# NUPARM and NUCGEN: software for analysis and generation of sequence dependent nucleic acid structures

M.Bansal[1], D.Bhattacharyya[2] and B.Ravi

## Abstract

*Software packages NUPARM and NUCGEN, are described, which can be used to understand sequence directed structural variations in nucleic acids, by analysis and generation of non-uniform structures. A set of local inter basepair parameters (viz. tilt, roll, twist, shift, slide and rise) have been defined, which use geometry and coordinates of two successive basepairs only and can be used to generate polymeric structures with varying geometries for each of the 16 possible dinucleotide steps. Intra basepair parameters, propeller, buckle, opening and the C6...C8 distance can also be varied, if required, while the sugar phosphate backbone atoms are fixed in some standard conformation in each of the nucleotides. NUPARM can be used to analyse both DNA and RNA structures, with single as well as double stranded helices. The NUCGEN software generates double helical models with the backbone fixed in B-form DNA, but with appropriate modifications in the input data, it can also generate A-form DNA and RNA duplex structures.*

## Introduction

Sequence dependent variation of DNA structure has been demonstrated quite conclusively in recent years by several different experimental studies (Hagerman, 1990). In view of the obvious importance of this property of DNA, several groups have developed algorithms to analyse these features in nucleic acid structures obtained from crystallography, NMR or theoretical molecular mechanics and dynamics studies. Almost all the algorithms use a basepaired dinucleotide as a basic unit, but the definition of local coordinate axial system and origin differ considerably. These definitions are important when comparing different structures as well as in generating molecular models, using these local dinucleotide parameters. We have developed a suite of software packages NUPARM, NUCGEN and NUVIEW, to analyse, generate and view (on a graphics monitor) sequence

directed structural variations in nucleic acids. A set of local doublet parameters, viz. tilt, roll, twist, shift, slide and rise, defined according to the convention of the EMBO Workshop, 1989, have been used successfully in understanding base sequence dependent structural variations in DNA and double helical regions of tRNA crystal structures (Bhattacharyya and Bansal, 1989,1990,1992, 1994). A notable feature of the algorithms is that the local doublet parameters are self consistent. Thus the values of the local doublet parameters obtained by analysis of a structure by NUPARM can be used to regenerate the structure using the program NUCGEN. The program generates three dimensional coordinates of all the atoms in the dinucleotide unit with the basepairs being in the geometry described by the local dinucleotide parameter file and the sugar phosphate backbone attached in a fixed conformation as defined by the input coordinate file for the four nucleotides, A, T, G and C (Bhattacharyya and Bansal, 1988,1989). Hence the software can be used to generate uniform, fiber type, polymer models, by assigning the same values to all dinucleotide steps, as well as non-uniform or curved models, by defining appropriate values for each of the 16 dinucleotide steps.

The NUCGEN program also estimates the bending or curvature of the generated model by calculating three different parameters and generates several output files which can be used to analyse the structure (using NUPARM), view a schematic or skeletal model (using the NUVIEW software described in the accompanying paper) as well as to obtain plots.

## Systems and methods

NUPARM and NUCGEN have been written in FORTRAN77 and successfully compiled and executed using Microsoft Fortran compiler V 5.0 (under the DOS 5.0 environment) and Green Hills Fortran compiler (under the UNIX V 5.3 OS), on IBM-PC compatible machines. They have also been successfully run on VAX/ VMS machines and several UNIX workstations (e.g. Silicon Graphics Indigo and IBM R4000 machines). The NUCGEN program creates an HPGL plotfile of the generated molecule, that can be used to produce a hardcopy output, either on an HP pen-plotter or an HP-PCL5 compatible laser printer (tested on

[1] To whom correspondence and reprint requests should be sent, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India e-mail address: mb@mbu.iisc.ernet.in

[2] Present address· Div. Computer Research & Technology, Building 12A, NIH, Bethesda, MD 20892, USA

