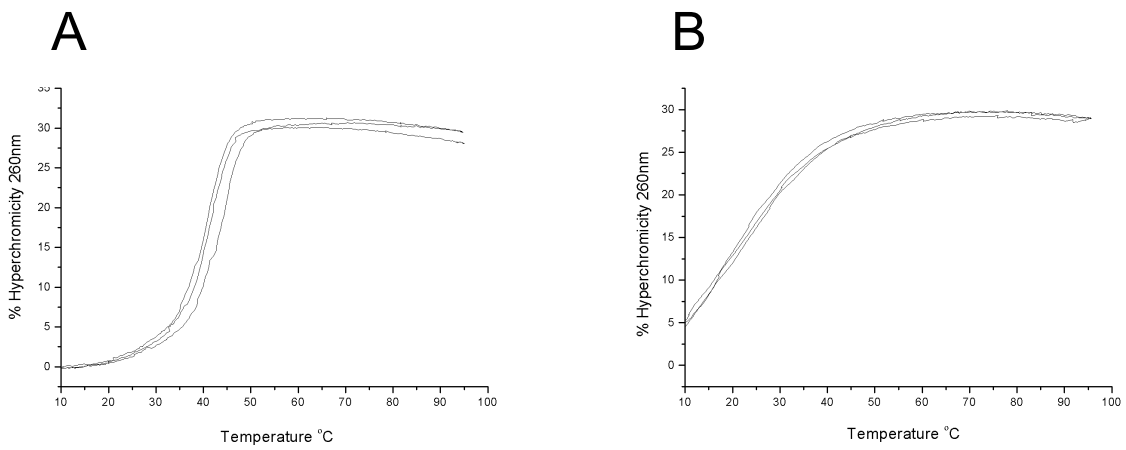
Supporting Information

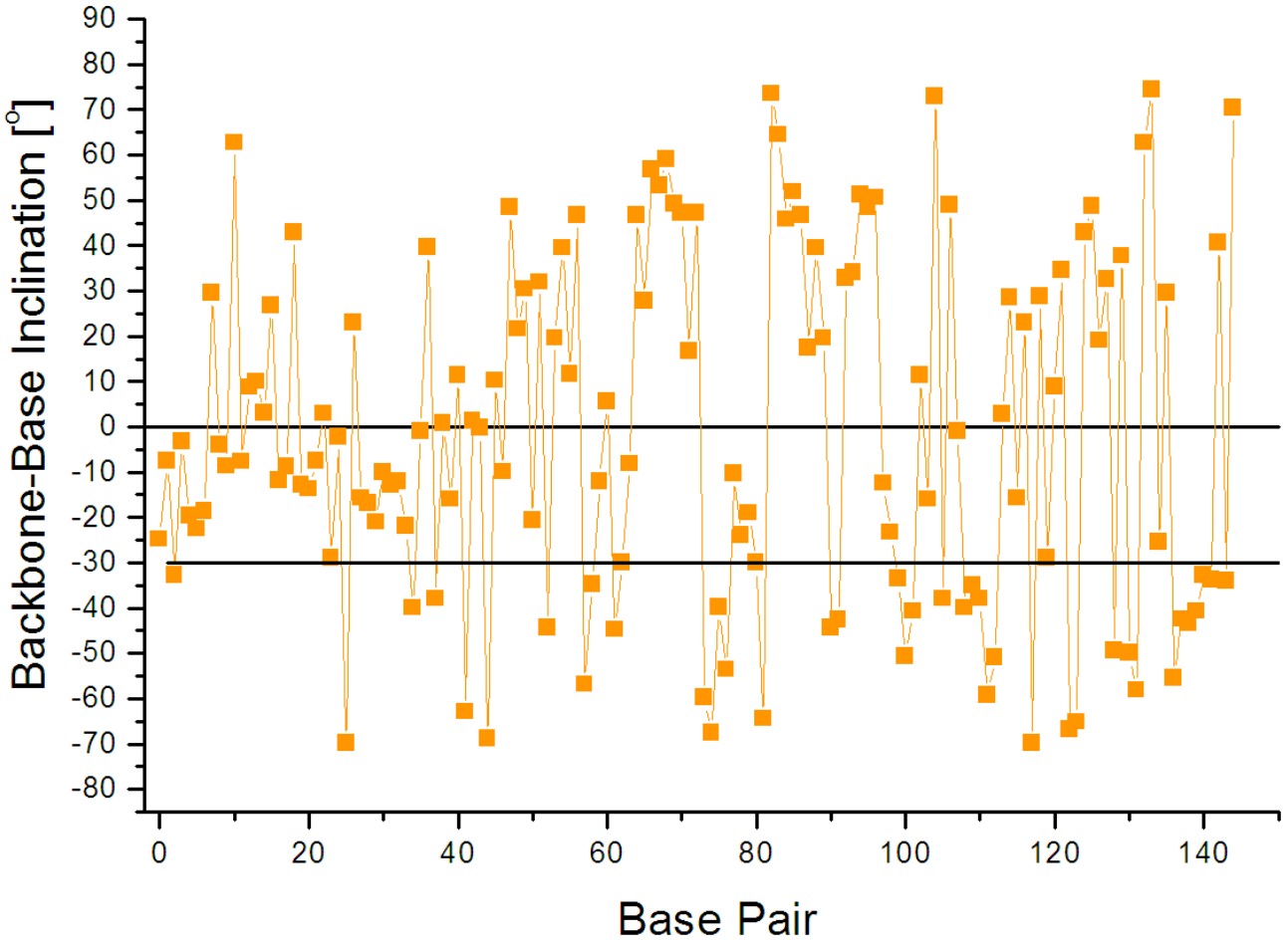
Backbone-base inclination as a fundamental determinant of nucleic acid self- and cross-pairing

