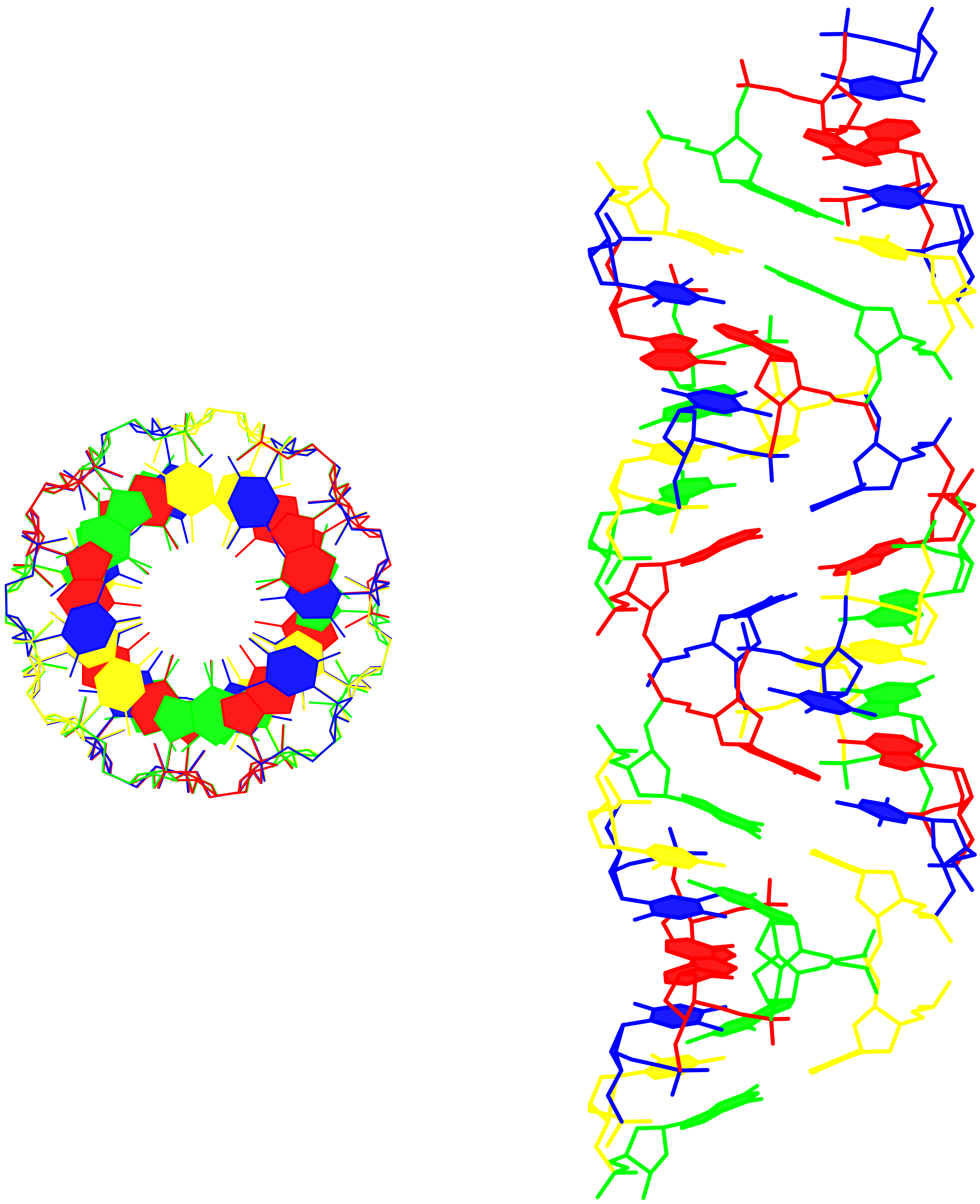


Figure 6: Top and side view images of the 20 base-pair long, $(ATCG)_5$, fiber A-DNA color coded by residue: A-red, T-blue, G-green, and C-yellow.

3DNA. The different scenes can then be concatenated together for `render` as detailed in Raster3D document.

Most of the images in 3DNA homepage were generated this way.

5.5 Build a DNA structure with sugar-phosphate backbone

To be completed.

Please refer to README file in directory `Examples/Analyze_Rebuild` for details.

6 Citation

Xiang-Jun Lu, Zippora Shakked & Wilma K. Olson (2000). “A-DNA Conformational Motifs in Ligand-bound Double Helices.” *J. Mol. Biol.* **300(4)**, 819-840.

7 Acknowledgments

The development of 3DNA benefit greatly from extensive testing and enthusiastic feedbacks of an increasing user community. In particular, we would like to thank Andrew Colasanti, Haim Rozenberg, A. R. Srinivasan, Patrick Furrer, Karolin Luger and Zukang Feng for validating the program, and Zippi Shakked for constructive suggestions.

This work has been generously supported by the U. S. Public Health Service under research grant GM20861. Computations were carried out at the Rutgers University Center for Computational Chemistry and through the facilities of the Nucleic Acid Database project (NSF grant DBI 9510703).

