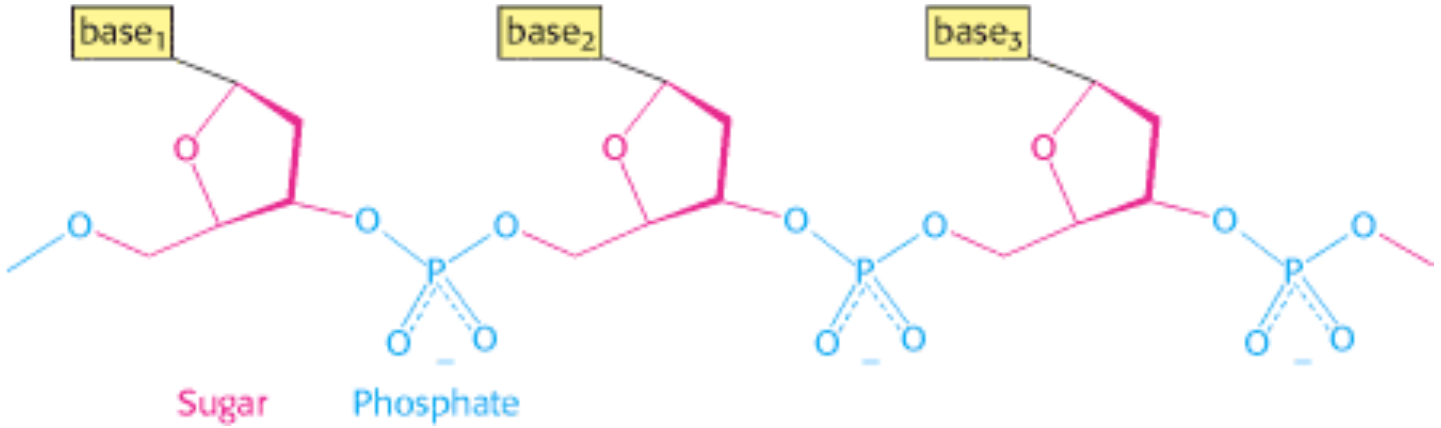
Nucleic acid structural principles

September 29, 2009

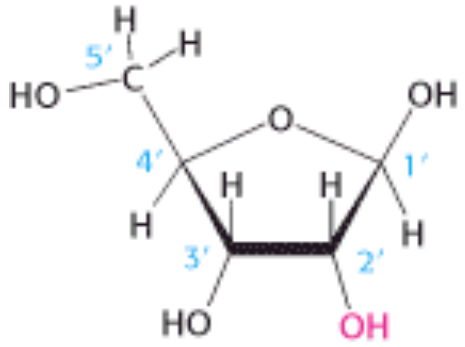
Professor Wilma K. Olson

Chemical make-up of nucleic acids

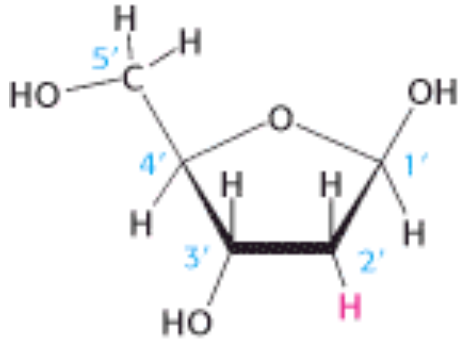
Nucleic acids are linear polymers made up of concatenated sugars, phosphates, and bases. The sugars and phosphates alternate along the chain backbone and the bases are laterally attached to the sugars.



The furanose sugars are of two types:
ribose in RNA and 2′-deoxyribose in DNA.

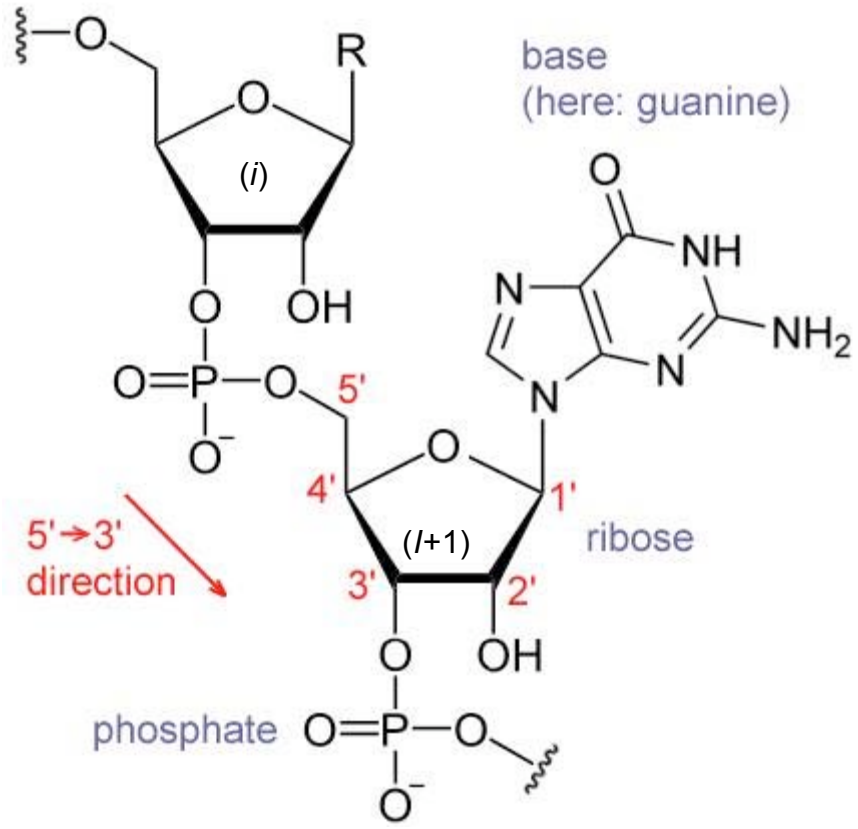


Ribose

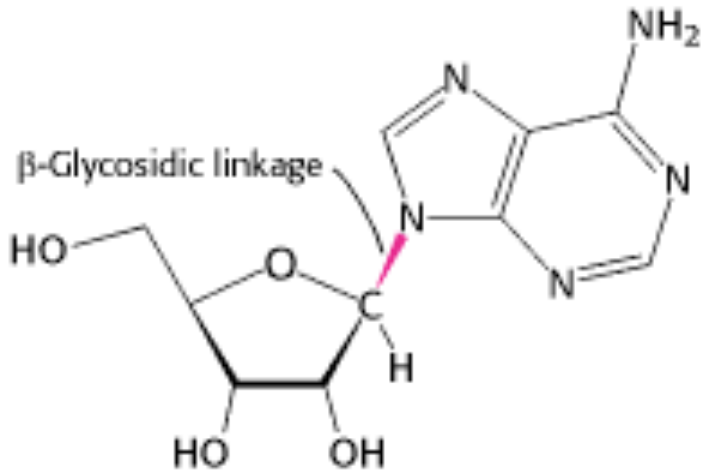


Deoxyribose

The phosphodiester linkage is directional.
The 3´-oxygen of nucleotide *i* is joined to the 5´-oxygen of nucleotide *i*+1.

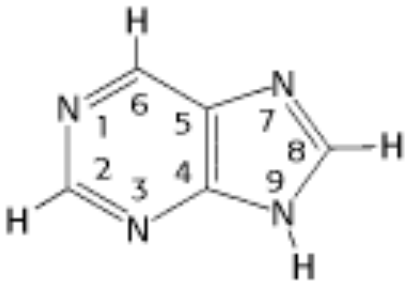


The sugar-base (glycosidic) linkage is stereo specific. The base is attached to the same side of the sugar ring as the exocyclic C5´ atom.

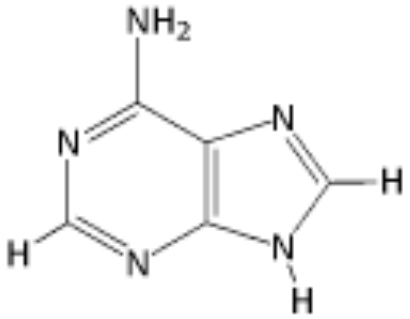


The heterocyclic bases fall into two categories: purines (R = *A* or *G*) and pyrimidines (Y = T/U or *C*).

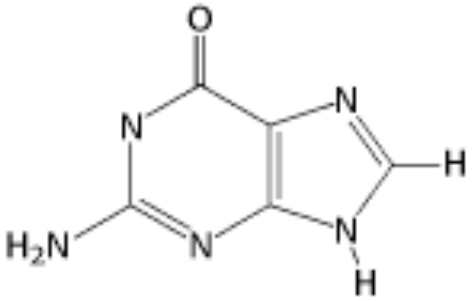
PURINES



Purine

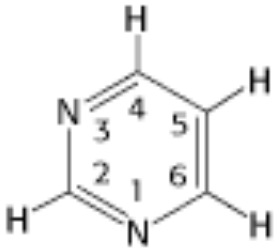


Adenine

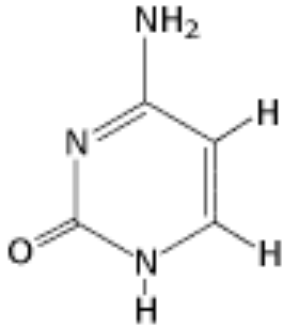


Guanine

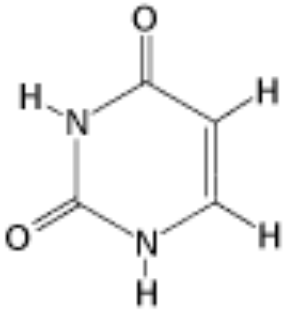
PYRIMIDINES



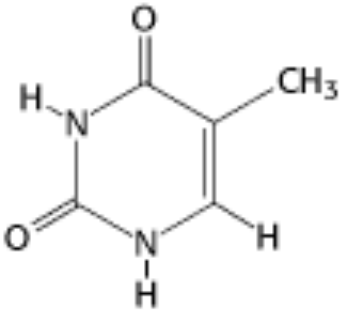
Pyrimidine



Cytosine



Uracil

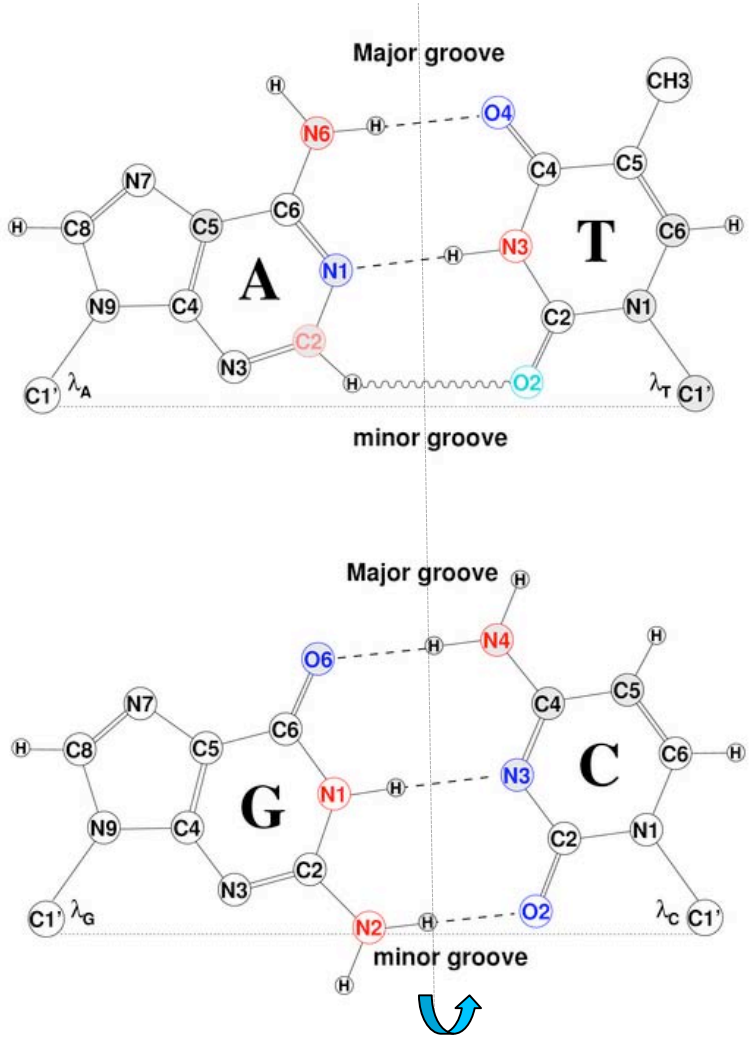


Thymine

The 5-methyl group of T in DNA is replaced by H in RNA.

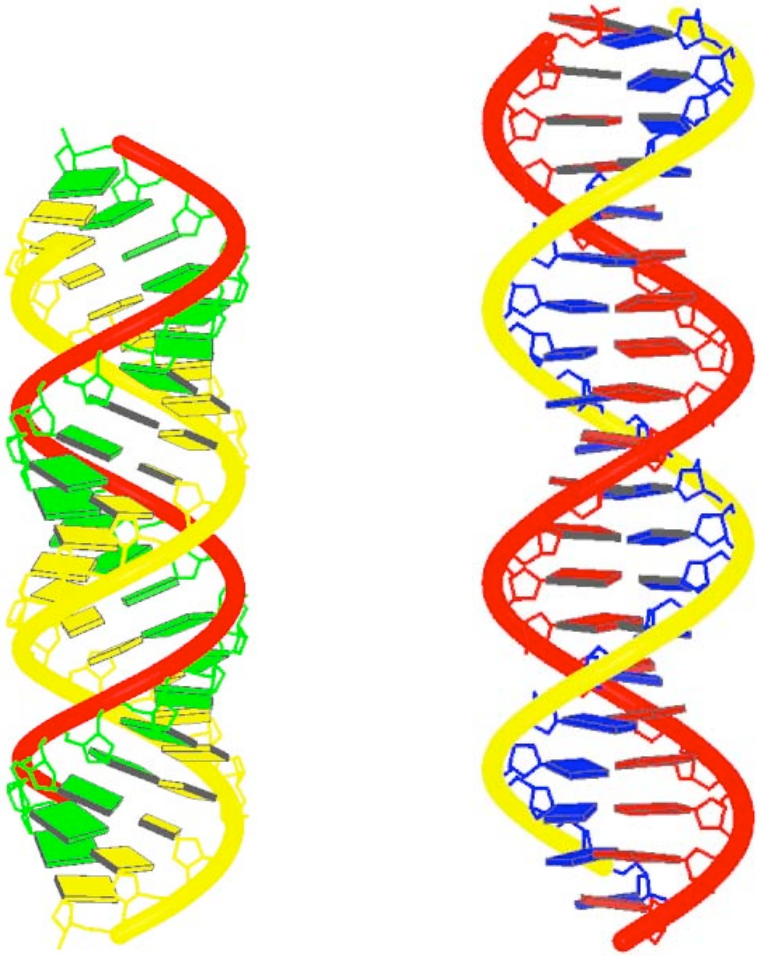
Watson-Crick hydrogen-bonding and double-helical DNA structure

The heterocyclic bases associate as hydrogen-bonded pairs, the most common of which are the canonical Watson-Crick A·T (A·U) and G·C pairs.