**Table I.** The definitions of the intra-basepair as well as local doublet and local helical parameters are briefly described, along with the relevant axial system used in the NUPARM algorithm

| Parameter | Symbol | Definition |
|---|---|---|
| Base normal | $\bar{N}_i$ | Vector normal to mean base plane defined by all the ring atoms of the base |
| BP normal | $\bar{Z}_i$ | Mean of the normals of both the bases in the basepair, normalized to unit vector |
| BP Long axis | $\bar{Y}_i$ | Vector along the line joining C8 (of purine) and C6 (of pyrimidine) or the two C1' atoms in the basepair |
| BP Short axis | $\bar{X}_i$ | Vector normal to the $\bar{Y}_i$ and $\bar{Z}_i$ axes and pointing towards the major groove side of the basepair |
| BB vector | $\bar{M}$ | Vector joining the basepair centers of two consecutive basepairs |
| Mean doublet z-axis | $\bar{Z}_m$ | $= \dfrac{(\bar{X}_1 + \bar{X}_2) \times (\bar{Y}_1 + \bar{Y}_2)}{\lvert(\bar{X}_1 + \bar{X}_2)\rvert \cdot \lvert(\bar{Y}_1 + \bar{Y}_2)\rvert}$ |
| Helical doublet z-axis | $\bar{Z}_h$ | $= \dfrac{(\bar{X}_1 - \bar{X}_2) \times (\bar{Y}_1 - \bar{Y}_2)}{\lvert(\bar{X}_1 - \bar{X}_2)\rvert \cdot \lvert(\bar{Y}_1 - \bar{Y}_2)\rvert}$ |
| **Intra-basepair parameters** | | |
| Propeller | $\pi$ | $= \cos^{-1}[(\bar{N}_1 \times \bar{Y}_i) \cdot (\bar{N}_2 \times \bar{Y}_i)]$ |
| Buckle | $\kappa$ | $= -\cos^{-1}[(\bar{N}_1 \times \bar{X}_i) \cdot (\bar{N}_2 \times \bar{X}_i)]$ |
| **Local doublet step parameters** | | |
| Tilt | $\tau$ | $= -\sin^{-1}(\bar{Z}_m \cdot \bar{X}_1)$ |
| Roll | $\rho$ | $= \sin^{-1}(\bar{Z}_m \cdot \bar{Y}_1)$ |
| Twist | $\omega_l$ | $= \cos^{-1}[(\bar{X}_1 \times \bar{Z}_m) \cdot (\bar{X}_2 \times \bar{Z}_m)]$ |
| Shift | Dx | $= \bar{M} \cdot (\bar{X}_1 + \bar{X}_2)/\lvert\bar{X}_1 + \bar{X}_2\rvert$ |
| Slide | Dy | $= \bar{M} \cdot (\bar{Y}_1 + \bar{Y}_2)/\lvert\bar{Y}_1 + \bar{Y}_2\rvert$ |
| Rise | Dz | $= \bar{M} \cdot \bar{Z}_m$ |
| **Local helical parameters** [a] | | |
| Inclination | $\eta$ | $= \sin^{-1}(\bar{Z}_h \cdot \bar{X}_1)$ |
| Tip | $\theta$ | $= -\sin^{-1}(\bar{Z}_h \cdot \bar{Y}_1)$ |
| Twist | $\omega_h$ | $= \cos^{-1}[(\bar{X}_1 \times \bar{Z}_h) \cdot (\bar{X}_2 \times \bar{Z}_h)]$ |
| x-displacement | dx | $= \bar{M} \cdot (\bar{X}_1 - \bar{X}_2)/\lvert\bar{X}_1 - \bar{X}_2\rvert$ |
| y-displacement | dy | $= \bar{M} \cdot (\bar{Y}_1 - \bar{Y}_2)\|\bar{Y}_1 - \bar{Y}_2\|$ |
| Helical rise | dz | $-\bar{M} \cdot \bar{Z}_h$ |

[a] The basepair parameters w.r t. a global helical axis are given by identical relations, but with $\bar{Z}_h$ axis now defined for the whole molecule.

HP-LJET III). Other output files are in formatted tabular form. Any general purpose graphing software like Grapher or Cricket Graph (on IBM-PC) and TOP-DRAW (on IBM workstations) can be used to obtain plots of the various parameters listed by NUPARM or the bending angle (sequential and cumulative), or end-to-end distance values listed by NUCGEN.