8 Appendix: 3DNA output for Tc3 transposon DNA

Detailed structural parameters for the analysis of Tc3 transposon DNA (data file `pde0128.out`)

```
****************************************************************************
****************************************************************************
    3DNA (v1.5, Nov. 2002) by Xiang-Jun Lu at Wilma K. Olson's Lab.
****************************************************************************
1. The list of the parameters given below correspond to the 5' to 3' direction
   of strand I and 3' to 5' direction of strand II.

2. All angular parameters, except for the phase angle of sugar pseudo-
   rotation, are measured in degrees in the range of [-180, +180], and all
   displacements are measured in Angstrom units.
****************************************************************************
File name: pde0128.pdb
Date and time: Tue Nov 12 00:26:08 2002

Number of base-pairs: 19
Number of atoms: 1287
****************************************************************************
HEADER    PROTEIN/DNA                             07-JUL-97   1TC3
TITLE     TRANSPOSASE TC3A1-65 FROM CAENORHABDITIS ELEGANS
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: TC3 TRANSPOSASE;
COMPND   3 CHAIN: C;
COMPND   4 FRAGMENT: SPECIFIC DNA BINDING DOMAIN, RESIDUES 2 - 52;
COMPND   5 ENGINEERED: YES;
COMPND   6 MUTATION: C-TERMINAL 6-HIS TAG;
COMPND   7 MOL_ID: 2;
COMPND   8 MOLECULE: DNA
COMPND   9 (5'-D(*AP*GP*GP*GP*GP*GP*GP*GP*TP*CP*CP*TP*AP*TP*AP*GP*A
COMPND  10 P*AP*CP*TP*T)-3');
COMPND  11 CHAIN: A;
COMPND  12 ENGINEERED: YES;
COMPND  13 MOL_ID: 3;
COMPND  14 MOLECULE: DNA
COMPND  15 (5'-D(*AP*GP*TP*TP*CP*TP*AP*TP*AP*GP*GP*AP*CP*CP*CP*CP*C
COMPND  16 P*CP*CP*T)-3');
COMPND  17 CHAIN: B;
COMPND  18 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: CAENORHABDITIS ELEGANS;
SOURCE   3 STRAIN: BERGERAC;
SOURCE   4 VARIANT: TR679;
```

```
SOURCE   5 ORGANELLE: NUCLEUS;
SOURCE   6 GENE: TC3A;
SOURCE   7 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   8 EXPRESSION_SYSTEM_STRAIN: BL21 (DE3) PLYSS;
SOURCE   9 EXPRESSION_SYSTEM_CELLULAR_LOCATION: CYTOPLASM;
SOURCE  10 EXPRESSION_SYSTEM_VECTOR_TYPE: PET3C;
SOURCE  11 EXPRESSION_SYSTEM_PLASMID: PRP1200;
SOURCE  12 EXPRESSION_SYSTEM_GENE: TC3A N1-65;
SOURCE  13 MOL_ID: 2;
SOURCE  14 SYNTHETIC: YES;
SOURCE  15 MOL_ID: 3;
SOURCE  16 SYNTHETIC: YES
KEYWDS    TRANSPOSASE, DNA BINDING, HELIX-TURN-HELIX, TC1/MARINER
KEYWDS   2 FAMILY, COMPLEX (TRANSPOSASE/DNA), PROTEIN/DNA
EXPDTA    X-RAY DIFFRACTION
AUTHOR    G.VAN POUDEROYEN,R.F.KETTING,A.PERRAKIS,R.H.A.PLASTERK,
AUTHOR   2 T.K.SIXMA
REVDAT   1   12-NOV-97         0
JRNL        AUTH   G.VAN POUDEROYEN,R.F.KETTING,A.PERRAKIS,
JRNL        AUTH 2 R.H.A.PLASTERK,T.K.SIXMA
JRNL        TITL   CRYSTAL STRUCTURE OF THE SPECIFIC DNA-BINDING
JRNL        TITL 2 DOMAIN OF TC3 TRANSPOSASE OF C. ELEGANS IN COMPLEX
JRNL        TITL 3 WITH TRANSPOSON DNA
JRNL        REF    EMBO J.                      V.  16  6044 1997
JRNL        REFN   ASTM EMJODG  UK ISSN 0261-4189                0897
HELIX    1   1 ASP C  209  LEU C  220  1
HELIX    2   2 LEU C  225  ILE C  232  1
HELIX    3   3 ARG C  236  LYS C  244  1
****************************************************************************
RMSD of the bases (----- for WC bp, + for isolated bp, x for helix change)

            Strand I                    Strand II          Helix
   1   (0.025) A:...2_:[..G]G-----C[..C]:.119_:B (0.016)     |
   2   (0.028) A:...3_:[..G]G-----C[..C]:.118_:B (0.019)     |
   3   (0.022) A:...4_:[..G]G-----C[..C]:.117_:B (0.014)     |
   4   (0.017) A:...5_:[..G]G-----C[..C]:.116_:B (0.017)     |
   5   (0.027) A:...6_:[..G]G-----C[..C]:.115_:B (0.020)     |
   6   (0.020) A:...7_:[..G]G-----C[..C]:.114_:B (0.014)     |
   7   (0.034) A:...8_:[..G]G-----C[..C]:.113_:B (0.019)     |
   8   (0.019) A:...9_:[..T]T-----A[..A]:.112_:B (0.020)     |
   9   (0.009) A:..10_:[..C]C-----G[..G]:.111_:B (0.017)     |
  10   (0.010) A:..11_:[..C]C-----G[..G]:.110_:B (0.034)     |
  11   (0.023) A:..12_:[..T]T-----A[..A]:.109_:B (0.016)     |
  12   (0.023) A:..13_:[..A]A-----T[..T]:.108_:B (0.021)     |
  13   (0.018) A:..14_:[..T]T-----A[..A]:.107_:B (0.024)     |
```

```
  14   (0.026) A:..15_:[..A]A-----T[..T]:.106_:B (0.018)     |
  15   (0.017) A:..16_:[..G]G-----C[..C]:.105_:B (0.025)     |
  16   (0.033) A:..17_:[..A]A-----T[..T]:.104_:B (0.019)     |
  17   (0.031) A:..18_:[..A]A-----T[..T]:.103_:B (0.017)     |
  18   (0.012) A:..19_:[..C]C-----G[..G]:.102_:B (0.026)     |
  19   (0.021) A:..20_:[..T]T-----A[..A]:.101_:B (0.023)     |
****************************************************************************
Detailed H-bond information: atom-name pair and length [ON]
   1 G-----C  [3]  O6 - N4  2.53  N1 - N3  2.85  N2 - O2  3.00
   2 G-----C  [3]  O6 - N4  2.61  N1 - N3  2.80  N2 - O2  2.77
   3 G-----C  [3]  O6 - N4  2.86  N1 - N3  2.80  N2 - O2  2.62
   4 G-----C  [3]  O6 - N4  2.72  N1 - N3  2.64  N2 - O2  2.40
   5 G-----C  [3]  O6 - N4  2.89  N1 - N3  2.76  N2 - O2  2.59
   6 G-----C  [3]  O6 - N4  2.64  N1 - N3  2.84  N2 - O2  2.87
   7 G-----C  [3]  O6 - N4  2.57  N1 - N3  2.67  N2 - O2  2.83
   8 T-----A  [2]  N3 - N1  2.51  O4 - N6  2.90
   9 C-----G  [3]  O2 - N2  2.38  N3 - N1  2.64  N4 - O6  2.83
  10 C-----G  [3]  O2 - N2  2.66  N3 - N1  2.74  N4 - O6  2.78
  11 T-----A  [2]  N3 - N1  2.74  O4 - N6  2.77
  12 A-----T  [2]  N6 - O4  2.67  N1 - N3  2.61
  13 T-----A  [2]  N3 - N1  2.60  O4 - N6  2.60
  14 A-----T  [2]  N6 - O4  3.52  N1 - N3  2.86
  15 G-----C  [3]  O6 - N4  3.43  N1 - N3  2.94  N2 - O2  2.60
  16 A-----T  [2]  N6 - O4  3.13  N1 - N3  2.91
  17 A-----T  [2]  N6 - O4  2.91  N1 - N3  2.63
  18 C-----G  [3]  O2 - N2  2.47  N3 - N1  2.56  N4 - O6  2.57
  19 T-----A  [2]  N3 - N1  2.65  O4 - N6  2.66
****************************************************************************
Overlap area in Angstrom^2 between polygons defined by atoms on successive
bases. Polygons projected in the mean plane of the designed base-pair step.

Values in parentheses measure the overlap of base ring atoms only. Those
outside parentheses include exocyclic atoms on the ring. Intra- and
inter-strand overlap is designated according to the following diagram:

                    i2  3'       5'  j2
                       /|\       |
                        |        |
               Strand I |        | II
                        |        |
                        |        |
                        |       \|/
                    i1  5'       3'  j1
```

