Volume 17 Number 5 1989 Nucleic Acids Research

Definitions and nomenclature of nucleic acid structure components

R.E.Dickerson *et al.*

Molecular Biology Institute, University of California, Los Angeles, CA 90024, USA

Received February 7, 1989; Accepted February 7, 1989

ABSTRACT

We report here recommendations for the definitions and nomenclature of nucleic acid structure parameters. These recommendations result from discussions at an EMBO Workshop on DNA Curvature and Bending held at Churchill College, Cambridge, in September 1988.

At an EMBO Workshop on DNA Curvature and Bending, held at Churchill College, Cambridge, on 10–15 September 1988, two sessions were scheduled on definitions of parameters used to describe the geometry of nucleic acid chains and helices, and a common nomenclature for these parameters. The most widely used library of helix analysis programs, HELIB (1,2), suffers from the fact that the translations and rotations as defined are not fully independent and depend to a certain extent on the choice of overall helix axis. Several research groups have been engaged independently in developing alternative programs for the geometrical analysis of polynucleotide chains, but with different definitions of quantities calculated and with widely different nomenclature even when the same parameter was involved. The EMBO workshop sessions involved four such programming groups and other potential users, and had as its goal the introduction of a common set of concepts and common language for greater ease of communication. It is hoped that the standards agreed will prove acceptable to others in the field, and ultimately will be approved by the IUPAC/IUB Commission on Chemical Nomenclature(3).

The following points were agreed upon at the EMBO Workshop:

(i) The stated goal in program development is the creation of a new and optimised library routines for the analysis and description of polynucleotide structure, especially, but not exclusively, the DNA double helix. This general purpose library is intended mainly for those who solve nucleic acid structures using X-rays, NMR or other physical techniques and those who are interested in these results. It is recognised that theoretical studies will require more elaborate and more specialised programs, but it is felt that the first order of business is the creation of a simple, easily used library calculating easily understood and physically meaningful structure parameters. In particular, the new analysis routines should involve the least possible change from the familiar HELIB library that has been in use for more than five years, consonant with making the three rotations and three translations leading from one base pair to the next mathematically independent.

(ii) Trial programs developed by several research groups are to be circulated to interested users over the next 12–18 months for practical evaluation, before any attempt is made

to select a 'standard' library. Programs will be contributed by the following four groups, but others may participate if desired:

a. Richard Lavery and Heinz Sklenar (4)
b. D.M. Soumpasis and Chang-Shung Tung (5)
c. E. von Kitzing and S. Diekmann (6)
d. Manju Bansal (7)

Each of the programming participants will apply the resulting routines to three test cases: the Drew Native B-DNA of sequence C-G-C-G-A-A-T-T-C-G-C-G, an A-DNA structure to be supplied by Andrew H.-J. Wang and a tRNA molecule. The tables of parameters generated will be circulated to all interested parties for comparison.

(iii) All programs should have as a user option the choice of base pairs versus individual bases of a single strand. They should also allow calculations to be carried out relative to the local helix axes (from one base pair to the next), and relative to a long-range or global axis.

(iv) The x direction of a local or base pair coordinate set should point along the short axis of the base pair, the y direction along the long axis and the z direction perpendicular to the plane of the pair, in a right-handed orthogonal axial set. (Directions of the x, y and z are considered below, following definitions of the parameters.) The long axis of a base pair can be defined either by the line from the C6 of a pyrimidine to a C8 of a purine, or alternatively by the line from C6 of a pyrimidine to a hypothetical C8* atom on the purine, chosen so that the C6-C8* vector is parallel to the C1′-C1′ vector. (The choice should be stated explicitly.) If desired, employment of axes along the three principal moments of inertia of a base pair may be incorporated as an extra user option, but should not replace the simpler definitions.

(v) Axes for calculating parameters of each base step should be chosen so that the same numerical values result (with only a possible change of sign) when going from base pair 1 to base pair 2, as from base pair 2 to 1. One way in which this can be accomplished is by choosing a local reference axis set intermediate between those of the base pairs themselves.