## Algorithm

The detailed mathematics and parameter definitions used in the analysis and generation algorithms for NUPARM and NUCGEN have been described earlier (Bhattacharyya and Bansal, 1988, 1989) and are briefly summarised in Table I. The NUPARM analysis package has been deposited with the Brookhaven Protein Data Bank, as well as the Nucleic Acid Data Base (Berman *et al.*, 1992). This program can be run interactively and allows several choices for defining the basepair origin and y-axis, as well

as for choosing the global helix axis, a sample session output is shown in Figure 1. NUPARM calculates various intra- and inter-basepair parameters for any given single or double stranded nucleic acid structure and lists these for each step, along with the mean and standard deviation values for each parameter. The angles between successive basepair normals and local helix axes are also tabulated. A single global helix axis can be fitted to the whole molecule and the related set of basepair parameters, inclination, tip, displacement, helical twist and rise calculated with respect to this axis (EMBO Workshop Definitions, 1989). The interchain phosphate separation (which defines the groove width in a double helical structure) and the backbone torsion angles can also be obtained.

The related software package NUCGEN has been developed to generate a nucleic acid fragment, corresponding to any given sequence and with a pre-defined set of local doublet parameters. All six inter-basepair parameters (the three rotations tilt, roll, twist and the

```
Type in the name of the INPUT DATA file:
COORDINATES TAKEN FROM :
5dnbcg.ent

$Is the data file in the "BROOKHAVEN PDB" format? [Y]
   Y

$Is the molecule a double stranded helix? [Y]
   Y

$Are the residues numbered according to the Brookhaven format? [Y]
   Y

$Should the base-pair ORIGIN be at the Centre of Gravity?
 (NOT RECOMMENDED!! Default is midpoint of C6--C8:) [N]
   N

$Should the Y-axis be along the line joining C1' atoms?
 (If "No",Y-axis is taken along the C6--C8 direction:) [N]
   N

$Are Single Strand parameters required?(Y/N)
   N

$Should the molecule be reoriented about a LINEAR GLOBAL Helix Axis? [Y]
   Y

Should the GLOBAL AXIS be fitted to:
1. Local HELIX ORIGINS (Type O)
2. Base-Pair CENTERS (Type C)
3. Any Backbone atom (Type the atom name, exact match required)
   C1'

Results of the calculation are written into:
5dnbcg.prm
Reoriented atomic coordinates written into:
5dnbcg.coor

$Are P--P distances and Cylindrical Polar Coordinates required?
Only meaningful for reoriented molecule. [Y]
   Y

$Is the torsion angle calculation required? [Y]
   Y
```

Fig. 1. A sample session output for the NUPARM software (stored in the file RUN.ANS) showing the various options available for detailed analysis of a nucleic acid fragment.

three translations, shift, slide and rise) as well as the intra-basepair parameters propeller, buckle, opening angle and C6 ... C8 distance can be varied, providing a very versatile and accurate generation procedure. In the case of long polymers, some of these parameters may be assigned standard values in the input parameter table (shown in Figure 2). The local doublet parameters are converted to local helical parameters for the purpose of generation and the molecular structure is generated in the $5'$ to $3'$ strand direction, with the required sequence (Bhattacharyya and Bansal, 1988,1989).

The bending or curvature of the generated model is estimated by calculating three different parameters (Bansal and Bhattacharyya, 1990). (i) The ratio of the end-to-end distance $d$ to the actual path length $l_{max}$ is calculated; (ii) the three principle moments of inertia are obtained and the ratio of the largest to the smallest moments, which is an indicator of the length to diameter ratio of the inertia ellipsoid fitted to the model structure, is then calculated; (iii) the angles between the vectors connecting every $n$th basepair centre are

calculated to trace the bending trajectory of the model structures.

## Implementation

NUCGEN can generate model structures in two ways:

1) Generate a molecular model corresponding to a given sequence, using some defined local doublet parameters for the 16 dinucleotide steps.
2) Regeneration of a particular sequence from the values of the local doublet parameters of all the doublet steps.

The NUCGEN program is very simple to use and functions interactively, a sample session output is shown in Figure 3.

### Input files

'COOR-B89.DAT' is a standard data file containing the coordinates of all the atoms in the four nucleoides (Adenosine, Thymidine, Guanosine and Cytidine), with standard B–DNA conformation (as given by