step		i1-i2	i1-j2	j1-i2	j1-j2	sum
1	GG/CC	4.01(2.20)	0.00(0.00)	0.03(0.00)	0.81(0.00)	4.85(2.20)
2	GG/CC	3.97(2.70)	0.00(0.00)	0.19(0.00)	0.00(0.00)	4.16(2.70)
3	GG/CC	3.18(1.65)	0.00(0.00)	1.16(0.00)	0.00(0.00)	4.34(1.65)
4	GG/CC	3.36(1.96)	0.00(0.00)	0.56(0.00)	0.00(0.00)	3.93(1.96)
5	GG/CC	4.26(2.90)	0.00(0.00)	0.43(0.00)	0.00(0.00)	4.69(2.90)
6	GG/CC	3.30(1.83)	0.00(0.00)	0.61(0.00)	0.00(0.00)	3.91(1.83)
7	GT/AC	6.08(1.72)	0.00(0.00)	0.00(0.00)	4.79(3.35)	10.87(5.06)
8	TC/GA	5.10(0.61)	0.00(0.00)	0.00(0.00)	4.69(1.63)	9.78(2.25)
9	CC/GG	1.18(0.00)	0.00(0.00)	0.00(0.00)	8.23(5.73)	9.41(5.73)
10	CT/AG	7.29(1.31)	0.00(0.00)	0.00(0.00)	2.86(1.70)	10.14(3.01)
11	TA/TA	2.56(0.00)	0.00(0.00)	0.00(0.00)	1.53(0.00)	4.08(0.00)
12	AT/AT	5.73(1.71)	0.00(0.00)	0.00(0.00)	6.31(2.41)	12.05(4.12)
13	TA/TA	1.76(0.00)	0.00(0.00)	0.00(0.00)	2.27(0.00)	4.03(0.00)
14	AG/CT	6.78(4.02)	0.00(0.00)	0.00(0.00)	4.40(0.27)	11.18(4.29)
15	GA/TC	3.83(0.87)	0.00(0.00)	0.00(0.00)	4.69(0.53)	8.52(1.40)
16	AA/TT	4.06(2.70)	0.00(0.00)	0.00(0.00)	5.91(0.32)	9.97(3.02)
17	AC/GT	5.32(4.25)	0.00(0.00)	0.00(0.00)	7.39(2.57)	12.70(6.81)
18	CT/AG	4.16(0.08)	0.00(0.00)	0.18(0.00)	3.04(2.60)	7.39(2.68)