The comparable size of the R·Y pairs makes them isosteric, allowing for their interchange and rearrangement in nucleic acid structures.

The regular repetition of paired nucleotide units generates double-helical structures, such as the right-handed A and B forms.

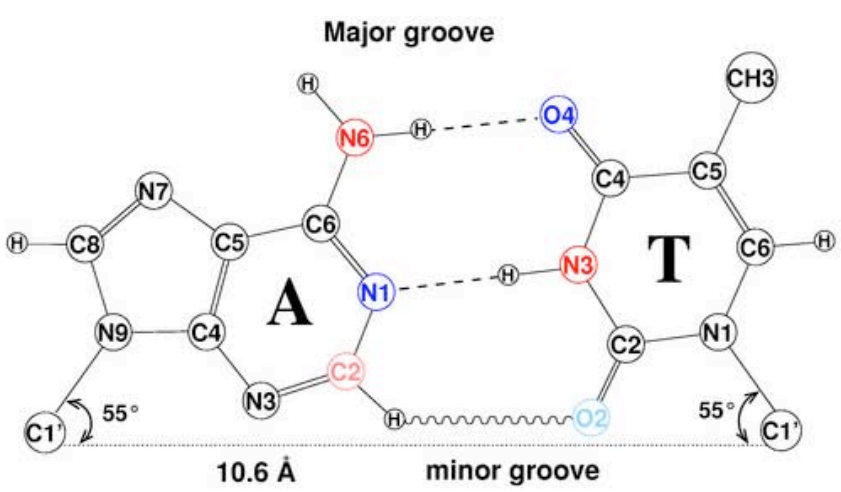


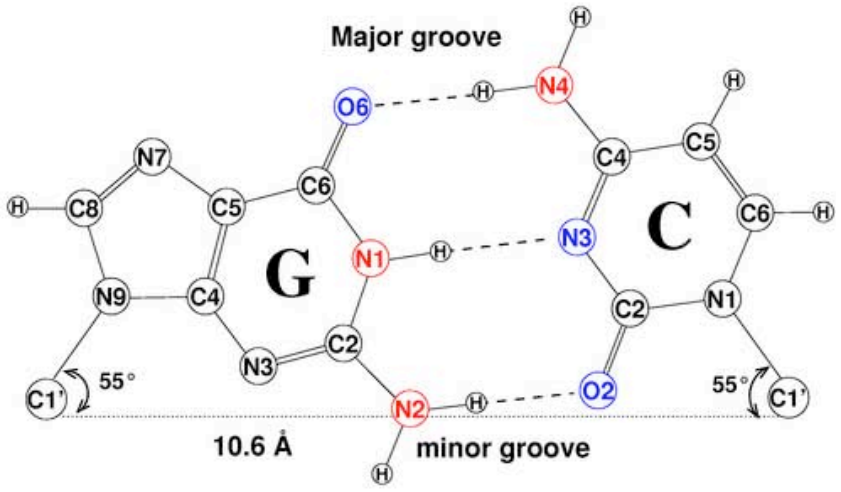
$dG_{20}{\cdot}dC_{20}$ and $dA_{20}{\cdot}dT_{20}$ in canonical A and B forms.

w3DNA.rutgers.edu

The hydrogen bonding distances are independent of double-helical form.

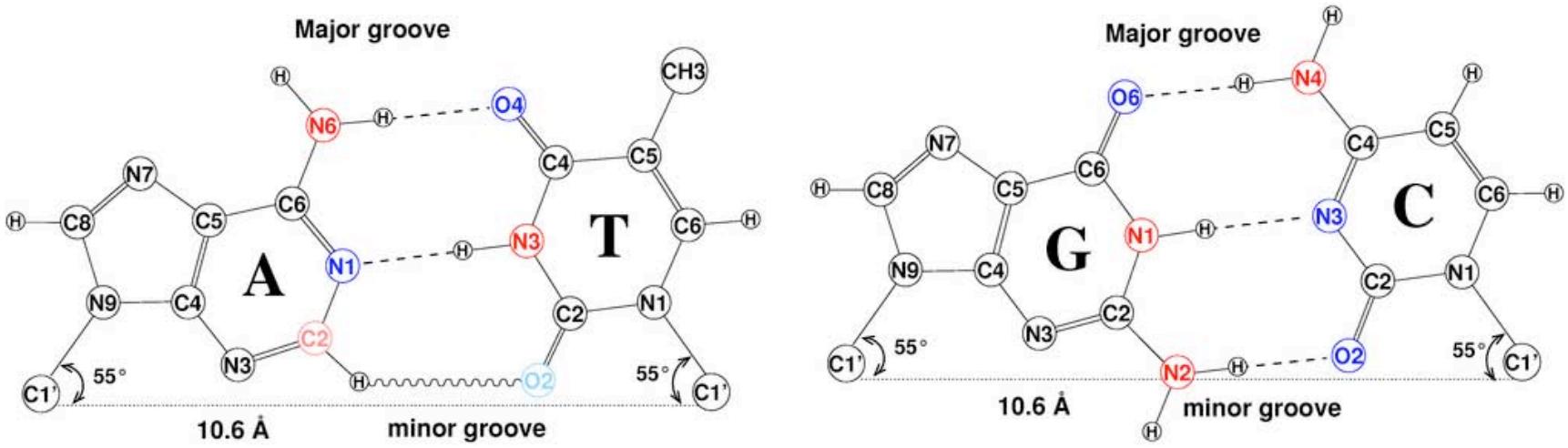
		N6···O4	N1···N3	O6···N4	N2···O2
A·T	A-DNA	2.90	2.88		
	B-DNA	2.95	2.88		
G·C			2.85	2.83	2.81
			2.85	2.85	2.85





The interstrand virtual distances and angles between the paired bases are also independent of helical form.

	λ_R	λ_y	C1′ ···C1′
A DNA	54.3	54.3	10.7
B DNA	54.2	54.2	10.7



The intrastrand virtual distances between successive P and *C1´* atoms along the same strand differ in the two helical forms.

	P···P	*C1´ ···C1´*
A DNA	5.5	5.4
B DNA	6.6	4.9

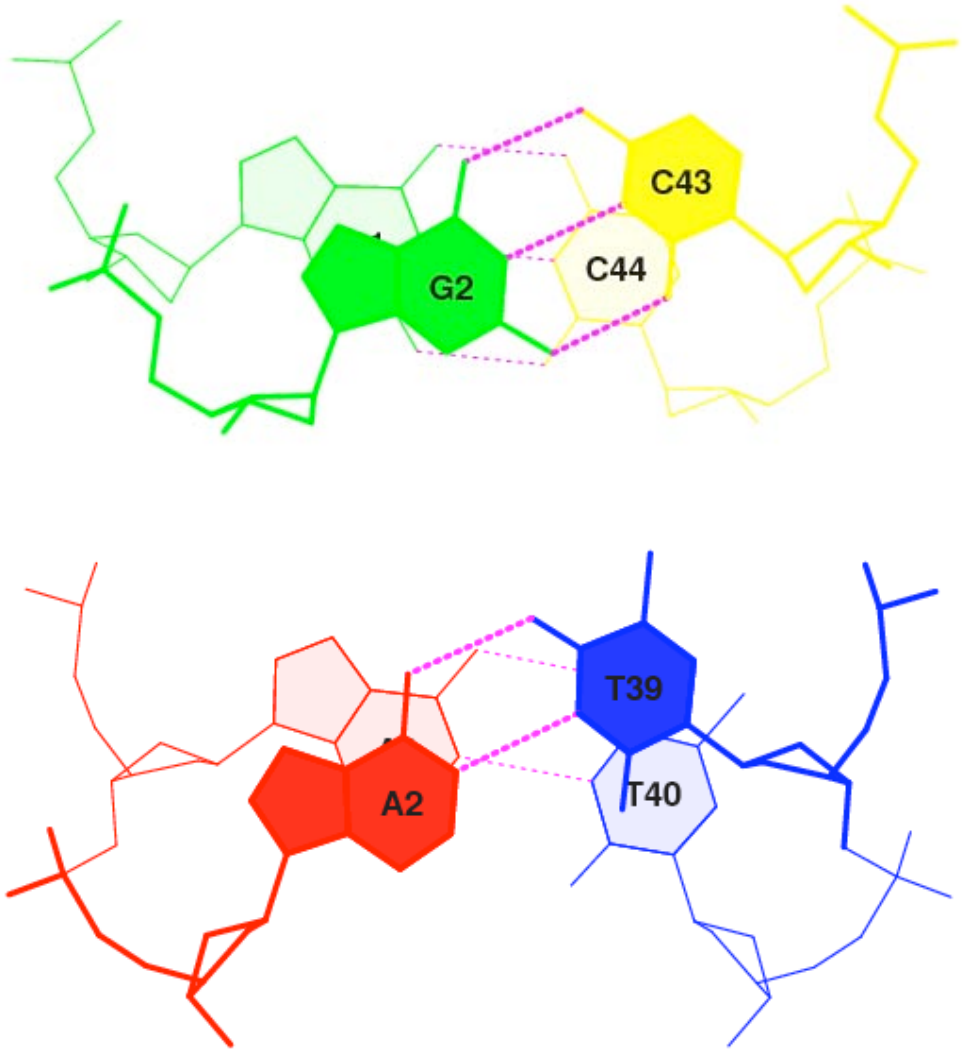
The cylindrical (helical) parameters differ in the two forms.

	base-pair inclination	helical twist	helical rise	*n*
A DNA	20.6	32.7	2.6	11
B DNA	–0.2	36	3.4	10

Groove widths and depths also differ in the two helical forms.

	minor groove	Major groove	$r_{C1'}$
A DNA	16.7	11.1	6.9
B DNA	11.7	17.2	1.9

The overlap of successive base pairs depends on duplex form.



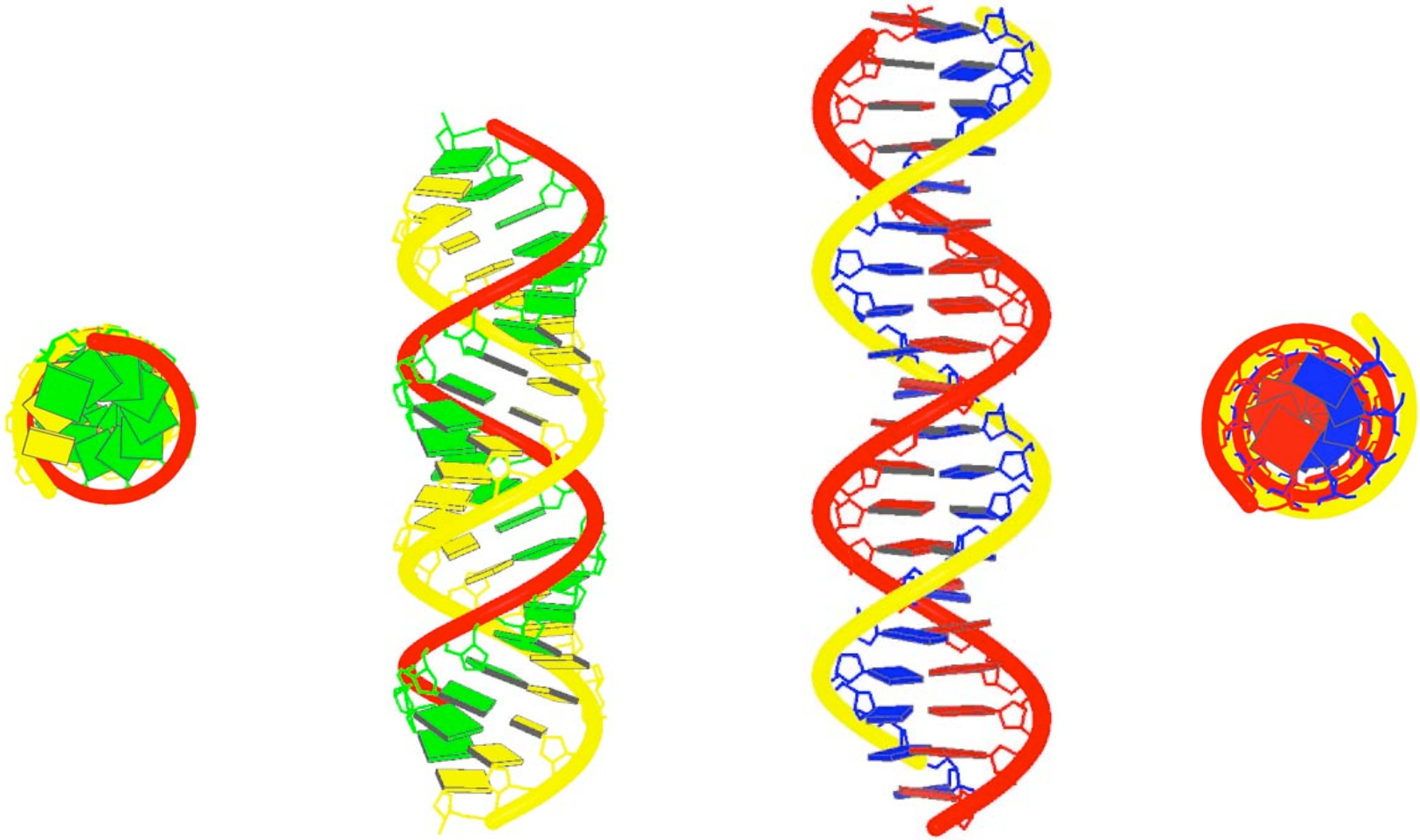
Top-down "stacking diagrams" of dG_2·dC_2 and dA_2·dT_2 units in canonical A and B forms.

Whereas the overlap of base rings is comparable, the overlap of side groups differs in the two helical forms.