Dinucleotide step params from crystal strs, basepair params from fiber model

| Seq | Tilt | Roll | Twist | Shift | Slide | Rise | Prop | Buckl | Glyc1 | Glyc2 | C6C8 |
|-----|------|------|-------|-------|-------|------|------|-------|-------|-------|------|
| G:G | -1.83 | 5.82 | 32.40 | -0.06 | 0.66 | 3.47 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| C:C | 1.83 | 5.82 | 32.40 | 0.06 | 0.66 | 3.47 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| G:C | -0.18 | -6.49 | 38.61 | -0.18 | 0.52 | 3.34 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| C:G | 0.05 | 3.21 | 35.06 | 0.13 | 0.55 | 3.53 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| A:A | -1.66 | 2.66 | 35.58 | -0.14 | -0.18 | 3.23 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| T:T | 1.66 | 2.66 | 35.58 | 0.14 | -0.18 | 3.23 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| A:T | 0.21 | -0.29 | 32.49 | 0.06 | -0.41 | 3.29 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| T:A | 0.04 | 0.54 | 39.04 | -0.19 | 0.22 | 3.22 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| A:C | -0.15 | -0.70 | 33.29 | -0.24 | -0.12 | 3.35 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| G:T | -0.15 | -0.70 | 33.29 | -0.24 | -0.12 | 3.35 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| G:A | -0.88 | 0.31 | 38.55 | -0.01 | -0.12 | 3.10 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| T:C | 0.88 | 0.31 | 38.55 | 0.01 | -0.12 | 3.10 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| A:G | -0.68 | 7.10 | 28.36 | -0.14 | 0.74 | 3.52 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| C:T | 0.68 | 7.10 | 28.36 | 0.14 | 0.74 | 3.52 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| C:A | 0.58 | -7.50 | 47.91 | 0.14 | 2.91 | 3.40 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| T:G | 0.58 | -7.50 | 47.91 | 0.14 | 2.91 | 3.40 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |

Comments:
B-DNA Mean Local parameters obtained from all the B-form crystal structures in
Nucleic acid database, July 1993 (except AG, AC and CA).
For AG/CT & AC/GT seqs, mean values obtained incl of Protein-DNA complexes.
For CA/TG seq, only steps which have B-II conformation and high twist are included.
For CA/TG steps with B-I conformation and low twist, the params are:

| | | | | | | | | | | | |
|-----|------|------|-------|------|------|------|--------|------|-------|-------|------|
| C:A | -0.31 | 5.09 | 31.68 | 0.25 | 0.72 | 3.58 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |
| T:G | -0.31 | 5.09 | 31.68 | 0.25 | 0.72 | 3.58 | -15.00 | 0.01 | 54.10 | 53.50 | 9.82 |

**Fig. 2.** The standard input parameter file 'crystal.loc' for NUCGEN package, showing the various inter and intra basepair parameters which can be defined for the sixteen possible dinucleotide steps and varied while generating nucleic acid molecules (Bhattacharyya,D. and Bansal,M., 1989).

Chandrasekaran and Arnott, 1989) and does not generally require modification for B-form DNA structures. To generate A-form models, the backbone atoms should be fixed in the appropriate conformation with C3'-endo sugar pucker, while for RNA structures, Thymidine should be replaced with Uridine. The input file '***.LOC' contains the relevant basepair generation information in the format '(1X,A3,3X,11F8.2)', which is similar to the output of NUPARM program.

If regeneration of a structure is required, the ***.LOC defined values of the local doublet step and intra basepair parameters are used for each of the steps. The file should contain the basepair residue names (e.g. A:T, G:C), the inter-basepair parameters for each doublet step (i.e. tilt, roll, twist, shift, slide and rise) as well as the intra-basepair parameters (i.e. propeller, buckle, the two glycosidic angles (C1'-C1'-N9/N1), which are related to the opening angle, and the distance between C8 and C6 atoms of the basepair), exactly as defined by the NUPARM algorithm.

If a polymer structure corresponding to a defined base sequence, contained in a file '***.SEQ' has to be generated, using standard dinucleotide step parameters, then the ***.LOC file can be assigned at run time and contains the local step parameters of the 16 possible dinucleotide sequences. The step parameter values have been obtained from crystal structures (Nagaich et al., 1994) (this is the default option, contained in the file

'crystal.loc' and shown in Figure 2) or empirical models (Bolshoy, et al., 1991). The file ***.LOC in this case defines the doublet sequences (e.g. A–A, A–T, G–A, G–G etc.) in the character field, instead of the basepairs (A : T, G : C). Furthermore, some standard values of intra basepair parameters are used, e.g. −15° propeller, 0° buckle, and 54° for glycosidic angles C1'– C1'– N9/N1, which corresponds to ≈20° opening angle in Watson–Crick basepairs. The C8–C6 distance is taken to be 9.8 Å. All the atoms in the nucleotide unit can be generated, but the generation can be speeded-up considerably, by generating only three atoms per base (C1', N9/N1 and C8/C6), and is recommended for large polymer structures.