Pradeep S. Pallan, Paolo Lubini,[1] Martin Bolli[2] and Martin Egli*

Department of Biochemistry, School of Medicine, Vanderbilt University, Nashville, TN 37232, USA, [1]Alta Scuola Pedagogica, CH-6600, Locarno, Switzerland and [2]Actelion Pharmaceuticals Ltd., CH-4123 Allschwil, Switzerland



UV-melting diagrams for 2'-*O*Me-RNA (A) antiparallel and (B) parallel.



Base pair-backbone inclinations of duplex DNA in the nucleosome core particle (base pairs T71•A71 to A0•T0 to T73•A73).

Program Inclination * Description * Dec. 19, 2006

Version for Silicon Graphics (PASCAL code)

This program computes the backbone-base inclination angle for nucleic acids.

1. First a BSpline curve based on the positions of backbone atoms (P or C1') is calculated. The default value of k is 3 but is variable (see: W.M. Newman, R.F. Sproull, Principles of interactive computer graphic, Second Edition, Mc Graw Hill, chapter 21). For each backbone atom the distance from the BSpline curve is calculated and listed in the output file.

2. The best plane (beta) through base atoms of an individual base or base pair is determined and the point where the BSpline curve intersects with beta is calculated.

3. The tangent to the BSpline curve at the point of intersection is calculated and the tangential vector is then projected onto the plane (alpha) that is normal to beta and contains the point of intersection as well as the center of gravity defined by base atoms.

4. The backbone-base inclination is defined as the angle between the normal vector to beta and the projected tangential vector.