(vi) The agreed-upon common nomenclature of parameters is as follows, using Greek letters for rotations and Roman letters for translations in accordance with IUPAC recommended practice:

a. From one base pair to the next:

Motion	Axis	Name	Symbol	Old Symbols
Rotation	z	Twist	Ω	t, Wdg, τ, θ, θ_w
Rotation	y	Roll	ϱ	R_{wd}, ϱ, ϱw, θ_R
Rotation	x	Tilt	τ	T_{wd}, t, τ_w, θ_T, λ
Translation	z	Rise	Dz	R_w, Z_{sh}, h, d, dz
Translation	y	Slide	Dy	S_l, Y_{sh}, Y_{dd}, dx, σ
Translation	x	Shift	Dx	S_s, X_{sh}, X_{dd}, dy

b. For individual base pairs:

Propeller twist about long y axis	ω	PrTw, Pro, TW, θ_p
Buckle about short x axis	$\varkappa$	
Inclination to overall helix axis	η	TL
Displacements from overall helix axis	dx, dy, Da $Da^2=dx^2+dy^2$	D, d

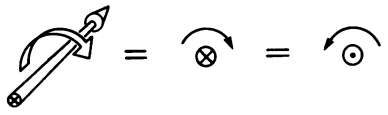

Fig. 1. Positive rotation as defined by the right-hand rule. Left, perspective; Centre, view along axis vector; Right, view down the end of the axis vector.

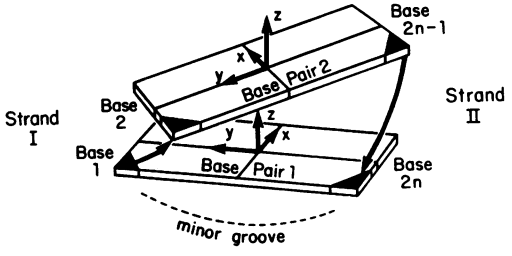

Fig. 2. Definitions of local reference axes (x,y,z) at the first two base pairs of an n-base pair double helix. View is into the minor groove. Shaded corners locate attachments of bonds to sugar C1′ atoms. Curved arrows indicate 5′−3′ direction of each backbone strand. Bases along strand **I** are numbered from 1 to n in a 5′−3′ direction, and bases back along strand **II** are numbered from $n+1$ to $2n$ also in a 5′−3′ direction. Base pairs are numbered from 1 to n, in agreement with the bases of strand **I**.

Fig. 3. Positive roll opens the angle between base pairs towards the minor groove. View is along the long base pair axis from strand **I** to strand **II**, or along $-y$. Curved arrow represents rotation of base 2 relative to base 1, about the y axis.

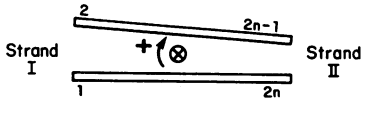

Fig. 4. Positive tilt opens the angle between base pairs toward strand **I**. View is from the minor groove side as in Figure 2, or along $+x$. Curved arrow represents rotation of base pair 2 relative to base pair 1, about the x axis.

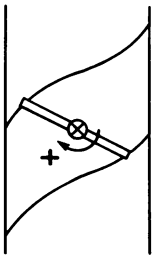

Fig. 5. Schematic view of the minor groove of an A helix, showing the positive inclination of one base pair.

Fig. 6. Illustration of positive propeller twist, viewed down the long axis of the base pair. The usual propeller twists in A- and B-DNA are negative with this convention, which is used because it is consistent with sign definitions for torsion angles.

Main chain torsion angles alpha through zeta, and glycosyl angle chi, are retained as in the standard IUPAC/IUB nomenclature.

(vii) The six parameters in group (a) above should always be clearly identified as to whether they are local helix axes between two base pairs, or global parameters calculated relative to an overall helix axis. The quantities can differ considerably; the rise for A-DNA along the global helix axis is $Dz_g=2.9$ Å, whereas the rise calculated from local axes will be the difference between stacked base pairs, $Dz_l=3.4$ Å. Subscripts sub-g