****

Origin (Ox, Oy, Oz) and mean normal vector (Nx, Ny, Nz) of each base-pair in the coordinate system of the given structure

bp		Ox	Oy	Oz	Nx	Ny	Nz
1	G-C	-0.32	155.89	29.45	0.62	-0.76	-0.20
2	G-C	2.94	154.00	28.78	0.70	-0.69	-0.15
3	G-C	5.94	152.64	29.63	0.62	-0.79	-0.01
4	G-C	8.66	150.82	31.66	0.55	-0.83	0.07
5	G-C	10.24	148.16	33.94	0.42	-0.90	0.15
6	G-C	10.19	144.79	35.97	0.30	-0.94	0.17
7	G-C	9.09	141.17	37.41	0.14	-0.99	0.07
8	T-A	9.15	137.99	37.55	0.15	-0.99	-0.03
9	C-G	10.37	134.67	37.58	0.12	-0.99	-0.08
10	C-G	9.49	130.73	38.76	0.14	-0.99	0.04
11	T-A	10.52	127.69	38.52	0.26	-0.97	-0.03
12	A-T	10.13	124.46	40.11	0.15	-0.99	0.01
13	T-A	10.83	121.09	39.68	0.18	-0.98	-0.05
14	A-T	9.50	117.55	38.32	0.19	-0.97	-0.17
15	G-C	8.98	114.42	36.82	-0.05	-1.00	-0.08
16	A-T	9.55	111.47	36.25	-0.05	-0.98	-0.17
17	A-T	9.56	108.56	35.71	0.04	-0.97	-0.24
18	C-G	8.48	104.79	34.55	-0.05	-0.97	-0.25
19	T-A	7.31	101.48	33.18	-0.03	-0.94	-0.33

****

Local base-pair parameters

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	-0.18	-0.27	0.58	8.57	-9.77	-8.26
2 G-C	0.16	-0.26	0.13	-5.72	-10.04	-3.52
3 G-C	-0.10	-0.25	0.42	-1.54	-7.63	-0.67
4 G-C	0.02	-0.33	0.08	-3.56	-6.03	1.74
5 G-C	-0.21	-0.24	-0.13	-3.78	-0.69	3.13
6 G-C	0.04	-0.22	-0.35	-8.83	-11.08	-3.03
7 G-C	-0.32	-0.40	-0.38	-16.06	-17.27	-3.53
8 T-A	0.43	-0.40	-0.05	-11.97	-14.80	12.03
9 C-G	0.52	-0.43	-0.34	-10.52	4.23	3.87
10 C-G	0.27	-0.33	0.52	-15.53	-4.19	-2.52
11 T-A	-0.18	-0.23	-0.13	-9.98	-5.91	-1.56
12 A-T	-0.57	-0.49	0.47	0.67	-7.27	-0.77
13 T-A	-0.34	-0.35	0.33	-3.99	-14.34	-5.83
14 A-T	-0.69	-0.02	-0.04	-2.98	-0.45	12.91
15 G-C	-0.51	-0.07	-0.61	3.47	-9.27	12.05
16 A-T	-0.66	-0.15	-0.12	12.32	-13.66	3.93
17 A-T	-0.41	-0.38	0.73	21.10	-23.76	-1.56
18 C-G	0.40	-0.55	0.33	-5.21	-10.24	-3.28
19 T-A	0.56	-0.49	0.46	-21.17	-8.34	-4.72
ave.	-0.09	-0.31	0.10	-3.93	-8.97	0.55
s.d.	0.40	0.14	0.38	10.22	6.42	6.13

****

Local base-pair step parameters

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GG/CC	0.19	-1.13	3.65	5.58	3.88	39.73
2 GG/CC	0.43	-1.67	2.94	-3.14	10.32	29.99
3 GG/CC	0.17	-2.24	3.13	1.31	6.72	31.31
4 GG/CC	0.08	-1.92	3.33	3.67	8.74	27.62
5 GG/CC	0.50	-1.91	3.41	4.23	5.60	31.56
6 GG/CC	0.18	-2.14	3.43	4.68	10.29	27.30
7 GT/AC	-0.07	-0.42	3.15	-0.69	5.91	35.56
8 TC/GA	0.25	0.76	3.46	3.31	0.88	37.83
9 CC/GG	-1.73	0.69	3.77	-7.18	-1.10	31.39
10 CT/AG	0.35	-0.32	3.19	5.36	6.16	26.98
11 TA/TA	0.41	1.89	3.06	-6.05	-2.74	45.60
12 AT/AT	-0.34	-0.18	3.44	1.73	-3.42	31.69
13 TA/TA	0.29	2.18	3.37	-2.75	-6.21	49.54
14 AG/CT	0.33	1.29	3.25	-0.15	15.17	23.41
15 GA/TC	-0.68	0.30	2.96	-4.96	0.68	35.87
16 AA/TT	-0.06	-0.03	2.97	-6.65	-0.95	34.68
17 AC/GT	0.37	-1.03	3.93	4.89	1.28	38.27
18 CT/AG	-0.19	-1.07	3.61	1.91	4.78	37.16

ave.	0.03	-0.38	3.34	0.28	3.67	34.19
s.d.	0.53	1.37	0.28	4.40	5.57	6.62

****

Local base-pair helical parameters

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 GG/CC	-2.13	0.43	3.53	5.66	-8.14	40.28
2 GG/CC	-4.56	-1.25	2.20	19.17	5.83	31.83
3 GG/CC	-5.13	-0.10	2.61	12.26	-2.38	32.03
4 GG/CC	-5.58	0.59	2.60	17.66	-7.41	29.17
5 GG/CC	-4.42	-0.14	3.07	10.15	-7.65	32.31
6 GG/CC	-6.26	0.59	2.47	20.70	-9.43	29.51
7 GT/AC	-1.48	0.03	3.05	9.60	1.12	36.04
8 TC/GA	1.05	0.06	3.48	1.35	-5.09	37.98
9 CC/GG	1.48	1.57	4.03	-2.00	13.05	32.20
10 CT/AG	-2.07	0.50	3.05	12.84	-11.16	28.17
11 TA/TA	2.64	-1.00	2.88	-3.51	7.75	46.05
12 AT/AT	0.32	0.95	3.42	-6.23	-3.15	31.91
13 TA/TA	3.03	-0.54	3.07	-7.37	3.26	49.97
14 AG/CT	-1.36	-0.72	3.43	33.28	0.33	27.84
15 GA/TC	0.40	0.46	3.03	1.09	8.01	36.20
16 AA/TT	0.09	-0.80	2.93	-1.57	11.03	35.30
17 AC/GT	-1.75	0.16	3.92	1.95	-7.42	38.59
18 CT/AG	-2.34	0.56	3.44	7.45	-2.98	37.50
	~~	~~~~~~~~~~	~~~~~~~~~~	~~~~~~~~~~	~~~~~~~~~~	~~~~~~~~~~
ave.	-1.56	0.07	3.12	7.36	-0.80	35.16
s.d.	2.81	0.73	0.48	10.73	7.47	5.99