Parameter definitions:

-	P	phosphorus atom in the backbone
-	BPS	BSpline curve points with minimal distance to P atoms
-	K	points that define the BSpline curve
-	PB	point of intersection between beta and BSpline curve
-	T0	tangent to BSpline at PB
-	S	center of gravity of base or base pair
-	B	normal vector to beta, pointing from 5' to 3'
-	PA	projection of tangential vector onto alpha, pointing from 5' to 3'
-	PbPa	projection of tangential vector onto alpha through S
-	Incl. angle	B--S--PbPa
-	Sign	Backbone-base inclination > 0 when dot product between PB -> S and PaPb > 0

****

Run program:

%> do_it.sh input 7 P

(to use C1* atoms for calculating the BSpline curve: %> do_it.sh input 7 C1*)

where:

- input is a PDB-format file with all atomic coordinates (use PDB or NDB code for input)
- 7 corresponds to the number of P atoms in strand 1 that bracket the base pairs; the latter need to be defined in a file input.bp (i.e. for NDB code adh027s; note that terminal base pairs in the octamer have been omitted):

G 2 C 15
G 3 C 14
C 4 G 13
G 5 C 12
C 6 G 11
C 7 G 10

The program recognizes the following bases (see getBP.awk or getB.awk):

A	Adenine
G	Guanine
C	Cytosine
T	Thymine
I	Inosine
U	Uracil

In the case of A, T, G, C, ADE, GUA, THY, CYT will also be accepted.

IMPORTANT (file input.bp): one base pair per line and no empty lines at the end

In the procedure do_it.sh the variable REFD needs to be set; it provides information regarding the location of programs and awk scripts:

- prepPDB.awk: Generates the input file for the program.

- getC.awk: Extracts backbone atoms (P or C1') from the file with the PDB coordinates and produces two files: P1.pdb contains the backbone atoms for strand 1 and P2.pdb contains the corresponding atoms for strand 2.

- BSPabstand: Calculates the distance between backbone atoms and the BSPline curve.

  INPUT: File with backbone atoms in PDB format.
  OUTPUT: Distance P (C1') - BSPline for each backbone atom as well as average distance and standard deviation.

- getBP.awk: Uses the input file with PDB-format coordinates for all atoms and the input.bp file to extract nucleobase atoms. For each base pair a separate file with atomic coordinates is generated (i.e. file name BP_G2-C15.pdb etc.).

- getB.awk: Analogous to getBP.awk, but the generated files only contain atoms of individual bases (i.e. file name BP_G5.pdb etc.).

- best_plane: Calculates the distance of individual base atoms to the best plane beta.

  INPUT: File with coordinates of base atoms.
  OUTPUT: r.m.s. deviations from the best plane.

- dihedral: Calculates the dihedral angles between normal vectors to the best planes through base pairs.

  INPUT1: File with coordinates for atoms of base A.
  INPUT2: File with coordinates for atoms of base B.
  OUTPUT: dihedral angle.

- BSScherung: Calculates the inclination based on the tangential vector to a BSpline curve through P atoms and torsion angles between the base normal vector and the tangent.

  INPUT1: File with coordinates of backbone atoms.
  INPUT2: File with coordinates of base atoms.
  OUTPUT: Inclination angle and torsion angle between normal vector to beta and the projected tangential vector (onto alpha) and the tangential vector, respectively.

- getPP.awk: Extracts from the input file two phosphorus atoms that belong to two successive residues in a 5' -> 3' direction, i.e. P of residue C4 and P of residue G5 from the adh027s NDB file.

- PPScherung: Calculates the inclination angle based on the base normal vector and the P->P vector projected onto alpha as well as the torsion angle between normal vector and P->P vector.

  INPUT1: File with coordinates of base atoms and two backbone atoms (see getPP.awk). The latter have to be at the end of the file.

OUTPUT: Inclination angle and torsion angle between normal vector to beta and the projected P->P vector (onto alpha) and the P->P vector, respectively.

- scherung.sh: Shell script. Contains all the information for executing the program and produces a file named q with the calculated parameters for each base pair or base.

- analyse.sh: Shell script. Contains information for extracting the calculated parameters that are stored in file q and produces a file result_input.

- evaluate: Program that is initiated with analyse.sh and computes average values and standard deviations.