Ring overlap	i_{R1}-i_{R2}	i_{R1}-j_{Y2}	j_{Y1}-i_{R2}	j_{Y1}-j_{Y2}	total
A DNA	2.4	0.0	0.0	0.1	2.4
B DNA	2.0	0.0	0.0	0.2	2.2

"All-atom" overlap	i_{R1}-i_{R2}	i_{R1}-j_{Y2}	j_{Y1}-i_{R2}	j_{Y1}-j_{Y2}	total
A DNA	3.8	0.0	0.0	1.0	4.8
B DNA	3.6	0.0	0.0	5.8	9.4

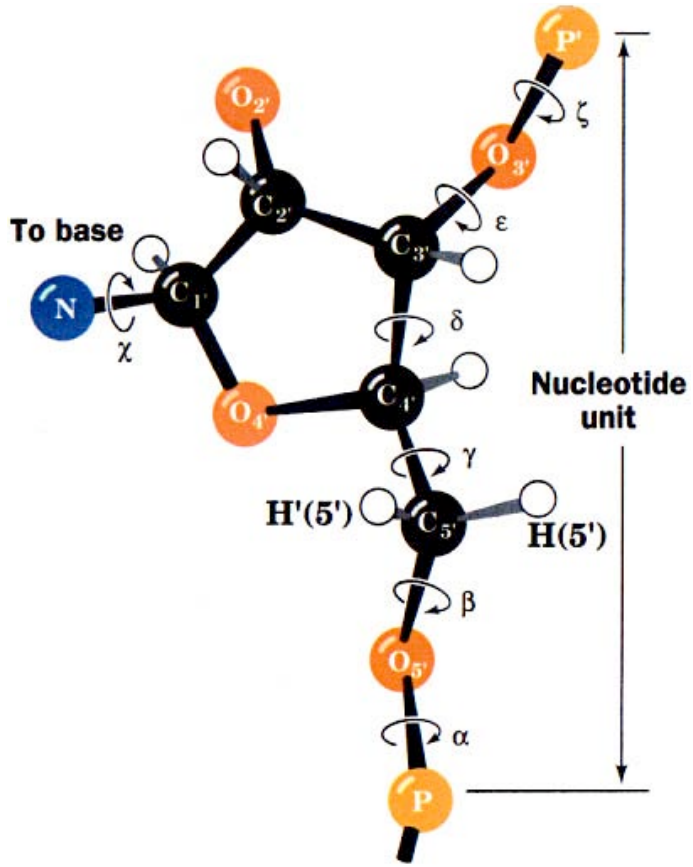
The differences in A vs. B groove widths, base-pair displacement and inclination, base-stacking overlap, and residues per turn are evident in molecular models.



$dG_{20}{\cdot}dC_{20}$ and $dA_{20}{\cdot}dT_{20}$ in canonical A and B forms.
w3DNA.rutgers.edu

Torsional preferences in double-helical A- and B-DNA structures

Nucleotide conformation is defined by seven torsion angles.



The canonical A- and B-DNA structures show large differences in three of the seven repeated nucleotide torsion angles.

	α	β	γ	δ	ε	ζ	χ
A DNA	-52	175	42	79	-148	-75	-157
B DNA	-30	136	31	143	-141	-161	-98

α: O3'(i-1)-P-O5'-C5'
β: P-O5'-C5'-C4'
γ: O5'-C5'-C4'-C3'
δ: C5'-C4'-C3'-O3'
ε: C4'-C3'-O3'-P(i+1)
ζ: C3'-O3'-P(i+1)-O5'(i+1)

χ pyrimidines(Y): O4'-C1'-N1-C2
χ purines (R): O4'-C1'-N9-C4

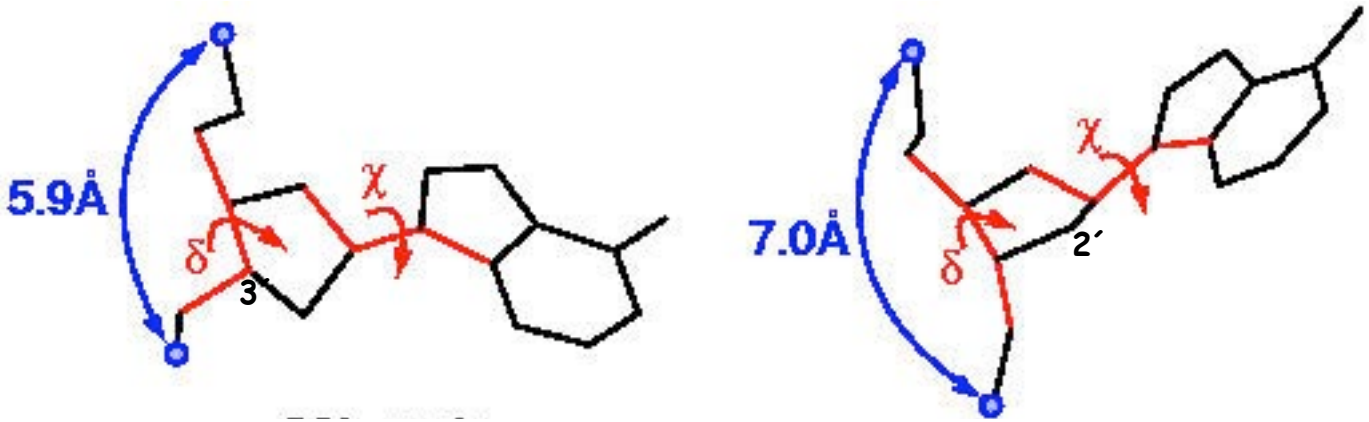
The sugar ring adopts two distinct conformational states (N or S, C3′-endo or C2′-endo) in the canonical A- and B-DNA duplexes.

	ν_0	ν_1	ν_2	ν_3	ν_4	P	τ_m
A DNA	8	-34	44	-40	21	8	44.5
B DNA	-33	45	-40	23	6	154	44.7

ν_0: C4′-O4′-C1′-C2′
ν_1: O4′-C1′-C2′-C3′
ν_2: C1′-C2′-C3′-C4′
ν_3: C2′-C3′-C4′-O4′
ν_4: C3′-C4′-O4′-C1′

τ_m: pseudorotation amplitude
P: pseudorotation phase angle

The differences in the sugar-base torsion angles (the backbone sugar torsion δ or the pseudorotation parameters *P* and τ_m and the glycosyl torsion χ) give rise to characteristic intrastrand P···P distances that distinguish A from B DNA.

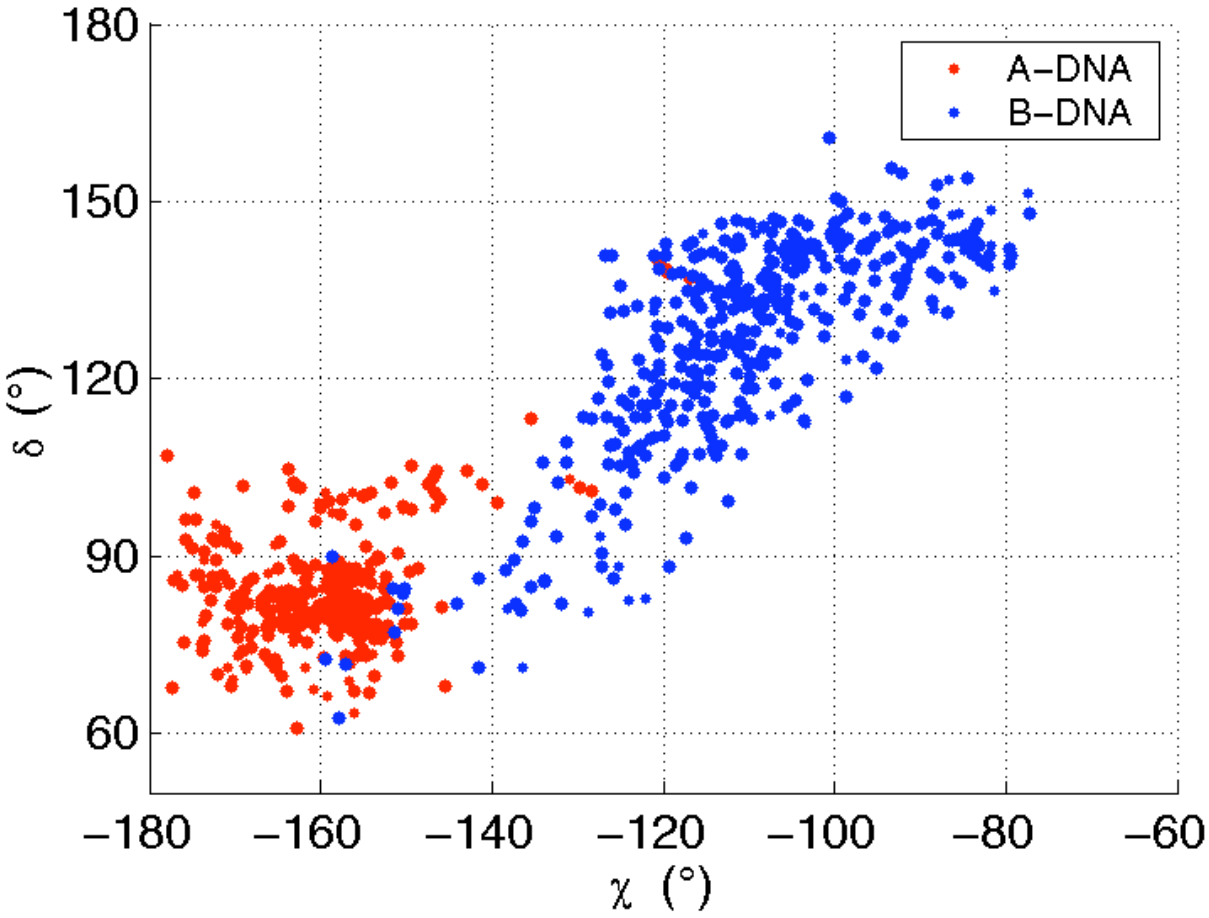


C3´-endo
$P \approx \pi/10$

C2´-endo
$P \approx 9\pi/10$

P··P distances cited here are average values found in high-resolution crystal structures.

The sugar and glycosyl torsion angles are the best chemical-level discriminators of high resolution A-DNA and B-DNA structures.



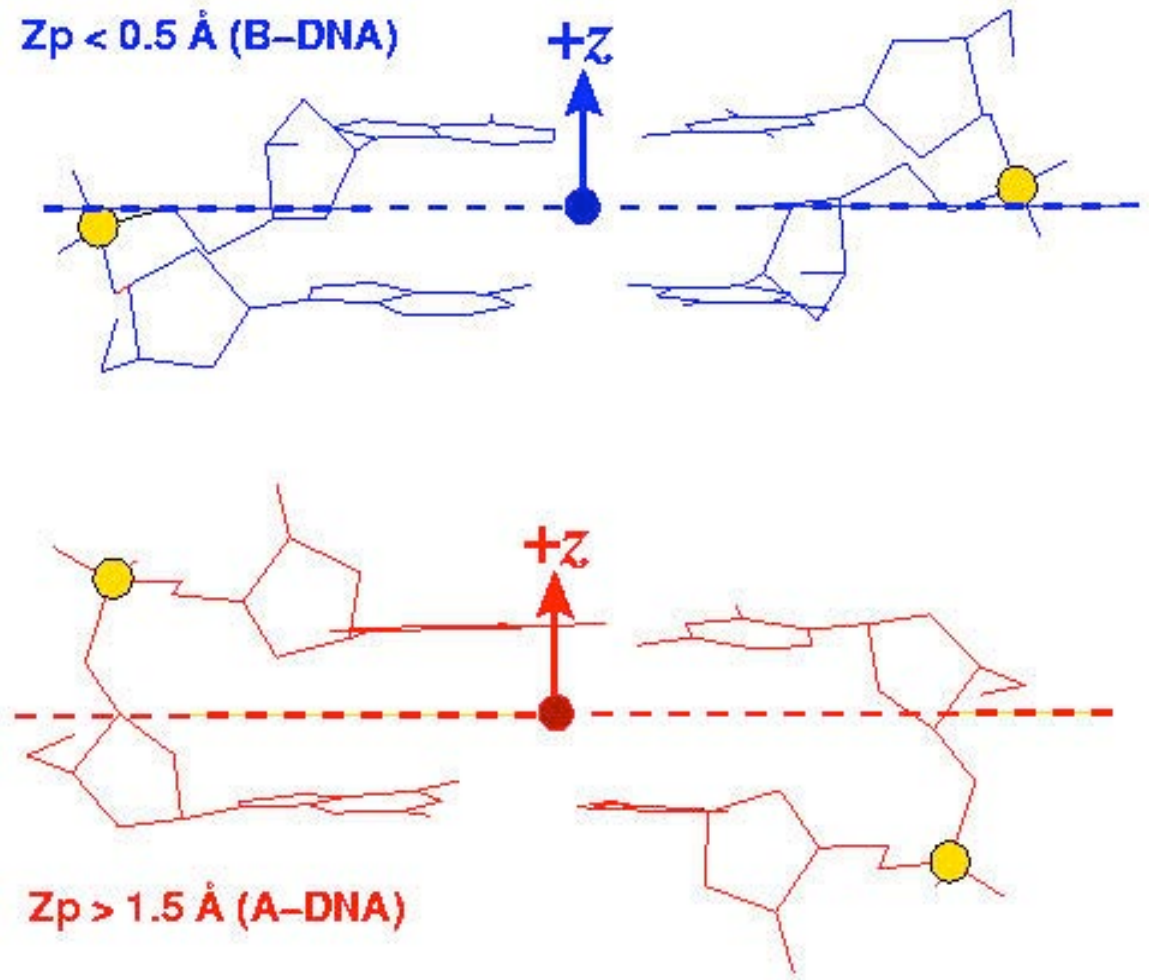
Lu *et al.* (2000) "A-form conformational motifs in ligand-bound DNA structures," *J. Mol. Biol.* **300**, 819-840.

The P atoms lie in two distinct locations in A and B duplex "steps".

	x_P	y_P	z_P
A DNA	–1.0	8.4	2.5
B DNA	–3.0	8.9	–0.6

Mean coordinates of P atoms
in the local dimer frames

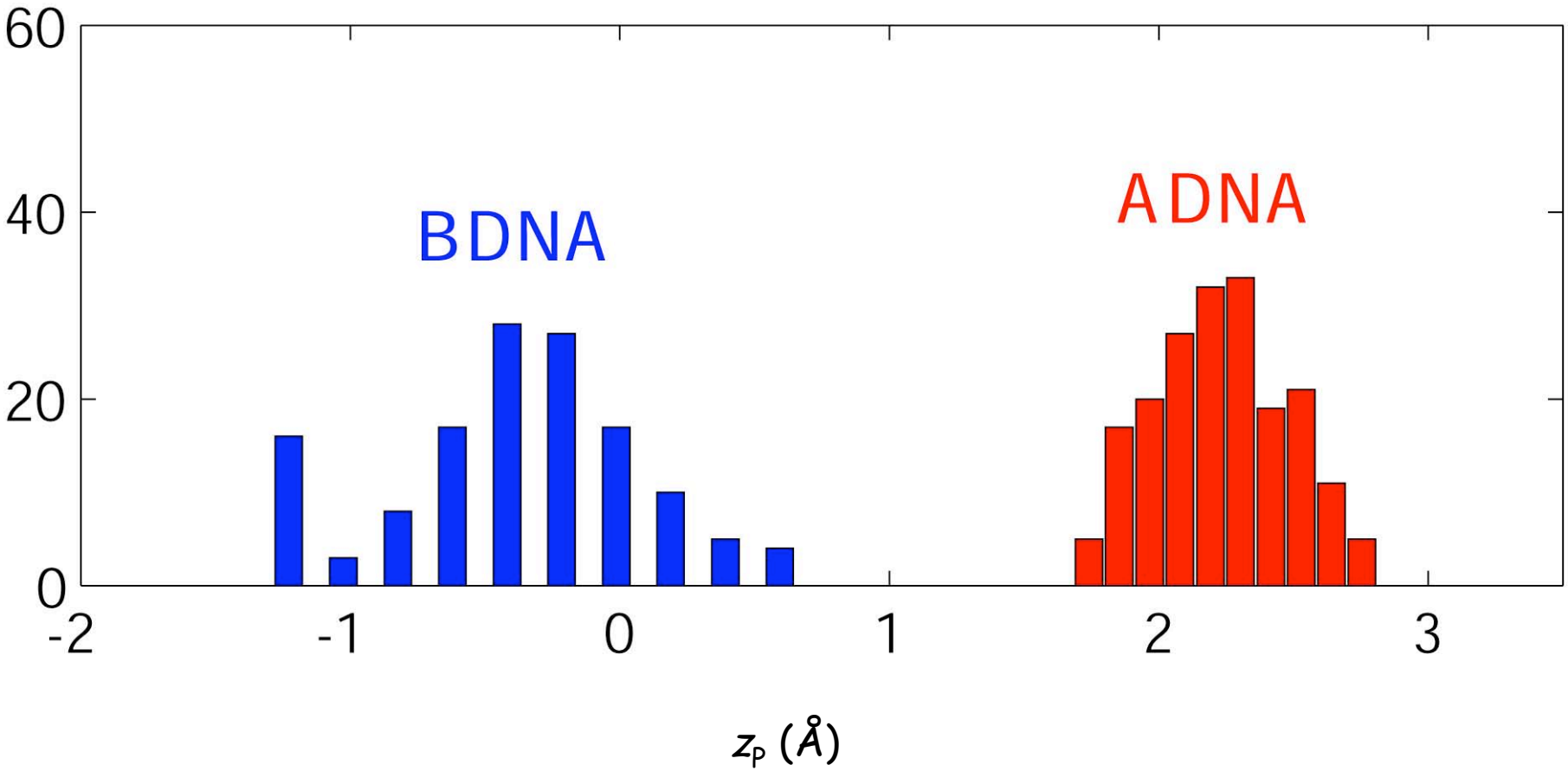
Phosphorus displacement (z_P) differs in A-DNA and B-DNA dimer steps.