****

Structure classification:

This is a right-handed nucleic acid structure

****

lambda: virtual angle between C1'-YN1 or C1'-RN9 glycosidic bonds and the
base-pair C1'-C1' line

C1'-C1': distance between C1' atoms for each base-pair
RN9-YN1: distance between RN9-YN1 atoms for each base-pair
RC8-YC6: distance between RC8-YC6 atoms for each base-pair

bp	lambda(I)	lambda(II)	C1'-C1'	RN9-YN1	RC8-YC6
1 G-C	51.0	58.7	10.7	9.0	9.8
2 G-C	58.2	48.2	10.6	8.9	9.7
3 G-C	62.1	54.4	10.3	8.8	9.8
4 G-C	51.7	52.9	10.5	8.7	9.7
5 G-C	51.6	55.9	10.4	8.7	9.8
6 G-C	54.2	60.3	10.5	8.9	9.8

7 G-C	62.5	57.0	10.2	8.7	9.5
8 T-A	57.3	53.0	9.9	8.3	9.5
9 C-G	52.6	54.7	10.3	8.6	9.6
10 C-G	49.3	52.7	10.6	8.8	9.6
11 T-A	56.0	57.1	10.4	8.8	9.7
12 A-T	54.9	59.9	10.1	8.6	9.5
13 T-A	55.0	50.7	10.5	8.8	9.7
14 A-T	49.2	62.3	10.4	8.7	9.9
15 G-C	59.0	60.5	10.2	8.8	9.9
16 A-T	53.9	62.2	10.3	8.8	9.8
17 A-T	45.7	51.3	10.5	8.5	9.5
18 C-G	57.3	57.8	10.2	8.7	9.5
19 T-A	46.6	45.9	10.4	8.5	9.3

****

Classification of each dinucleotide step in a right-handed nucleic acid structure: A-like; B-like; TA-like; intermediate of A and B, or other cases

step	Xp	Yp	Zp	XpH	YpH	ZpH	Form
1 GG/CC	-2.03	8.22	1.96	-4.14	8.00	2.77	A
2 GG/CC	-2.00	8.35	2.44	-6.36	7.15	4.91	A
3 GG/CC	-1.96	8.21	2.78	-6.89	7.47	4.41	A
4 GG/CC	-1.61	8.45	2.50	-6.98	7.34	4.87	A
5 GG/CC	-1.44	8.13	2.57	-5.65	7.57	3.94	A
6 GG/CC	-2.30	8.54	1.72	-8.31	7.41	4.67	A
7 GT/AC	-3.10	8.76	0.52	-4.51	8.57	1.89	
8 TC/GA	-3.37	8.71	-0.37	-2.26	8.72	-0.19	B
9 CC/GG	-2.98	8.85	-0.08	-1.58	8.84	-0.56	B
10 CT/AG	-3.21	8.87	-0.51	-4.98	8.79	1.29	B
11 TA/TA	-2.15	8.08	-0.04	0.28	8.06	-0.45	B
12 AT/AT	-3.22	8.85	-0.34	-2.89	8.76	-1.29	B
13 TA/TA	-2.31	8.32	-0.37	0.44	8.22	-1.32	B
14 AG/CT	-3.69	8.49	-0.96	-5.02	7.68	3.76	B
15 GA/TC	-3.87	8.83	-1.10	-3.34	8.85	-0.88	B
16 AA/TT	-3.52	8.81	0.03	-3.19	8.81	-0.12	B
17 AC/GT	-3.62	9.14	0.44	-5.27	9.12	0.70	B
18 CT/AG	-2.81	9.09	0.83	-5.00	8.91	1.98	

****

Minor and major groove widths: direct P-P distances and refined P-P distances which take into account the directions of the sugar-phosphate backbones

(Subtract 5.8 Angstrom from the values to take account of the vdw radii of the phosphate groups, and for comparison with FreeHelix and Curves.)

Ref: M. A. El Hassan and C. R. Calladine (1998). ``Two Distinct Modes of Protein-induced Bending in DNA.'' J. Mol. Biol., v282, pp331-343.

		Minor Groove		Major Groove	
		P-P	Refined	P-P	Refined
1	GG/CC	---	---	---	---
2	GG/CC	---	---	---	---
3	GG/CC	16.3	---	17.1	---
4	GG/CC	17.0	16.0	20.4	18.6
5	GG/CC	16.0	15.3	20.8	18.5
6	GG/CC	14.7	13.8	18.6	18.0
7	GT/AC	13.9	13.2	17.9	17.9
8	TC/GA	12.8	12.5	16.4	16.2
9	CC/GG	13.0	13.0	15.9	15.9
10	CT/AG	13.1	13.1	17.3	17.2
11	TA/TA	11.6	11.6	15.8	15.5
12	AT/AT	10.5	10.5	17.6	16.9
13	TA/TA	11.8	11.8	17.2	16.6
14	AG/CT	14.0	14.0	16.8	16.7
15	GA/TC	13.7	13.6	16.1	15.7
16	AA/TT	12.2	---	19.2	---
17	AC/GT	---	---	---	---
18	CT/AG	---	---	---	---