The third input file, '***.SEQ' contains the base sequence information in one strand (in 5' → 3'direction) and is required in the case of polymer generation option. This file should also specify the number of basepairs required in the first line and a title in second line, followed by the sequence in EMBL sequence data format (5X,6(10A1,1X)).

*Output files*

***.ANG contains the moments of inertia components and also the bending parameters for the generated molecule.
***.HLX sequentially lists the C1' atoms (in Brookhaven

D \>nucgen94

Is polymer generation option required? [default Y]  y

Is 3-ATOM fast generation option required? [Y]  y

Type in the name of input SEQUENCE DATA file  alu156.seq

180 mer alu156c seq
Enter fragment length for calc the bending angle.  5

Use standard CRYSTAL LOC parameter file? [Y]  y

Is PLOTTER file option required ? [N]  y

Local params from NDB crystal data (1994),BI & B-II CA steps, opp sign tilts
$GENERATING BASEPAIR 180 OF 180

WRITING EVERY nTH BP Centre COORDS IN  alu156 bpc

WRITING PATH LENGTH,DIST,d/l RATIO IN  alu156 ang

r3/r1 =  13 078    [r2+r3]/r1 =  25 279      Rad Gyr =  5060.97
No  basepair =  180  , d/l(helical) =  0 847  , d/l(local) =  0 838

WRITING COORDS OF C1 ATOMS INTO FILE  alu156 hlx

WRITING C1 COORS FOR HPGL-PLOTTER FILE IN  : alu156 plt

Fig. 3. A sample session output for the NUCGEN software, showing the options available at run-time.

PDB format) for each of the basepairs and inserts a title, as contained in the input ***.SEQ file. The ***.HLX is the input file for the display package NUVIEW (described in the accompanying paper).
***.PLT is the HPGL plotfile created for obtaining the hardcopy output of the path traced by the C1' atoms in the generated DNA molecule.
***.BPC contains the coordinates of all the basepair centres in the sequence.
***.COR lists the cartesian coordinates of all the atoms in the Brookhaven Protein Data Bank format, when the regeneration or all atom options are used.

## Applications and Discussion

All six inter-basepair parameters as well as intra-basepair parameters propeller, buckle, opening angle and C6 ... C8 distance can be varied or any of them can be assigned some standard value, thus providing an immensely versatile environment.

The NUCGEN algorithm has been tested in two ways:

1) Generating a structure using a set of local parameters, calculating the local parameters by NUPARM and comparing the two sets of parameters, which give an extremely good match (mostly differing in the second decimal place only).
2) Local doublet parameters calculated by NUPARM are used to generate the same structure and the basepairs in
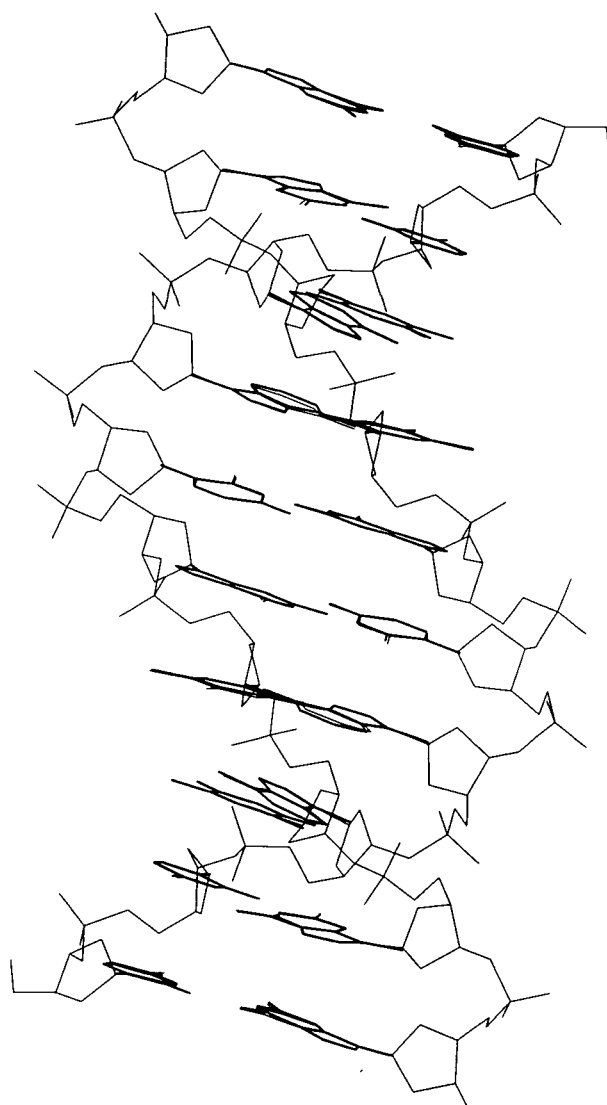