A-DNA: GG·CC step from d(GCCCGGGC)$_2$ (adh038)
B-DNA: AA·TT step from d(CGCGAATTCGCG)$_2$ (bdl084)

Lu *et al.* (2000)

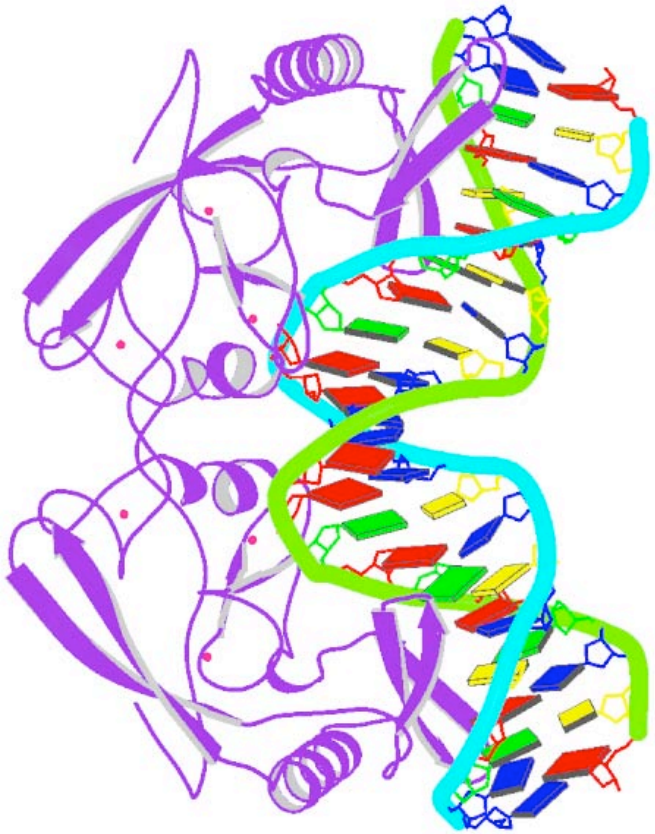
Phosphorus displacement discriminates A-DNA vs. B-DNA base-pair "steps"
(Histograms of observed values of z_P in high-resolution structures)



Lu *et al.* (2000)

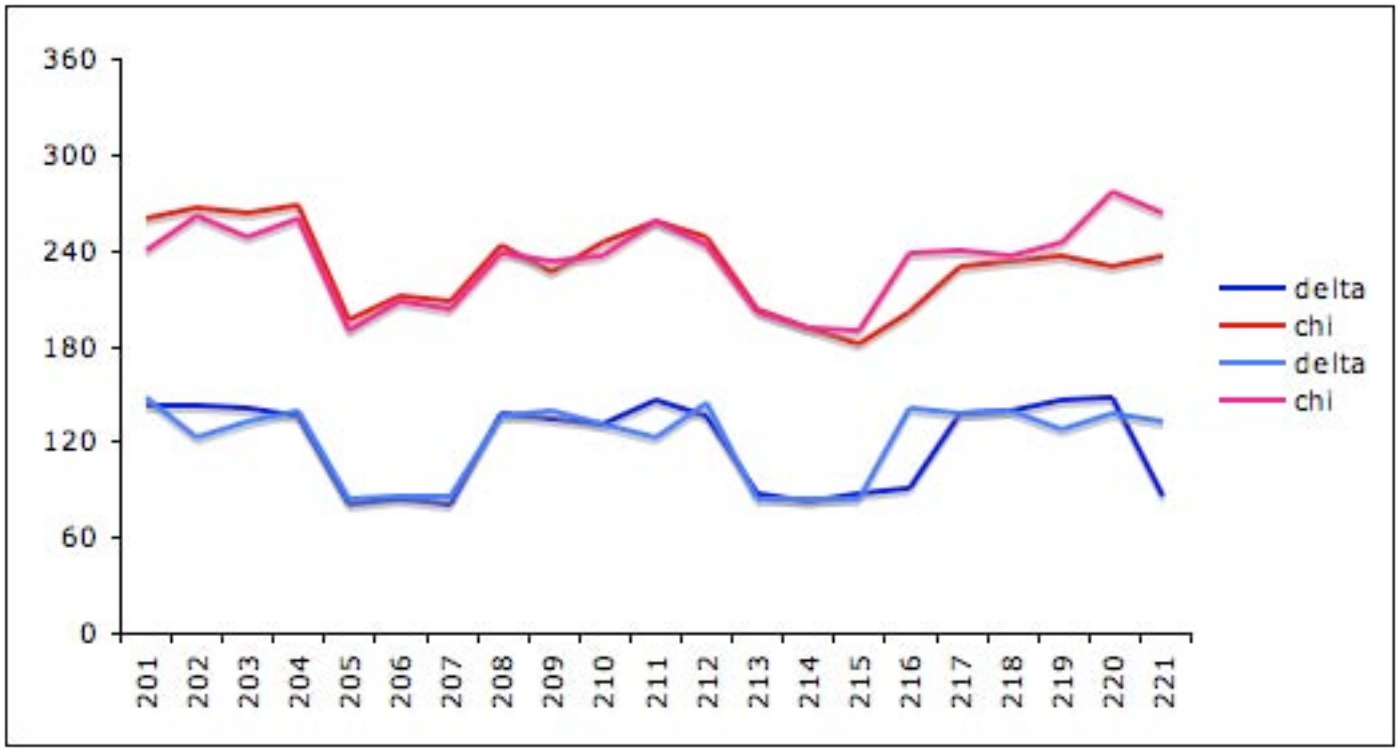
Mechanisms of DNA bending

Proteins often bend DNA without disruption of the double-helical structure.
Homing endonuclease I-*Ppo*I bend DNA by ~60° (PDB_ID: 1ipp).



Flick *et al.* (1998). "DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-*Ppo*I." *Nature* 394, 96–101.

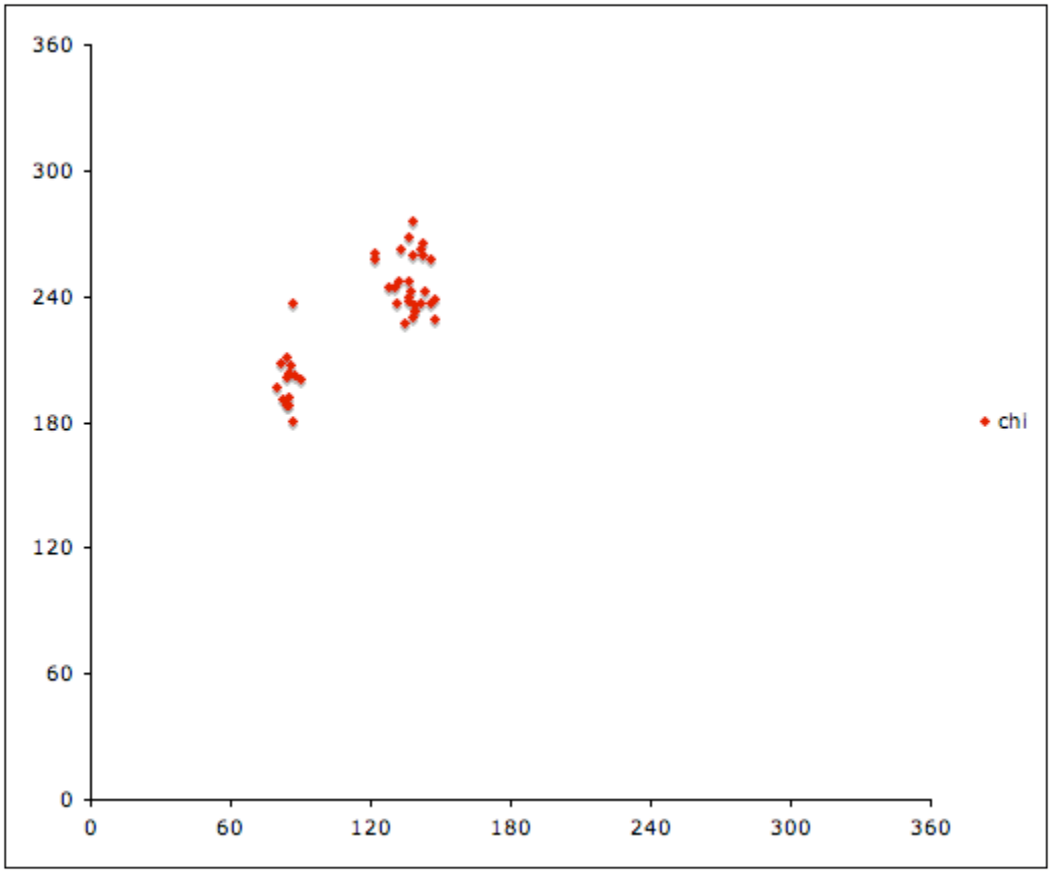
The sugar ring and glycosyl rotations appear to interconvert between A- and B-like forms in this complex.



δ: C5'-C4'-C3'-O3' (79° A-DNA vs. 143° B-DNA)

χ pyrimidines(Y): O4'-C1'-N1-C2
χ purines (R): O4'-C1'-N9-C4
(203° A-DNA vs. 262° B-DNA)

The sugar ring and glycosyl torsions are strongly coupled.



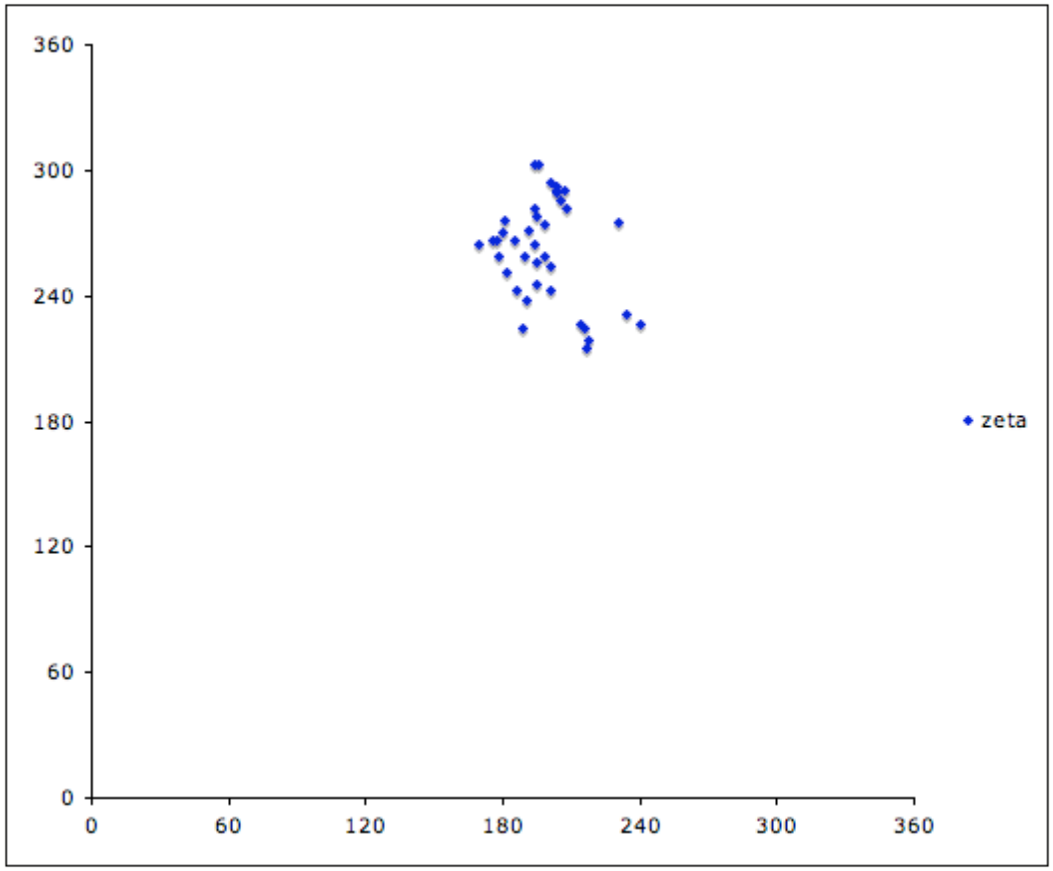
δ: C5'-C4'-C3'-O3' (79° A-DNA vs. 143° B-DNA)

χ pyrimidines(Y): O4'-C1'-N1-C2
χ purines (R): O4'-C1'-N9-C4
(203° A-DNA vs. 262° B-DNA)

One of the phosphodiester rotations (ζ) also appears to interconvert between A- and B-like forms (but in an opposite sense to δ and χ).