****

Global linear helical axis defined by equivalent C1' and RN9/YN1 atom pairs

Deviation from regular linear helix: 2.85(0.76)

****

Main chain and chi torsion angles:

```
Note: alpha:   O3'(i-1)-P-O5'-C5'
      beta:    P-O5'-C5'-C4'
      gamma:   O5'-C5'-C4'-C3'
      delta:   C5'-C4'-C3'-O3'
      epsilon: C4'-C3'-O3'-P(i+1)
      zeta:    C3'-O3'-P(i+1)-O5'(i+1)

      chi for pyrimidines(Y): O4'-C1'-N1-C2
          chi for purines(R): O4'-C1'-N9-C4
```

Strand I

base		alpha	beta	gamma	delta	epsilon	zeta	chi
1	G	---	-170.2	73.1	82.0	-152.8	-70.0	-172.9
2	G	-79.5	-178.4	49.9	90.6	-149.4	-67.0	-162.6
3	G	-78.2	169.8	57.7	73.5	-163.3	-64.3	-172.9
4	G	-92.9	-177.5	71.5	76.5	-140.6	-93.3	-176.2
5	G	-59.6	158.4	55.7	67.7	-155.6	-61.5	-168.2
6	G	-71.0	161.4	62.0	73.1	-156.0	-88.8	-177.2

7 G	123.5	-165.1	-137.7	125.7	-147.4	-82.7	-164.8
8 T	-176.7	-142.6	149.5	101.5	-148.3	-61.2	-153.5
9 C	-58.8	-174.1	21.1	159.4	-110.7	171.2	-79.4
10 C	-92.0	143.7	68.4	154.1	-156.8	-107.9	-122.7
11 T	-35.9	-174.3	-6.0	169.1	-117.1	167.1	-69.6
12 A	-93.0	148.2	66.1	161.4	-120.5	-151.0	-109.0
13 T	-33.2	120.8	38.5	116.9	-123.5	-172.6	-123.3
14 A	-57.0	149.4	48.3	158.9	-164.6	-84.7	-94.1
15 G	-82.2	-176.7	40.0	154.1	-96.9	159.0	-69.5
16 A	-75.0	135.8	50.2	121.8	-147.3	-154.9	-112.8
17 A	6.5	98.0	45.3	80.0	177.5	-81.0	-136.9
18 C	105.5	-173.8	-165.4	81.5	83.6	54.5	-153.6
19 T	-142.8	170.5	85.0	80.9	---	---	-167.3

Strand II

base	alpha	beta	gamma	delta	epsilon	zeta	chi
1 C	-72.4	-176.5	67.2	128.1	---	---	-129.0
2 C	-43.5	172.3	43.4	148.7	-173.4	-112.6	-122.5
3 C	142.3	-163.6	-168.5	97.5	-136.0	-81.8	-170.0
4 C	-60.4	174.5	48.5	83.2	-179.6	-85.3	-166.2
5 C	-43.3	172.4	34.5	79.4	-151.6	-66.2	-155.8
6 C	107.3	-161.6	-159.1	89.0	-141.1	-88.0	-165.4
7 C	161.3	-116.0	148.8	88.1	-155.2	-103.4	-151.3
8 A	-78.6	126.0	59.1	86.8	175.4	-81.2	-135.3
9 G	-69.2	179.3	44.6	141.5	-109.5	-178.9	-92.7
10 G	-99.0	-157.7	76.8	111.6	-174.7	-74.4	-108.6
11 A	-64.6	168.0	47.6	139.1	153.4	-72.2	-88.7
12 T	-82.1	-174.6	63.1	146.1	-132.6	-163.4	-112.7
13 A	-70.6	130.0	55.8	140.4	175.2	-94.7	-97.2
14 T	-82.3	-167.0	69.2	156.5	-102.6	171.9	-88.1
15 C	-92.1	153.6	77.9	78.5	-170.8	-83.1	-129.7
16 T	38.9	90.6	38.3	104.6	-156.8	-86.5	-119.6
17 T	154.2	-128.5	165.8	80.1	-145.5	-126.3	-179.1
18 G	98.2	-172.9	-121.4	136.6	176.7	-104.8	-131.4
19 A	---	---	-157.3	94.7	-173.6	-100.2	-137.4

****

Sugar conformational parameters:

Note: v0: C4’-O4’-C1’-C2’
      v1: O4’-C1’-C2’-C3’
      v2: C1’-C2’-C3’-C4’
      v3: C2’-C3’-C4’-O4’
      v4: C3’-C4’-O4’-C1’

      tm: amplitude of pseudorotation of the sugar ring

P: phase angle of pseudorotation of the sugar ring

Strand I

base	v0	v1	v2	v3	v4	tm	P	Puckering
1 G	1.7	-26.8	40.6	-40.2	24.6	42.3	16.2	C3’-endo
2 G	11.7	-33.7	42.2	-36.2	15.7	42.2	2.8	C3’-endo
3 G	-5.5	-25.6	45.2	-50.0	35.1	49.9	25.1	C3’-endo
4 G	-22.3	-11.0	38.7	-52.3	46.7	52.7	42.8	C4’-exo
5 G	-5.4	-28.9	50.6	-55.1	38.0	55.4	24.1	C3’-endo
6 G	-4.0	-27.0	47.5	-51.1	34.3	51.6	23.1	C3’-endo
7 G	39.7	-32.4	14.6	7.4	-28.7	38.0	292.5	C1’-endo
8 T	-21.3	7.8	7.4	-19.8	25.5	25.3	73.1	O4’-endo
9 C	-22.0	44.1	-47.8	36.3	-9.4	48.3	172.1	C2’-endo
10 C	-32.4	50.5	-49.1	31.8	0.3	51.9	161.2	C2’-endo
11 T	-18.0	43.1	-48.9	41.5	-15.1	48.9	178.3	C2’-endo
12 A	-15.5	35.7	-41.9	34.2	-11.8	41.9	177.7	C2’-endo
13 T	-51.6	50.3	-29.7	0.8	31.5	52.3	124.6	C1’-exo
14 A	-7.3	30.1	-40.2	36.9	-19.3	40.6	188.6	C3’-exo
15 G	-23.0	41.8	-42.1	30.7	-5.1	43.2	167.4	C2’-endo
16 A	-36.4	38.4	-26.2	6.3	18.6	38.6	132.8	C1’-exo
17 A	-55.8	34.1	-0.4	-32.8	55.7	58.0	90.3	O4’-endo
18 C	-6.0	-19.6	36.1	-40.6	29.0	40.4	26.8	C3’-endo
19 T	-41.2	10.7	21.0	-45.1	52.9	53.0	66.6	C4’-exo