Fig. 4. The basepairs generated using the local parameters obtained from the NUPARM analysis of the decamer crystal structure of d(CCAACGTTGG) (Prive et al., 1991) and the regeneration option in NUCGEN are shown (in thick line) superposed on the original structure (in thin line). The r.m.s. d between the C1' atoms in the two sets of coordinates is 0.13 Å.

the original and regenerated structures are superimposed, which again show very good accuracy as seen in Figure 4 (r.m.s. deviation between the two sets of coordinates are generally less than 0.2 Å).

The software has also been used to generate models of several polymeric sequences, using a set of dinucleotide parameters, obtained from an analysis of all available B–DNA oligonucleotide crystal structures and standard intra basepair parameters (Figure 2). The DNA curvature predicted for these model structures was found to be in reasonably good agreement with that obtained from gel mobility studies (Nagaich et al., 1994). A representative

```
180 mer
    180 mer alu156 (-150 to +30) seq
        CCCCGGATCC GTCCTTCCAA GAAAGATATC CTAACAGCAC AAGAGCGGAA ACACGTTTTG
        TTCTACATCC AGAACAACCT CTGCTAAAAT TCCTGAAAAA TTTTGCAAAA AGTTGTTGAC    (a)
        TTTCTCTACG AGGTGTGGCA TAATAATCTT AACAACAGCA GGACGCTAGG ACGGATCCGG
```
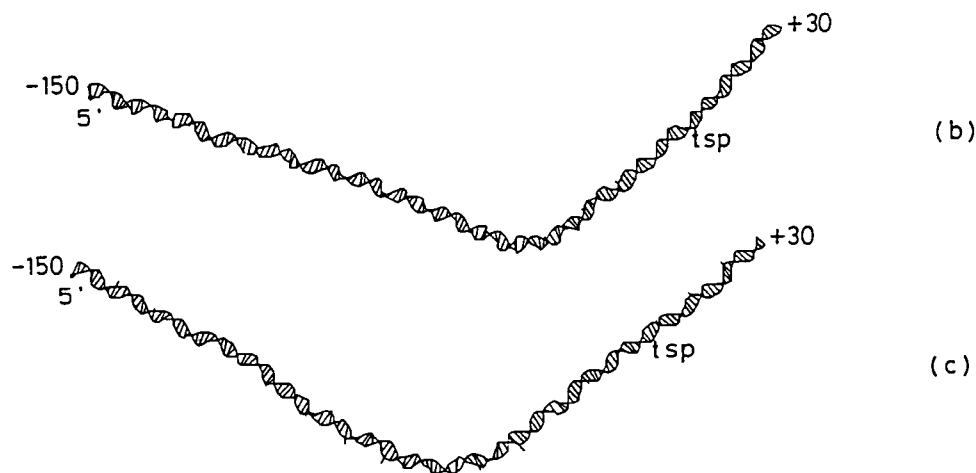


Fig. 5. (a) The input file 'alu156.seq', containing the 180 basepair long Alu156 promoter sequence (−150 to +30) from *B. subtilis* (McAllister and Achberger, 1988) (b) The model structure generated for the Alu156 promoter sequence, using the NUCGEN software and the local doublet parameters listed in Figure 2, shows bending upstream of the transcription start position(tsp), also indicated by the bending parameters listed in Table II. (c) The model structure generated using the equivalent parameters proposed by Bolshoy *et al.* (1991), which shows similar bending.

model structure for the Alu 156 promoter sequence (Figure 5a), is shown in Figure 5b and the bending parameters are listed in Table II. Since in principle any set of dinucleotide parameters, which are defined according to the EMBO convention can be used, we have also generated models using the parameter values given by Bolshoy *et al.* (1991) and the model structure for the same sequence is shown in Figure 5c. It is interesting to note that both sets of parameters predict very similar structures for this sequence, though some other sequences (e.g. the kinetoplast sequence from *L. tarentolae*, Diekmann and Wang, 1985) are predicted to show greater bending with the parameters of Bolshoy *et al.*, (1991).