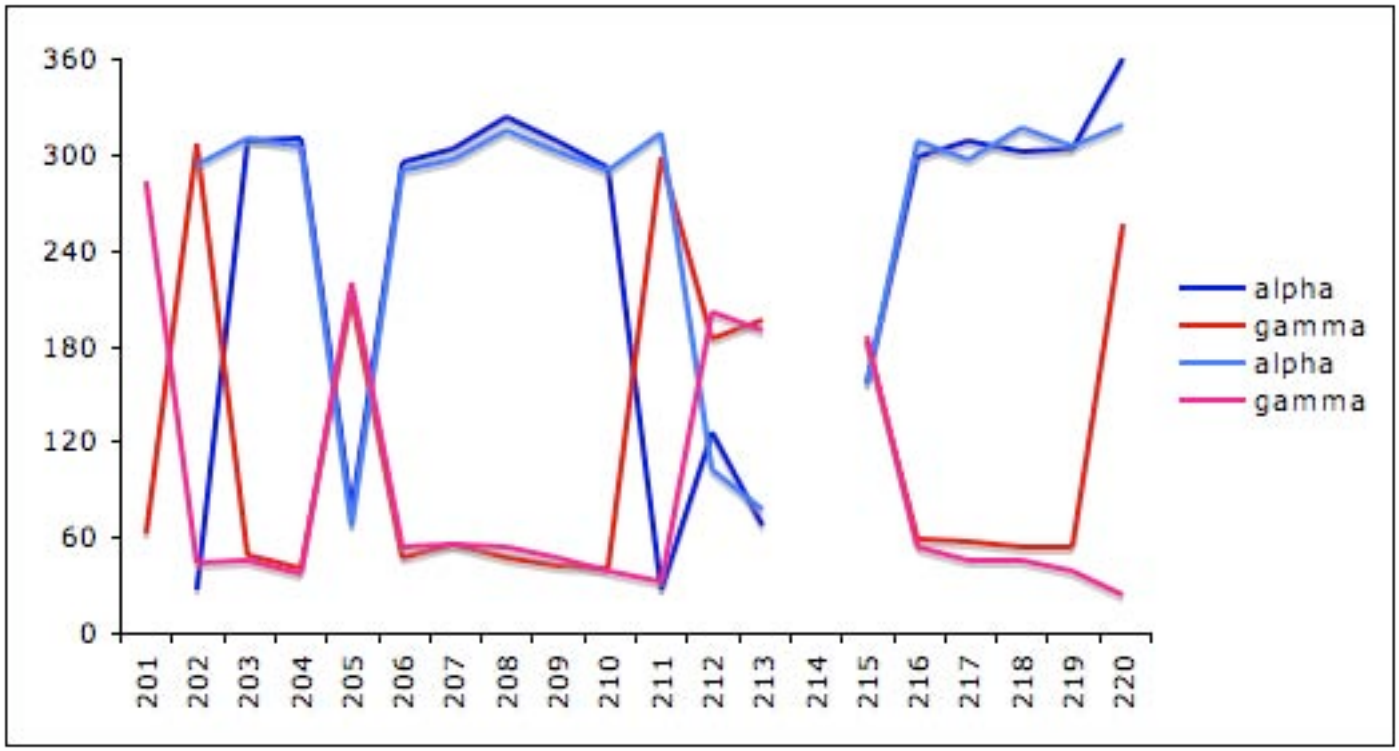
ε: C4'-C3'-O3'-P(i+1)
ζ: C3'-O3'-P(i+1)-O5'(i+1)
(285° A-DNA vs. 161° B-DNA)

The εζ angle pair exhibits slight coupling.



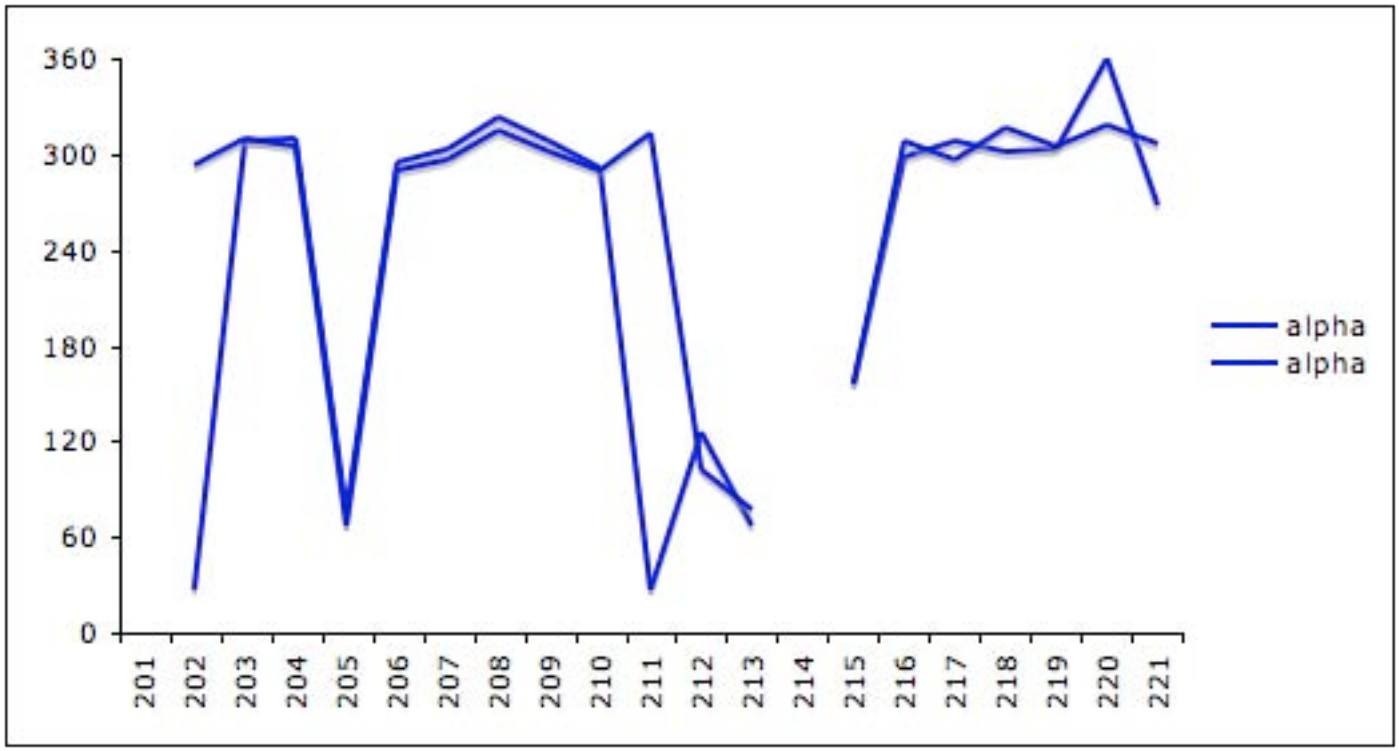
ε: C4'-C3'-O3'-P(i+1)
ζ: C3'-O3'-P(i+1)-O5'(i+1)
(285° A-DNA vs. 161° B-DNA)

Although the α and γ angles adopt similar values in the canonical *A* and B helices, they show large coupled changes in the I-*Ppo*I-DNA complex.



α: O3'(i-1)-P-O5'-C5´(308° *A*-DNA vs. 330° B-DNA)
γ: O5'-C5'-C4'-C3' (42° *A*-DNA vs. 31° B-DNA)

Although the α and γ angles adopt similar values in the canonical *A* and B helices, they show large coupled changes in the I-*Ppo*I-DNA complex.

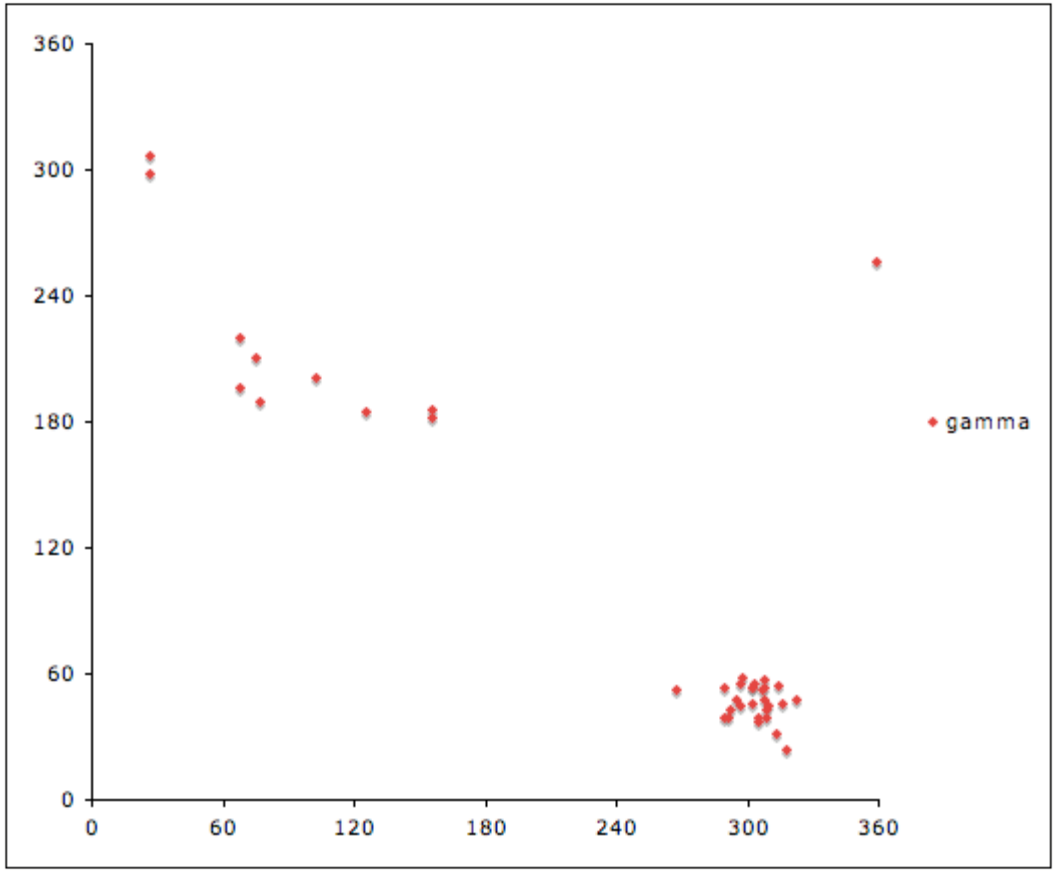


α: O3'(i-1)-P-O5'-C5´ (308° *A*-DNA vs. 330° B-DNA)

Although the α and γ angles adopt similar values in the canonical *A* and B helices, they show large coupled changes in the I-*Ppo*I-DNA complex.

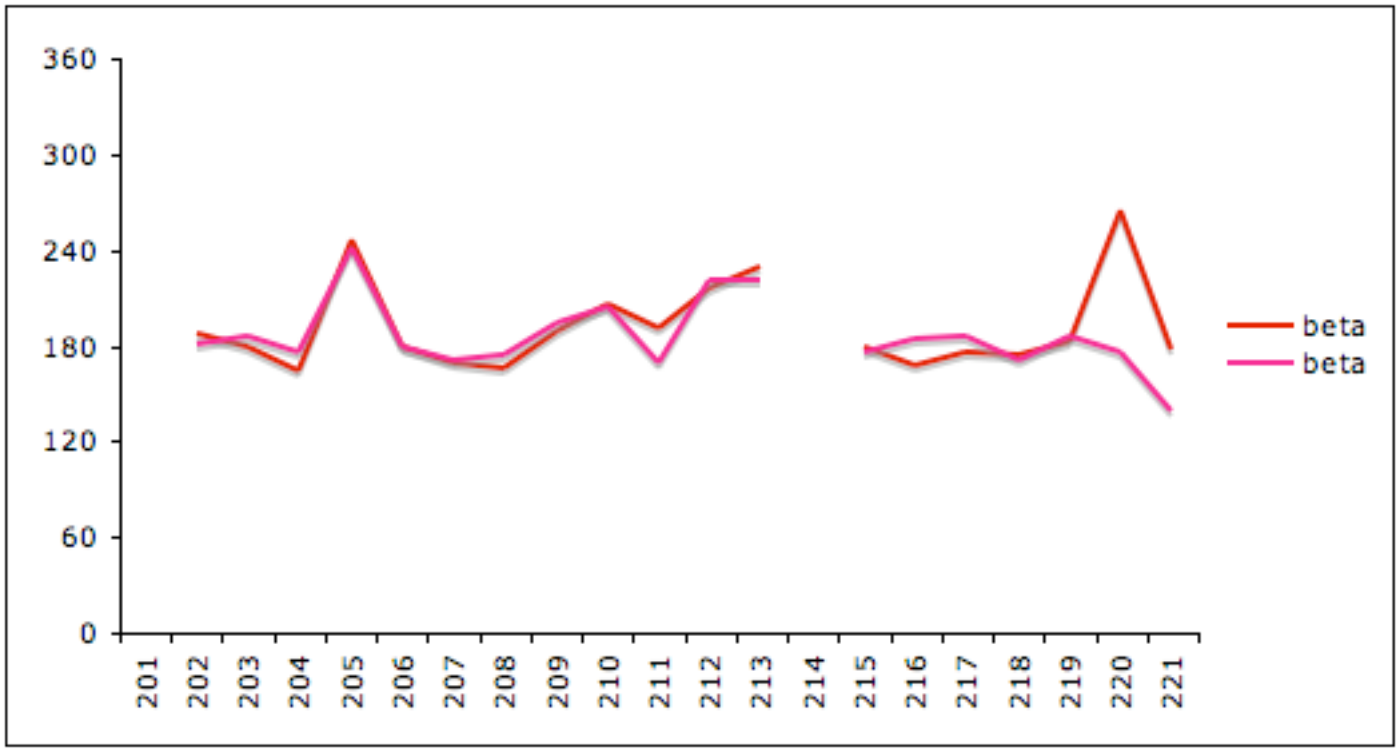
γ: O5'-C5'-C4'-C3' (42° A-DNA vs. 31° B-DNA)

The anticorrelation of the $\alpha\gamma$ torsions preserves the stacked geometry of DNA base pairs in the I-*Ppo*I-DNA complex.