Strand II

base	v0	v1	v2	v3	v4	tm	P	Puckering
1 C	-21.4	29.0	-26.2	14.7	4.0	29.2	153.8	C2’-endo
2 C	-35.3	47.2	-41.3	22.0	8.1	46.9	151.6	C2’-endo
3 C	8.7	-24.7	30.5	-26.9	11.4	30.5	3.0	C3’-endo
4 C	-18.5	-9.5	31.4	-43.1	39.0	43.1	43.3	C4’-exo
5 C	-1.5	-21.6	36.9	-38.7	24.3	39.4	20.7	C3’-endo
6 C	4.2	-27.5	39.4	-38.2	21.3	40.4	12.9	C3’-endo
7 C	-44.0	21.3	6.9	-32.2	47.4	47.6	81.7	O4’-endo
8 A	-18.2	-4.7	24.3	-35.3	33.2	36.0	47.6	C4’-exo
9 G	-39.6	49.4	-40.7	19.0	12.6	48.7	146.6	C2’-endo
10 G	-33.8	28.7	-14.6	-4.5	23.9	32.9	116.3	C1’-exo
11 A	13.5	4.9	-19.9	28.4	-27.3	28.9	226.4	C4’-endo
12 T	-39.1	48.1	-38.2	17.2	13.5	46.9	144.6	C2’-endo
13 A	-27.6	39.7	-37.0	22.0	3.3	40.2	156.9	C2’-endo
14 T	-10.5	30.5	-37.5	33.0	-14.4	37.5	183.2	C3’-exo
15 C	-29.8	3.6	22.2	-40.0	44.0	44.2	59.8	C4’-exo
16 T	-37.5	34.3	-18.7	-2.4	25.0	37.3	120.2	C1’-exo
17 T	-28.6	1.7	24.4	-41.8	43.4	44.7	57.0	C4’-exo
18 G	-3.0	15.4	-21.2	20.1	-11.0	21.6	190.9	C3’-exo
19 A	-44.0	27.3	-1.9	-23.3	41.5	44.3	92.5	O4’-endo

****
Same strand P--P and C1'--C1' virtual bond distances

	Strand I			Strand II	
base	P--P	C1'--C1'	base	P--P	C1'--C1'
1 G/G	5.8	5.6	1 C/C	6.6	5.3
2 G/G	5.7	5.3	2 C/C	6.3	5.6
3 G/G	5.8	5.4	3 C/C	6.3	5.5
4 G/G	5.4	5.7	4 C/C	5.4	5.4
5 G/G	5.9	5.1	5 C/C	6.6	5.6
6 G/G	5.8	5.7	6 C/C	6.7	5.4
7 G/T	6.7	5.3	7 C/A	6.7	4.6
8 T/C	6.9	5.2	8 A/G	6.1	4.8
9 C/C	6.4	5.4	9 G/G	6.7	4.4
10 C/T	7.2	4.9	10 G/A	7.0	4.7
11 T/A	6.2	5.0	11 A/T	6.5	5.9
12 A/T	7.1	4.3	12 T/A	7.0	4.5
13 T/A	6.3	5.7	13 A/T	6.7	5.1
14 A/G	7.1	4.4	14 T/C	6.6	5.6
15 G/A	6.6	4.9	15 C/T	6.3	4.7
16 A/A	6.7	4.4	16 T/T	6.6	5.2
17 A/C	6.5	4.9	17 T/G	7.1	4.5
18 C/T	6.7	4.9	18 G/A	---	5.0

****
Helix radius (radial displacement of P, O4', and C1' atoms in local helix frame of each dimer)

	Strand I			Strand II		
step	P	O4'	C1'	P	O4'	C1'
1 GG/CC	8.9	8.1	7.2	9.1	7.2	6.5
2 GG/CC	8.0	8.3	7.8	11.1	10.1	9.4
3 GG/CC	9.9	9.6	9.1	10.4	9.6	9.1
4 GG/CC	10.7	10.4	10.0	9.5	9.4	9.0
5 GG/CC	9.4	9.1	8.6	9.5	9.2	8.6
6 GG/CC	10.9	10.9	10.4	11.4	9.8	9.4
7 GT/AC	9.4	7.4	6.5	10.0	7.3	6.5
8 TC/GA	9.3	6.0	5.3	8.7	6.1	5.3
9 CC/GG	10.3	7.3	6.9	7.8	4.3	3.8
10 CT/AG	10.8	7.7	7.3	9.4	6.9	6.1
11 TA/TA	7.2	4.3	4.1	9.0	6.6	6.1
12 AT/AT	10.3	6.9	6.4	8.1	5.0	4.6
13 TA/TA	7.4	5.0	4.6	9.0	6.0	5.8
14 AG/CT	9.5	6.8	6.0	9.0	7.6	6.6
15 GA/TC	10.0	6.8	6.2	9.0	5.7	5.1
16 AA/TT	9.0	6.0	5.1	9.8	7.0	6.2