Description of file result_input:

*** S C H E R U N G ***

INPUT File : adh027s

Backbone atoms : P

BSpline constant : 3

Base plane

Base pair	rms AB	rms A	rms B	Tau
G2 - C15	0.17	0.01	0.01	9.56
G3 - C14	0.18	0.00	0.00	13.12
C4 - G13	0.10	0.01	0.01	6.91
G5 - C12	0.10	0.01	0.01	6.91
C6 - G11	0.18	0.00	0.00	13.12
C7 - G10	0.17	0.01	0.01	9.56
Mean	0.15	0.01	0.01	9.86
Sig	0.04	0.00	0.00	2.79

where: rms AB: r.m.s. distance for atoms of base pair A:B from best plane (i.e. line 1: A=G2 and B=C15 etc.)
Tau: Dihedral angle between best planes through bases A and B

BSpline inclination angles

Base pair	AB1	A1	B1	AB2	A2	B2	Tau1	Tau2
G2 - C15	-	-	-	45.58	47.21	50.29	-	-65.51
G3 - C14	29.73	29.13	33.05	45.00	46.45	44.39	-52.03	-55.64
C4 - G13	24.44	32.52	23.25	25.73	16.37	28.12	-52.65	-54.46
G5 - C12	25.73	28.12	16.37	24.44	23.25	32.52	-54.46	-52.65
C6 - G11	45.00	44.39	46.45	29.73	33.05	29.13	-55.64	-52.03
C7 - G10	45.58	50.29	47.21	-	-	-	-65.51	-
Mean	34.10	36.89	33.27	34.10	33.27	36.89	-56.06	-56.06
Sig	10.41	9.90	13.73	10.41	13.73	9.90	5.48	5.48

where: AB1: Inclination angle for base pair A:B based on a BSpline curve through phosphorus atoms of strand 1
A1, B1: Inclination angles for bases A and B based on a BSpline curve through phosphorus atoms of strand 1
AB2: Inclination angle for base pair A:B based on a BSpline curve through phosphorus atoms of strand 2
A2, B2: Inclination angles for bases A and B based on a BSpline curve through phosphorus atoms of strand 2
Tau1: Torsion angles between base normals and tangential vectors to the BSpline curve through phosphorus atoms of strand 1 at the points of intersection between curve and base planes.
Tau2: Torsion angles between base normals and tangential vectors to the BSpline curve through phosphorus atoms of strand 2 at the points of intersection between curve and base planes.

Backbone inclination angles

Base pair	sA1	sAB1	sB2	sAB2	pp1AB	pp2AB
G2 - C15	45.80	44.27	55.63	54.47	53.17	63.72
G3 - C14	39.29	42.20	53.70	56.57	51.43	62.40
C4 - G13	41.06	35.10	44.31	45.58	54.18	54.18
G5 - C12	44.31	45.58	41.06	35.10	54.18	54.18

C6 - G11	53.70	56.57	39.29	42.20	62.40	51.43
C7 - G10	55.63	54.47	45.80	44.27	63.72	53.17
Mean	46.63	46.36	46.63	46.37	56.51	56.51
Sig	6.66	7.99	6.66	7.99	5.19	5.19

Where: sA1: Inclination angle for base A based on the P->P vector of strand 1 (see getPP.awk)
sAB1: Analogous to sA1 but for base pair A:B
sB2: Inclination angle for base A based on the P->P vector of strand 2 (see getPP.awk)
sAB2: Analogous to sA2 but for base pair A:B
pp1AB: Angle between base normal and P->P vector from strand 1
pp2AB: Angle between base normal and P->P vector from strand 2

Distance backbone atoms BSpline curve

Distance: 42	P	G	3	0.5300
Distance: 64	P	C	4	0.2185
Distance: 83	P	G	5	0.5823
Distance: 105	P	C	6	0.2583
Distance: 124	P	C	7	0.3638

Average distance :	0.3906
Standard deviation:	0.1442

Distance: 204	P	G	11	0.5300
Distance: 226	P	C	12	0.2185
Distance: 245	P	G	13	0.5823
Distance: 267	P	C	14	0.2583
Distance: 286	P	C	15	0.3638

Average distance :	0.3906
Standard deviation:	0.1442