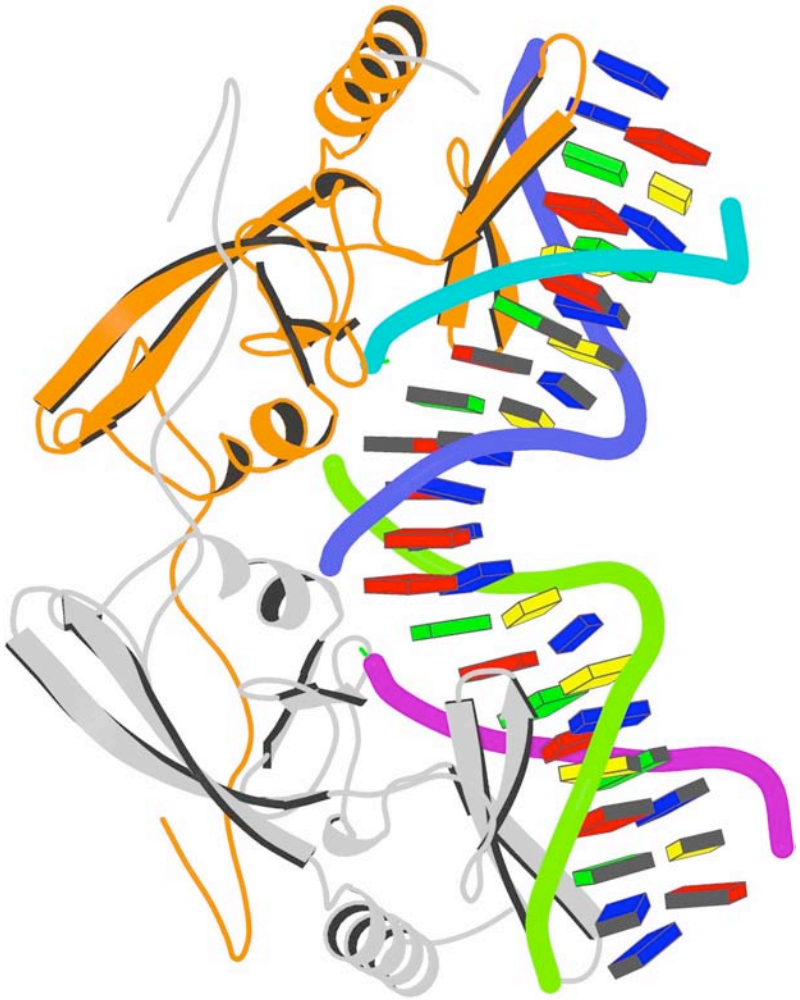
α: O3'(i-1)-P-O5'-C5'
γ: O5'-C5'-C4'-C3'

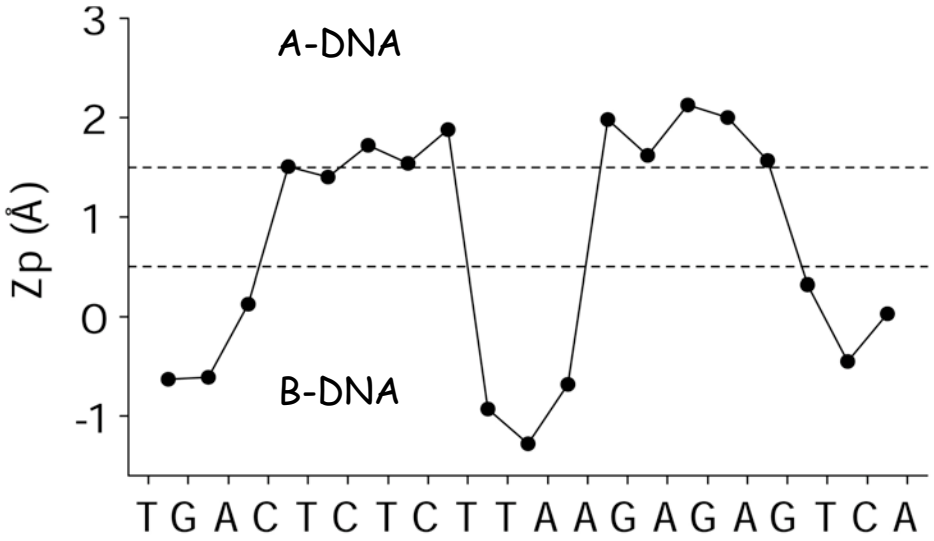
The excursions in the β torsion in the I-*Ppo*I-DNA complex differ from the changes characteristic of changes from the canonical B to A forms.



β: P-O5'-C5'-C4' (175° A-DNA vs. 136° B-DNA)

Multiple A/B junctions apparently contribute to the significant DNA bending in the I-*Ppo*I-DNA complex.





Flick *et al.* (1998). "DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-*Ppo*I." *Nature* 394, 96–101.

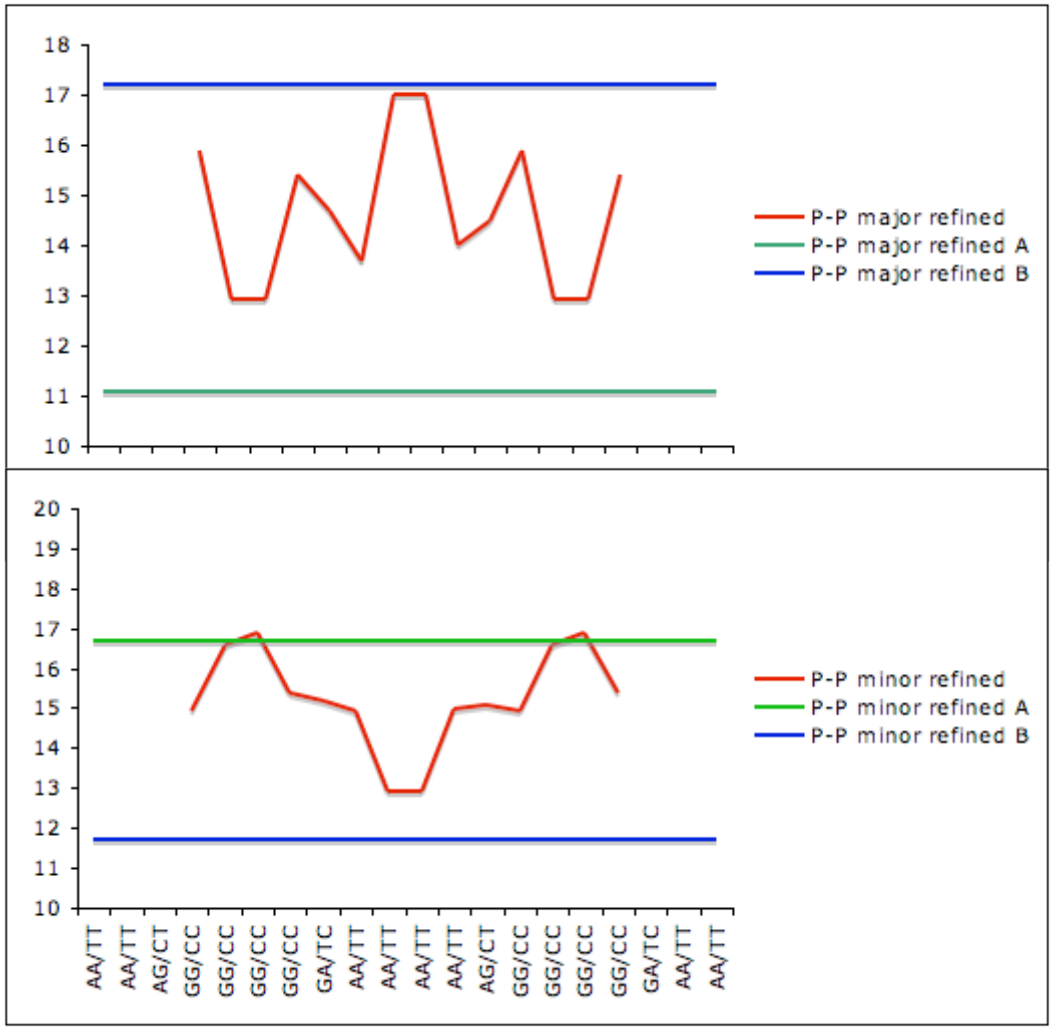
Lu *et al.* (2000)

Analysis of the I-*Ppo*I-DNA complex suggests that concatenation of A- and B-DNA helices generates a naturally curved structure.

If regularly repeated, the concatenation of A- and B-DNA helices generates a naturally curved structure.

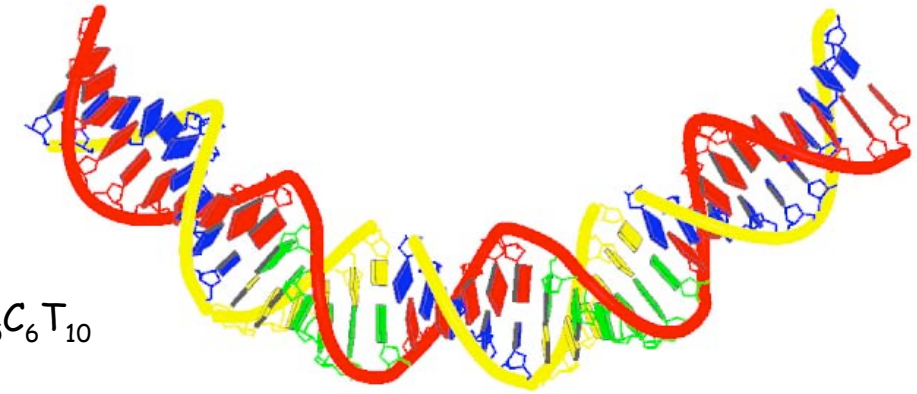
$A_3G_5A_5G_5A_3 \cdot T_3C_5T_5C_5T_3$ miniduplex
B-like AA·TT and AG·TC steps
A-like GG·CC and GA·CT steps

The concatenation of short A- and B-DNA helices alters the groove structure at helix junctions.

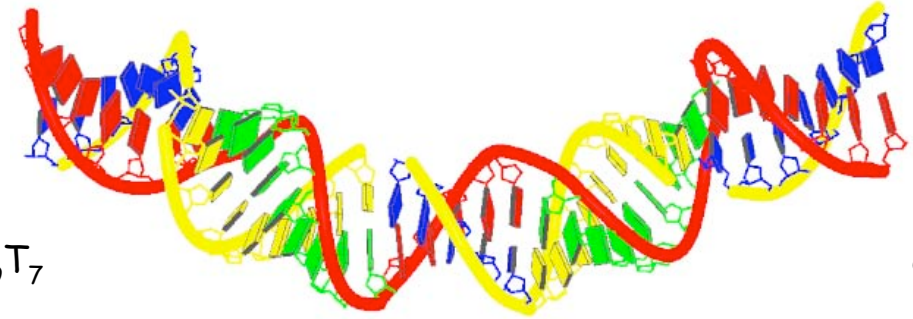


$A_3G_5A_5G_5A_3{\cdot}T_3C_5T_5C_5T_3$ miniduplex
B-like *AA·TT* and *AG·TC* steps
A-like *GG·CC* and *GA·CT* steps

The angle between the base pairs at the termini of concatenated helices depends upon the length of the A-DNA segment.



$A_{10}G_6A_5G_6A_{10}\cdot T_{10}C_6T_5C_6T_{10}$ $\cos^{-1}(\mathbf{n}_1\cdot\mathbf{n}_{37}) = 83°$



$A_7G_9A_5G_9A_7\cdot T_7C_9T_5C_9T_7$ $\cos^{-1}(\mathbf{n}_1\cdot\mathbf{n}_{37}) = 27°$

$AG_{15}A_5G_{15}A\cdot TC_{15}T_5C_{15}T$ $\cos^{-1}(\mathbf{n}_1\cdot\mathbf{n}_{37}) = 67°$

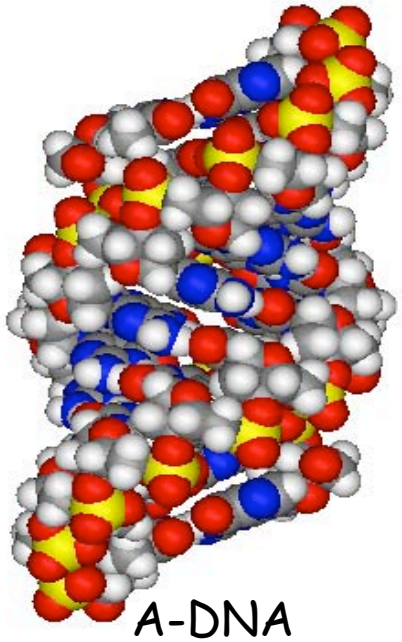
DNA phase transitions

DNA phase transitions

- The ionic character of the sugar-phosphate backbone makes DNA especially sensitive to changes in its local environment, *e.g.*, salt, alcohol.
- Interactions with other molecules, including proteins, may lead to a change of helical state.

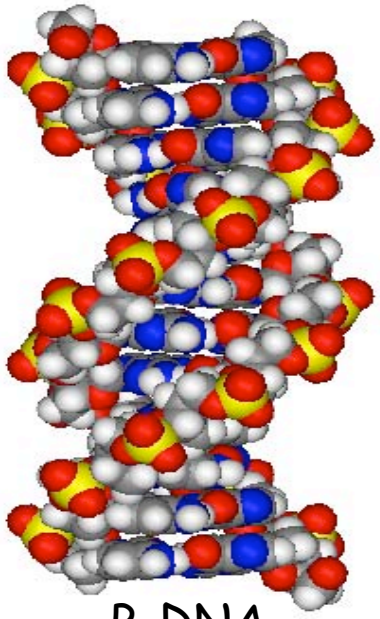
The A→B transition: first known change of DNA double-helical state.

11 bp/turn



A-DNA

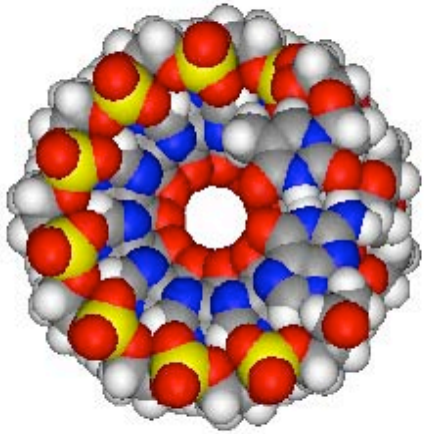
+H_2O
→
←
+salt and/or
+alcohol



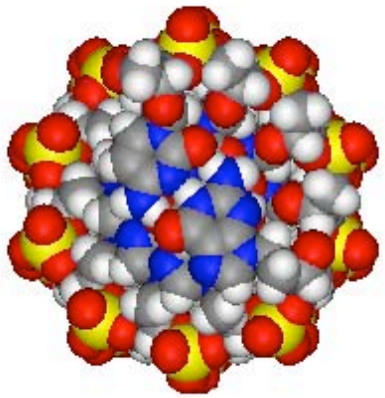
B-DNA

10 bp/turn

A-DNA base pairs inclined with respect to helical axis and untwisted *cf.* B DNA.



A-DNA minor groove wider and more shallow, major groove narrower and more deep *cf.* B DNA



Base pairs displaced from A-DNA helical axis.

Base composition of A- & B-DNA structures depends on sequence.

Dimer Step	A-DNA (2.0 Å)	All	B-DNA (2.0 Å)	All	$\Delta G_{B/A}$† (kcal/mole)
AA·TT	0	2	**31**	**112**	0.97
GG·CC	**92**	**243**	14	55	0.19
CA·TG	16	20	**17**	**48**	1.04
AC·GT	**42**	**80**	4	22	0.13
GA·TC	4	21	**27**	**80**	0.96
AG·CT	2	18	18	30	0.33
GC·GC	40	102	11	121	0.73
CG·CG	44	93	28	200	0.52
TA·TA	10	26	6	15	0.75
AT·AT	4	6	17	58	0.68

† Ivanov & Minchenkova (1995) "The A-form of DNA: in search of biological role," *Mol. Biol.* 28, 780-788.