17 AC/GT	11.0	7.5	6.8	10.1	7.3	6.4
18 CT/AG	10.5	7.9	7.2	10.0	7.2	6.6

****

Position (Px, Py, Pz) and local helical axis vector (Hx, Hy, Hz)
for each dinucleotide step

bp	Px	Py	Pz	Hx	Hy	Hz
1 GG/CC	1.94	155.38	31.00	0.62	-0.78	-0.03
2 GG/CC	2.47	151.13	32.67	0.43	-0.85	-0.30
3 GG/CC	3.81	149.58	33.38	0.45	-0.88	-0.11
4 GG/CC	4.52	147.45	33.83	0.35	-0.92	-0.17
5 GG/CC	6.63	145.33	32.97	0.33	-0.94	-0.05
6 GG/CC	5.82	142.28	32.01	0.41	-0.89	-0.20
7 GT/AC	9.07	139.52	36.07	0.30	-0.95	0.03
8 TC/GA	9.17	136.12	38.33	0.20	-0.97	-0.11
9 CC/GG	8.08	132.84	37.25	-0.09	-1.00	0.03
10 CT/AG	11.95	129.88	38.43	0.34	-0.91	0.25
11 TA/TA	7.96	125.84	38.27	0.16	-0.98	-0.14
12 AT/AT	11.32	123.00	39.49	0.27	-0.96	0.02
13 TA/TA	11.22	119.89	36.48	0.31	-0.95	-0.09
14 AG/CT	7.77	116.25	37.52	-0.12	-0.78	-0.61
15 GA/TC	9.57	112.91	37.03	0.08	-0.99	-0.12
16 AA/TT	10.14	110.11	35.51	0.11	-0.99	-0.06
17 AC/GT	8.61	107.25	33.63	0.01	-0.93	-0.36
18 CT/AG	8.50	103.79	31.77	0.09	-0.96	-0.27

References

Calladine, C. R. & Drew, H. R. (1997). *Understanding DNA; The Molecule & How It Works*, 2nd edn, Academic Press, London.

Chandrasekaran, R. & Arnott, S. (1989). The structures of DNA and RNA helices in oriented fibers, *in* W. Saenger (ed.), *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, Vol. VII/1b, Springer-Verlag, pp. 31–170.

Clowney, L., Jain, S. C., Srinivasan, A. R., Westbrook, J., Olson, W. K. & Berman, H. M. (1996). Geometric parameters in nucleic acids: Nitrogenous bases, *J. Am. Chem. Soc.* **118**, 509–518.

Dickerson, R. E. (1998). DNA bending: The prevalence of kinkiness and the virtues of normality, *Nucl. Acids Res.* **26**, 1906–1926.

Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O., Lavery, R., Nelson, H. C. M., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C.-S., von Kitzing, E., Wang, A. H.-J. & Zhurkin, V. B. (1989). Definitions and nomenclature of nucleic acid structure parameters, *J. Mol. Biol.* **205**, 787–791.

El Hassan, M. A. & Calladine, C. R. (1995). The assessment of the geometry of dinucleotide steps in double-helical DNA: A new local calculation scheme, *J. Mol. Biol.* **251**, 648–664.

El Hassan, M. A. & Calladine, C. R. (1998). Two distinct modes of protein-induced bending in DNA, *J. Mol. Biol.* **282**, 331–343.

Guzikevich-Guerstein, G. & Shakked, Z. (1996). A novel form of the DNA double helix imposed on the TATA-box by the TATA-binding protein, *Nature Struct. Biol.* **3**, 32–37.

Lavery, R. & Sklenar, H. (1989). Defining the structure of irregular nucleic acids: Conventions and principles, *J. Biomol. Struct. Dynam.* **6**, 655–667.

Leslie, A. G. W., Arnott, S., Chandrasekaran, R., Birdsall, D. L. & Ratlif, R. L. (1980). Left-handed DNA helices, *Nature* **283**, 743–746.

Lu, X. J., Babcock, M. S. & Olson, W. K. (1999). Mathematical overview of nucleic acid analysis programs, *J. Biomol. Struct. Dynam.* **16**, 833–843.

Lu, X. J., El Hassan, M. A. & Hunter, C. A. (1997a). Structure and conformation of helical nucleic acids: Analysis program (SCHNAaP), *J. Mol. Biol.* **273**, 668–680.

Lu, X. J., El Hassan, M. A. & Hunter, C. A. (1997b). Structure and conformation of helical nucleic acids: Rebuilding program (SCHNArP), *J. Mol. Biol.* **273**, 681–691.

Lu, X. J. & Olson, W. K. (1999). Resolving the discrepancies among nucleic acid conformational analyses, *J. Mol. Biol.* **285**, 1563–1575.

Lu, X. J., Shakked, Z. & Olson, W. K. (2000). A-dna conformational motifs in ligand-bound double helices, *J. Mol. Biol.* **300**, 819–840.

Merritt, E. A. & Bacon, D. J. (1997). Raster3d: Photorealistic molecular graphics, *Methods in Enzymology* **277**, 505–524.

Olson, W. K. & *et al.*, X. J. L. (2000). A standard reference frame for the description of nucleic acid base-pair geometry. http://ndb-mirror-0.rutgers.edu:5000/ (in press).

Shui, X., McFail-Isom, L., Hu, G. G. & Williams, L. D. (1998). The B-DNA dodecamer at high resolution reveals a spine of water on sodium, **37**, 8341–8355.

van Pouderoyan, G., Ketting, G., Perrakis, R. F. & Sixma, T. K. (1997). Crystal structure of the specific DNA-binding domain of Tc3 transposase of *c. elegans* in complex with transposon DNA, *EMBO J.* **16**, 6044–6054.