Table II. The parameters calculated for the model structure, generated using NUCGEN and shown in Figure 5(b), for the 180-mer Alu156 promoter sequence and which indicate the extent of bending. A fragment length of 5 basepairs was used to calculate the bending angle between the vectors connecting the basepair centres

| | |
|---|---|
| Ratio of principle moments of inertia $(I_{max}/I_{min}) \cdot 13.08$ | |
| Radius of gyration : 5060.97 Å | |
| Ratio of end-to-end distance to the path length (d/l) : 0.84 | |
| Largest cumulative bending angle : 83.1°, position : 156 | |
| Largest successive bending angle : 27.2°, position : 106 | |

Thus the set of programs described here can be used to generate model structures for any chosen sequence and look for sequence dependent variations, as well as structure–function correlations in nucleic acid structures.

## Acknowledgements

## References

Bansal,M. and Bhattacharyya,D. (1990) Real and apparent curvature in oligonucleotides. In Sarma,R H. and Sarma,M.H. (eds). *Structure & Methods, Volume 3 DNA & RNA*, Adenine Press, New York, pp. 139–153.

Berman,H.M., Olson,W.K., Beveridge,D.L., Westbrook,J., Gelbin,A., Demeny,T., Hsieh,S H , Srinivasan,A.R. and Schneider,B. (1992) Nucleic acid Data Base. *Biophys J.*, 63, 751–759.

Bhattacharyya,D. and Bansal,M. (1988) A general procedure for generation of curved DNA molecules. *J Biomol. Str Dyn.*, 6, 93–104.

Bhattacharyya,D. and Bansal,M. (1989) A self-consistent formulation for analysis and generation of non-uniform DNA structures. *J. Biomol. Str. Dyn* , 6, 635–653.

Bhattacharyya,D. and Bansal,M. (1990) Local variability and base sequence effects in DNA crystal structures, *J Biomol. Str. Dyn.*, 6, 539–572.

Bhattacharyya,D. and Bansal,M. (1992) Groove width and depth of B-DNA structures depend on local variation in slide. *J. Biomol. Str Dyn.*, 10, 213–226.

Bhattacharyya,D. and Bansal,M. (1994) Analysis of sequence dependent

variations in secondary and tertiary structure of tRNA molecules, *J. Biomol. Str. Dyn.*, **11**, 1251–1275.

Bolshoy,A., McNamara,P., Harrington,R.E. and Trifonov,E.N. (1991) Curved DNA without A-A: Experimental estimation of all 18 DNA wedge angles, *Proc. Natl Acad. Sci., USA*, **88**, 2312–2316.

Chandrasekaran,R. and Arnott,S 1989. The structures of DNA and RNA helices in oriented fibers. In W.Saenger (ed.), Landolt Bornstein New Series, Group VII, Vol. 1b, Springer-Verlag, Berlin, p.31.

Diekmann,S. and Wang,J.C. (1985) On the sequence determinants and flexibility of the kinetoplast DNA fragment with abnormal gel electrophoretic mobilities, *J. Mol. Biol.*, **186**, 1–11.

EMBO Workshop (1989) Definitions and nomenclature of nucleic acid structure parameters, *EMBO J.*, **8**, 1–4, *J. Mol Biol.*, **205**, 787–791.

Hagerman,P.J. (1990) Sequence directed curvature of DNA, *Annu. Rev. Biochem.*, **59**, 755–781.

McAllister,C.F. and Achberger,E.C. (1988) Effect of polyadenine-containing curved DNA on promoter utilization in *Bacillus subtilis*, *J. Biol. Chem.*, **263**, 11743–11749.

Nagaich,A.K. Bhattacharyya,D., Brahmachari,S.K. and Bansal,M. (1994) CA/TG sequence at the 5′-end of oligo(A)-tracts strongly modulates DNA curvature*, *J. Biol. Chem*, **269**, 7824–7833.

Prive,G.G., Yanagi,K. and Dickerson,R.E. (1991) Structure of the B–DNA decamer C-C-A-A-C-G-T-T-G-G, *J. Mol Biol.*, **217**, 177–190.