A/B helical motifs are common in complexes of DNA with enzymes that make or break the O3´-P phosphodiester linkage

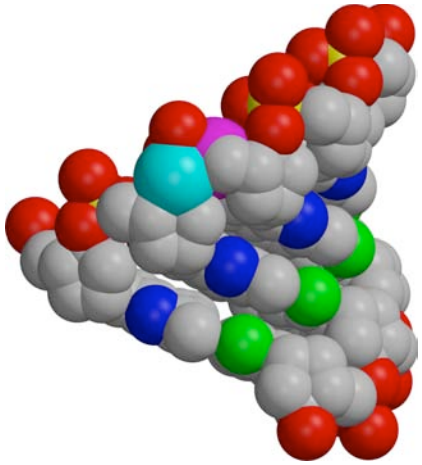
Tc3 transposase	A G G G G G G G T C C T A T A G A A C T T T C C C C C C C A G G A T A T C T T G A
I-PPOI homing endonuclease	T T G A C T C T C T T A A G A G A G T C A A C T G A G A G A A T T C T C T C A G T T
PVUII restriction endonuclease	T G A C C A G C T G G T C C T G G T C G A C C A G
Eco RV endonuclease	G G G A T A T C C C C C C T A T A G G G
TAQ polymerase	G A C C A C G G C G C C C T G G T G C C G C C C
Bacillus polymerase I	G C A T G A T G C C G T A C T A C G A
HIV-1 RT + FAB 28	G T C C C T G T T C G G G C G C C A C A G G G A C A A G C C C G C G G T A

Lu *et al.* (2000)

DNA helical form influences atomic exposure as well as global shape.

A-DNA

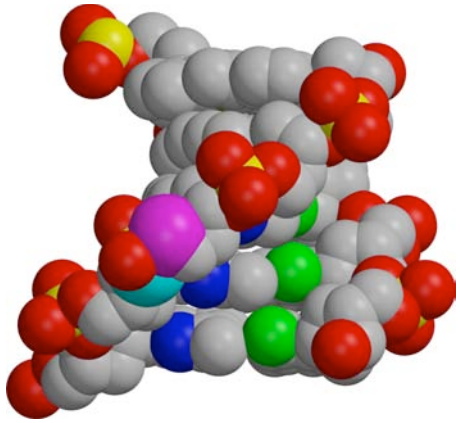
11 res/turn
Roll > 0
Slide < 0



Wide/shallow minor groove exposes O3´ and base-pair edges

B-DNA

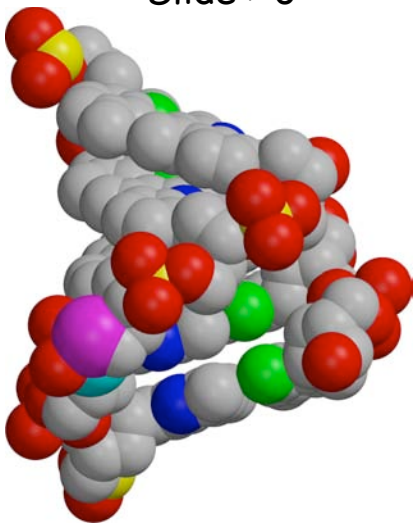
10 res/turn
Roll ≈ 0
Slide ≈ 0



~12 Å minor groove exposes O5´ vs. O3´, partial base-pair edges

C-DNA

9 res/turn
Roll < 0
Slide > 0

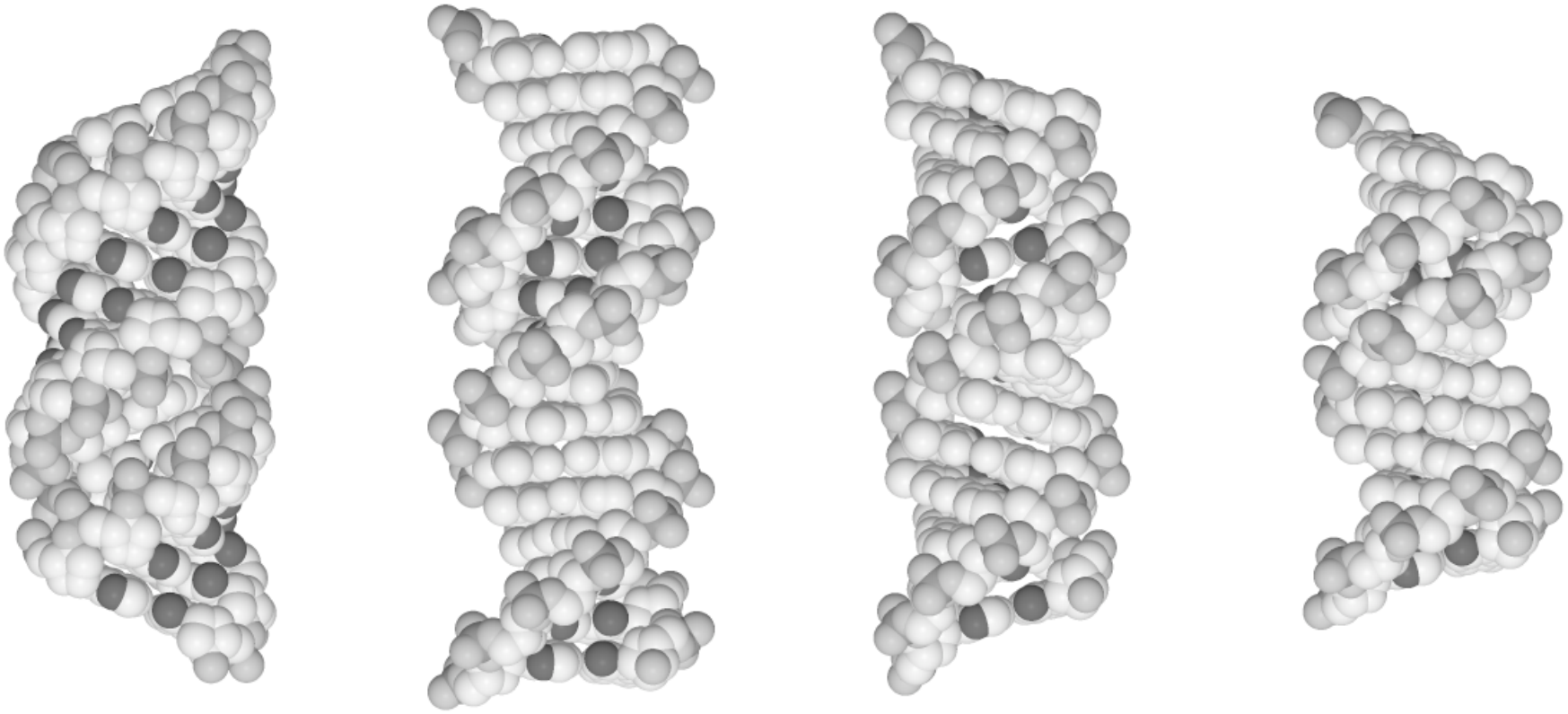


Deep/narrow minor groove exposes O5´, hides base-pair edges

Pseudo-symmetric R(N3), Y(O2) proton-acceptor atoms of Watson-Crick base pairs

Transformations within the ABCD family of right-handed double helices affect:

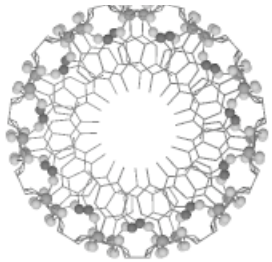
(i) the inclination of Watson-Crick base pairs
(ii) the widths and exposure of atoms on the major and minor-groove edges
(iii) the overall helical extension.



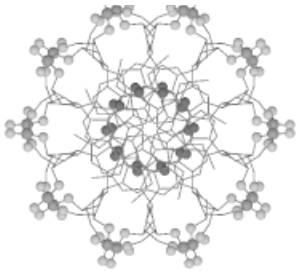
A-DNA B-DNA C-DNA D-DNA

Transformations within the ABCD family of structures also alter:

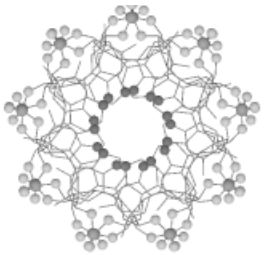
(i) the number of residues per helical turn;
(ii) the width of the solvent "channel" through the center of the duplex.



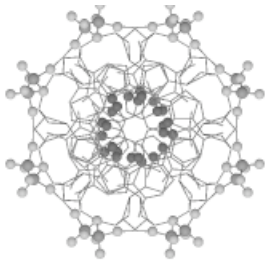
A-DNA



B-DNA



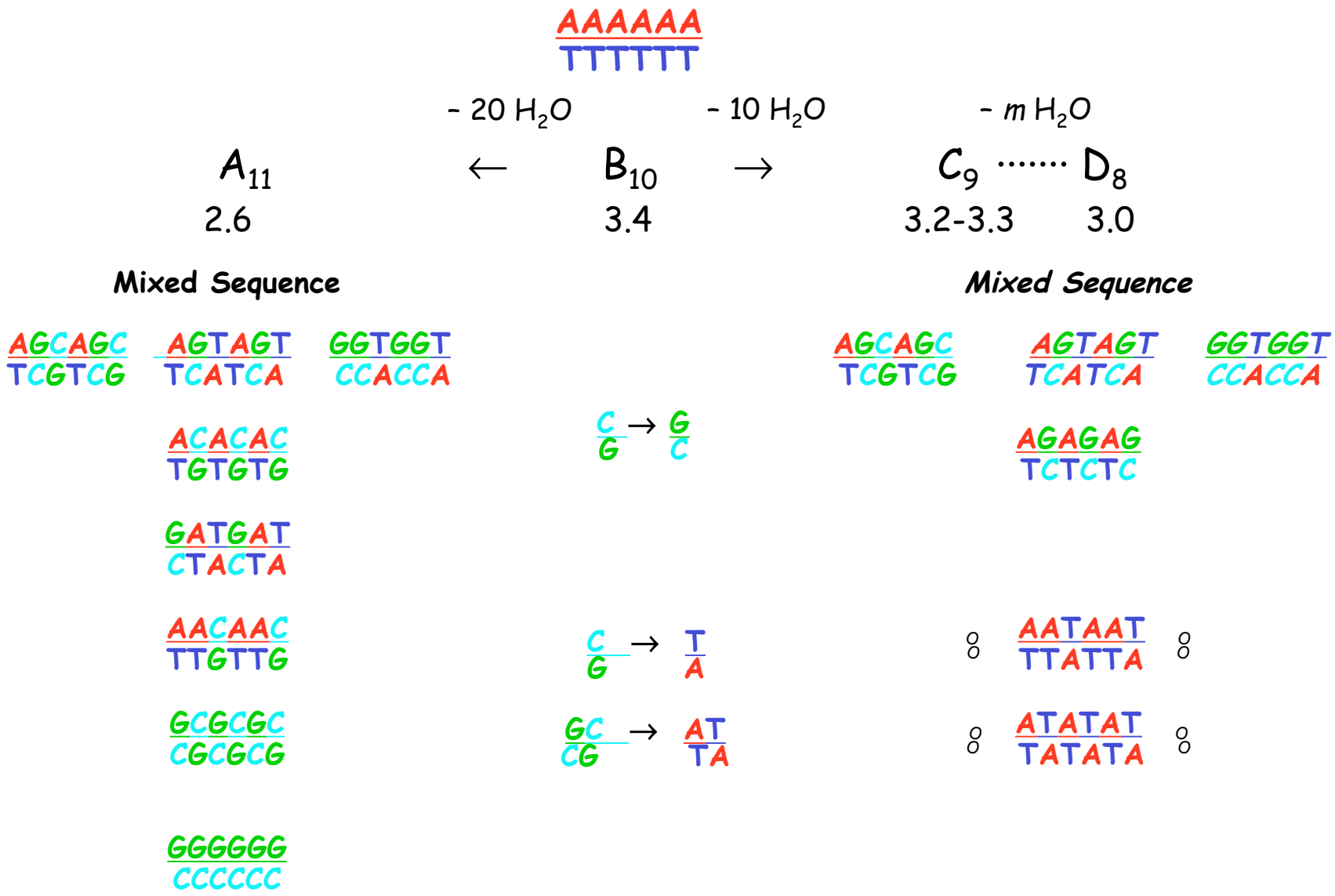
C-DNA



D-DNA

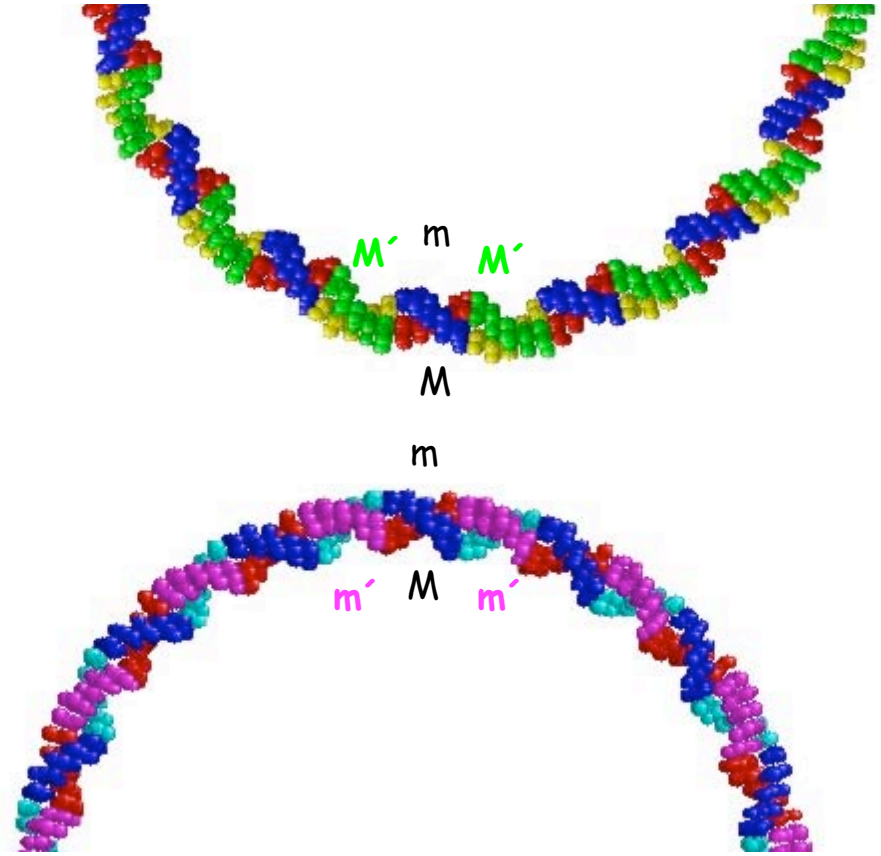
The tendency to adopt these helical forms depends upon sequence: poly dG· poly dC is "A philic"; repetition of A·T or I·C bases promotes formation of the C and D forms .

Sequence-dependent responses of DNA helical structure



Composite data of Leslie *et al.* (1980) *J. Mol. Biol.* **143**, 49-72; Harmouchi *et al.* (1990) *Eur. Biophys. J.* **19**, 87-92.

Deformations toward the A and C forms bend DNA in the opposite sense.