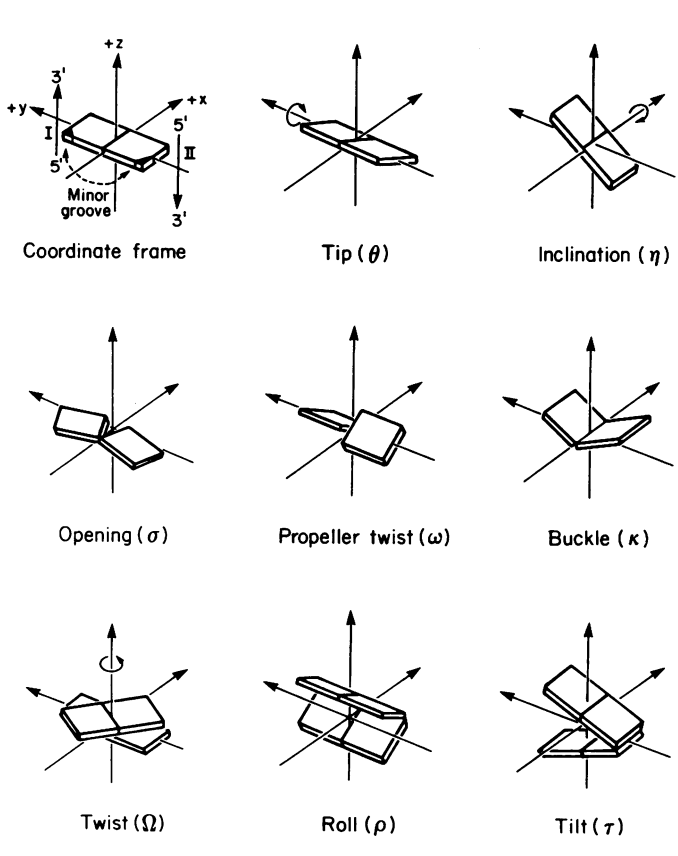

Fig. 7. Definitions of various rotations involving two bases of a pair (upper two rows) or two successive base pairs (bottom row). In the top row the motions of the two bases are coordinated, and in the middle row their motions are opposed. Columns at left, centre and right describe rotations about the z, y and x axes respectively. The standard coordinate frame is defined at upper left.

for global and sub-l for local should be used whenever parameters of both types are being discussed.

(viii) Careful attention was given to the employment of a consistent sense of positive rotations, using the right-hand rule diagrammed in Figure 1. Positive (x,y,z) axis directions were chosen so that all the rotational parameters except for propeller twist had the same sign as in the present HELIB library and in the published literature of the past decade. To achieve this, axes were chosen as in Figure 2: positive x in the direction of the major groove, positive y from helix strand **II** to helix strand **I** (or from the base to the backbone strand, for single chains), and positive z pointing in the $5'-3'$ direction of strand **I**.

With this choice roll (ϱ) is positive when the angle between base pairs opens towards the minor groove as in the past (Figure 3), tilt (τ) is positive when the angle between base pairs opens towards strand **I** (Figure 4) and the individual base pairs of A-DNA have

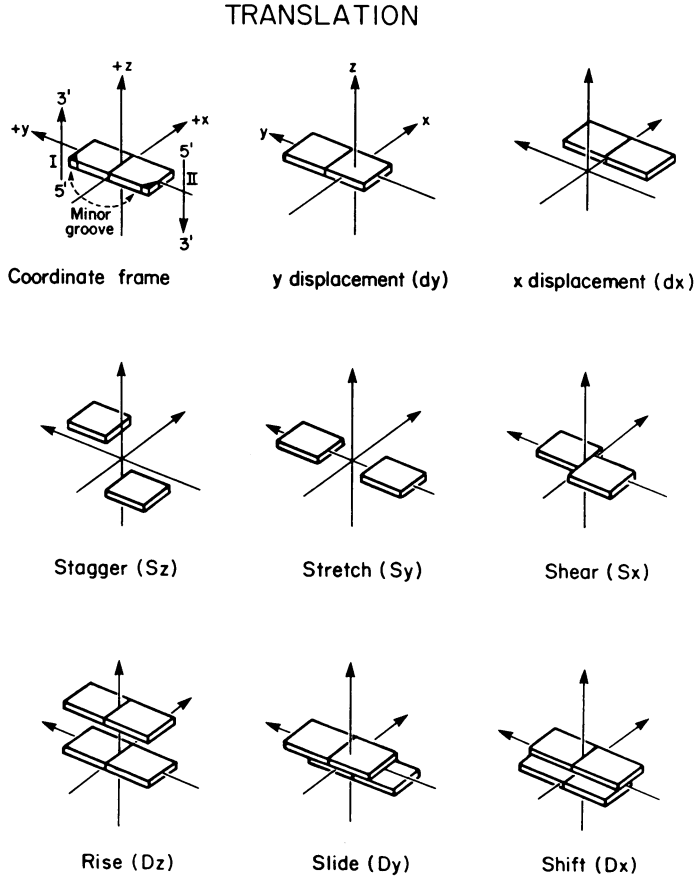

Fig. 8. Definitions of various translations involving two bases of a pair (upper two rows) or two successive base pairs (bottom row). In the top row the motions of the two bases are coordinated, and in the middle row their motions are opposed. Columns at left, centre and right describe translations along the z, y and x axes respectively. The standard coordinate frame is defined at the upper left.

positive inclination, η (Figure 5). Analogy between propeller twist (ω) and the standard torsion sign conventions forced a positive ω to be as shown in Figure 6, reversing common literature practice to date. But a decision was made in favour of consistency, with a suggestion that people obtain the now familiar signs merely by tabulating the quantity $-\omega$.

With these conventions , the signs of helical twist Ω, propeller twist ω, displacement dx, and inclination η, for the three families of DNA double helix typically are as follows:

Helix type	A	B	Z
Helical twist:	+	+	−
Propeller twist:	−	−	~0
x Displacement, dx:	−	~0	+
Inclination:	+	~0	~0

Note that both the propeller twist and x displacement have opposite signs from what has been customary in the past.

(ix) User-friendly input/output is crucial in securing general acceptance of a program library. The user at large is unwilling to invest the same effort in understanding a program as was its author. The Brookhaven atom list format should be used to avoid having to list every individual atom by sequential number, as is done presently in HELIB. The most user-friendly programs probably will have the highest likelihood of eventual acceptance.

The parameters listed and named in section (vi) above obviously do not constitute a complete set of descriptive parameters for polynucleotide structures. A more comprehensive set is found in Lavery and Sklenar(4) and in revised form in Figures 7 and 8. The subset of parameters that are listed in section (vi) above are those that have shown proven utility in past analyses of polynucleotide structures.

Those interested in participating in these helix program trials and finding out how to obtain copies of the new programs when ready should contact the coordinator of the consortium:

Stephan Diekmann,
Abteilung Molekulare Biologie,
Max Planck-Institute fur Biophysikalische Chemie,
Am Faßberg,
D-3400 Gottingen-Nikolausberg, FRG.
Electronic mail address : SDIEKMA@DGOGWDG1

Consortium participants

Richard E. Dickerson, University of California, Los Angeles, USA
Manju Bansal, Indian Institute of Science, Bangalore, India
Christopher R. Calladine, University of Cambridge, UK
Stephan Diekmann, Max Planck-Institute, Gottingen, FRG
William N. Hunter, University of Cambridge, UK
Olga Kennard, University of Cambridge, UK.
Eberhard von Kitzing, Max Planck-Institute, Gottingen, FRG
Richard Lavery, Institut de Biologie Pysico-Chimique, Paris, France
Hillary C.M. Nelson, MRC Laboratory of Molecular Biology, Cambridge, UK
Wilma K. Olson, Rutgers University, USA
Wolfram Saenger, Freie Universitat, Berlin, FRG

Zippora Shakked, Weizmann Institute, Israel
Heinz Sklenar, Academy of Sciences, Berlin, GDR
Dikeos Mario Soumpasis, Max Planck-Institute, Gottingen, FRG
Chang-Shung Tung, Los Alamos National Laboratory, USA
Andrew H.-J. Wang, University of Illinois, Urbana, USA
Victor B. Zhurkin, Academy of Sciences, Moscow, USSR

REFERENCES

1. Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982). J. Biol. Chem. **257**, 14686–14707.
2. Dickerson, R.E. (1985). In F. Jurnak and A. McPherson (eds), Biological Macromolecules and Assemblies: Vol. 2, Nucleic Acids and Interactive Proteins. Wylie, New York, Appendix, pp. 471–494.
3. IUPAC-IUB Joint Commission on Biochemical Nomenclature (1983) Eur. J. Biochem. **131**, 9–15.
4. Lavery, R. and Sklenar, H. (1988) J. Biomol. Struct. Dynam. **6**, 63–91.
5. Soumpasis, D.M. and Tung, C.-S. (1989) J. Biomol. Struct. Dynam., in press.
6. von Kitzing, E. and Diekmann, S. (1987) Eur. Biophys. J. **15**, 13–26.
7. Bhattacharya, D. and Bansal, M. (1988) J. Biomol. Struct. Dynam. **6**, 93–104.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.