A-DNA

B-DNA

Global bend:
360°/150 bp
circle

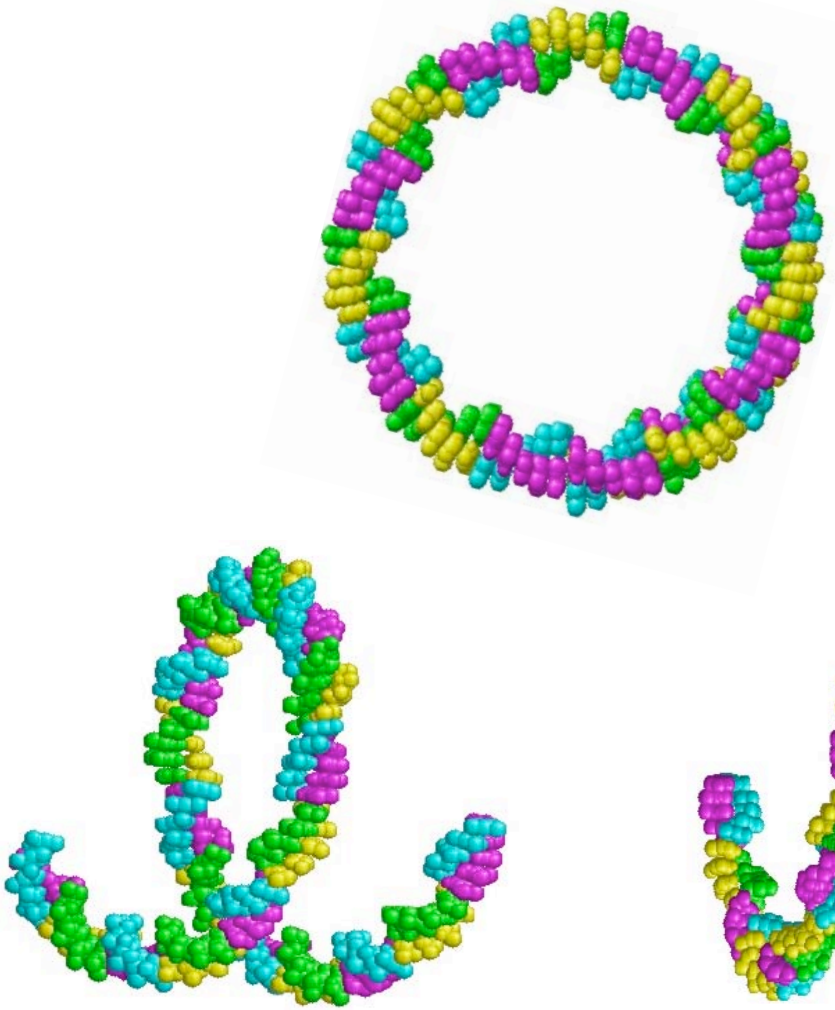
B-DNA

C-DNA

Major (M), minor (m) groove edges lie on opposite faces of B→A vs. B→C induced curves.

A.R. Srinivasan

Combined B→A and B→C deformations tighten the bending of DNA:



Global bend:
360°/75 bp
left-handed superhelix

Combined B→A and B→C deformations tighten the bending of DNA:

Unusual DNA structures

DNA sequences of repeated CG dinucleotides crystallize in an unusual left-handed Z-DNA (zig-zag) double-helical form.

The conformational parameters of Z-DNA differ at YR vs. RY steps.

Z-DNA backbone torsion angles

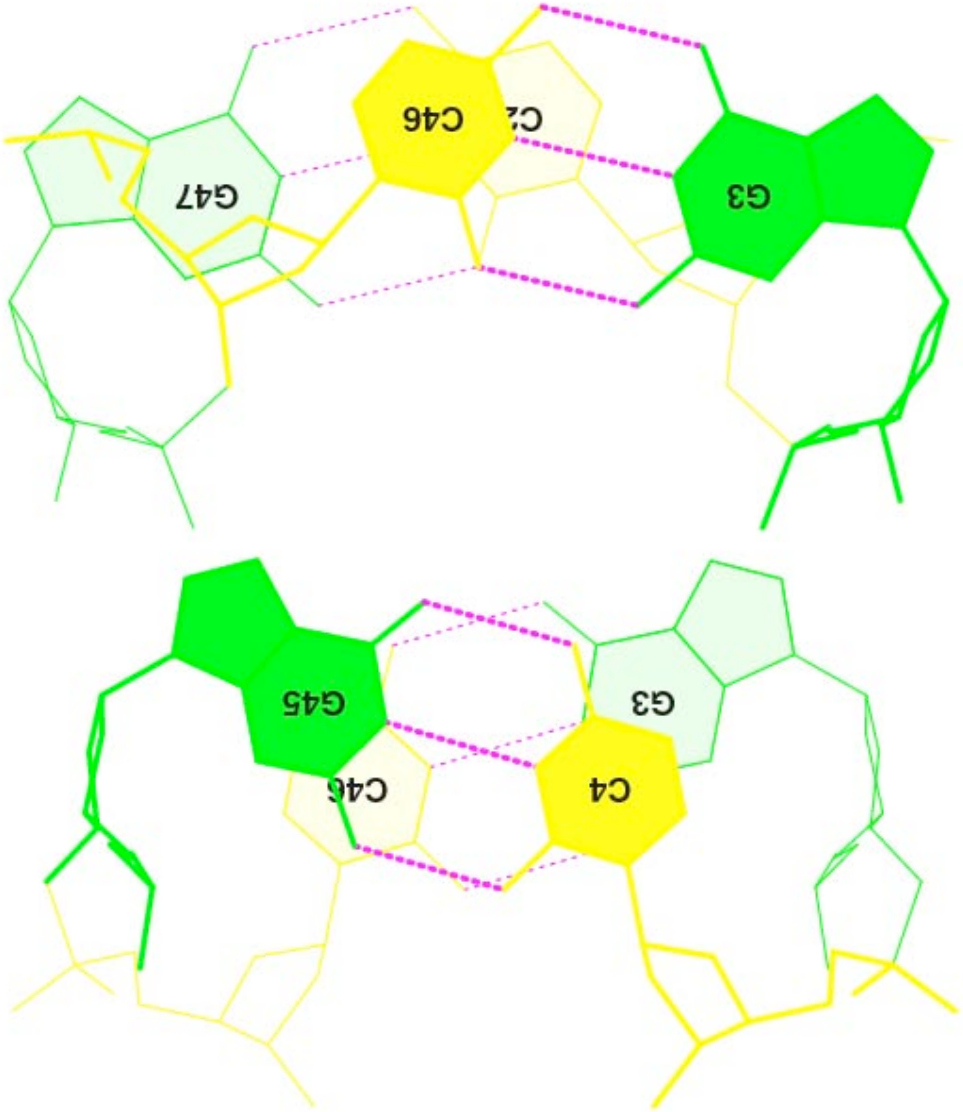
	α	β	γ	δ	ε	ζ	χ
C	-140	-137	51	138	-97	82	-154
G	52	179	-174	95	-104	-65	59

α: O3'(i-1)-P-O5'-C5'
β: P-O5'-C5'-C4'
γ: O5'-C5'-C4'-C3'
δ: C5'-C4'-C3'-O3'
ε: C4'-C3'-O3'-P(i+1)
ζ: C3'-O3'-P(i+1)-O5'(i+1)

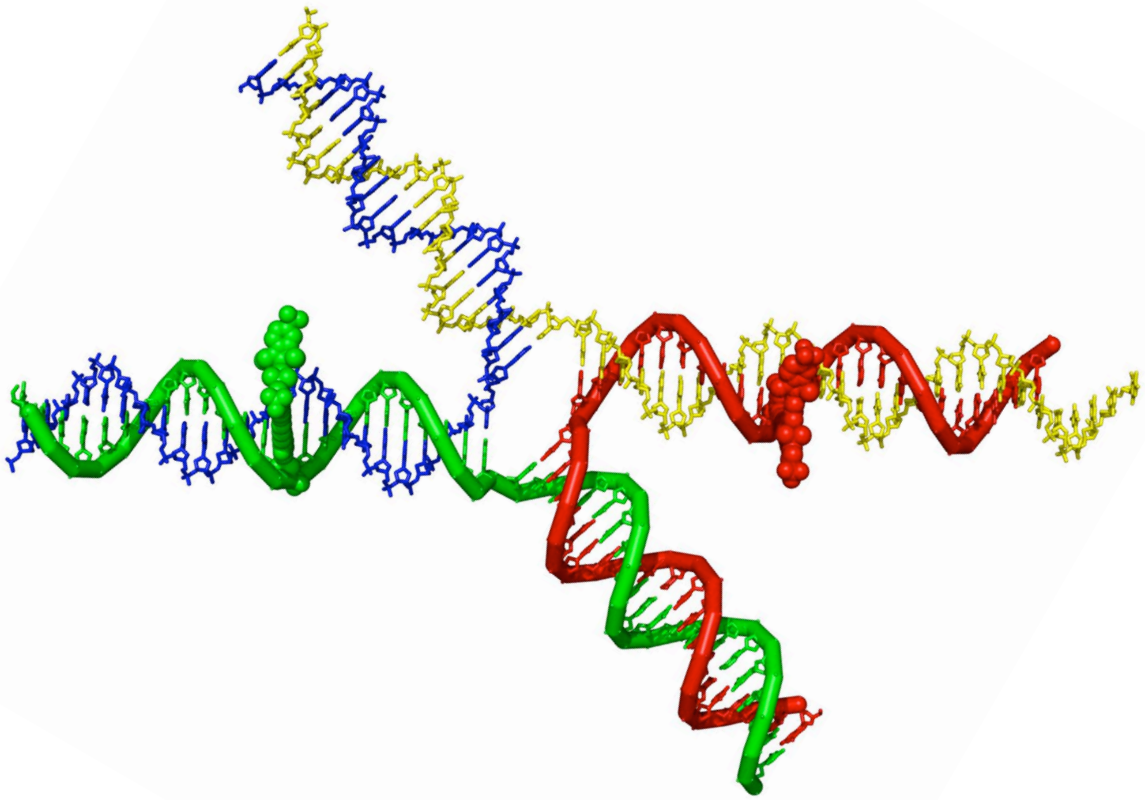
χ pyrimidines(Y): O4'-C1'-N1-C2
χ purines (R): O4'-C1'-N9-C4

Acyclic torsions of dimer steps noted by color coding: CpG GpC

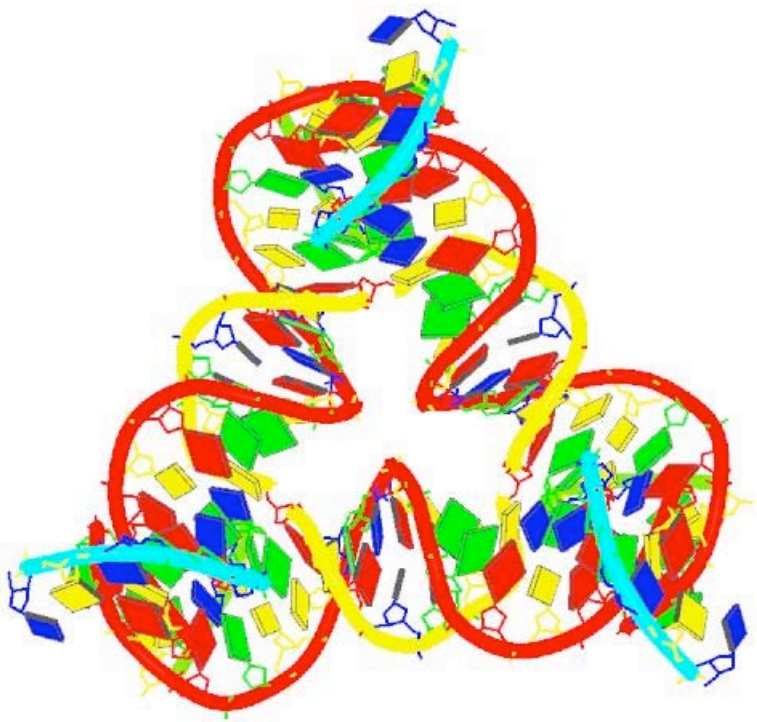
Z-DNA base-pair steps progress in an opposite direction from those of the ABCD family.

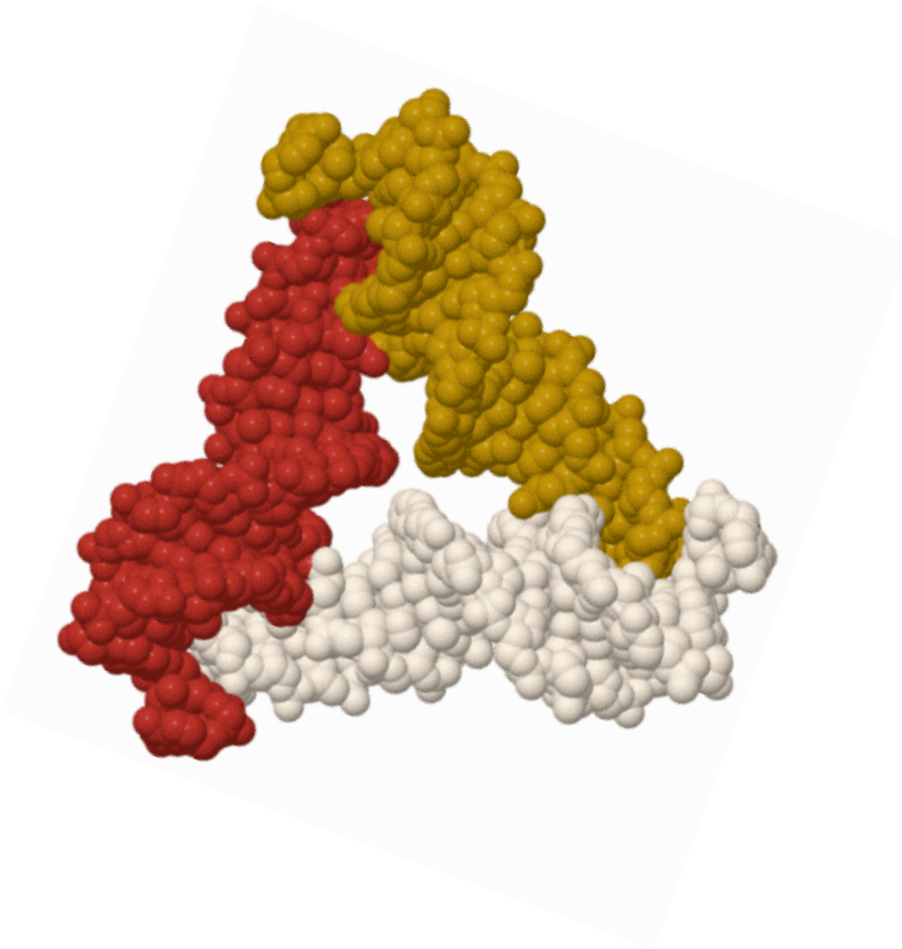


Some DNA sequences can be locked in 4-way Holliday junctions.



DNA junctions are the design elements of novel nanomaterials.

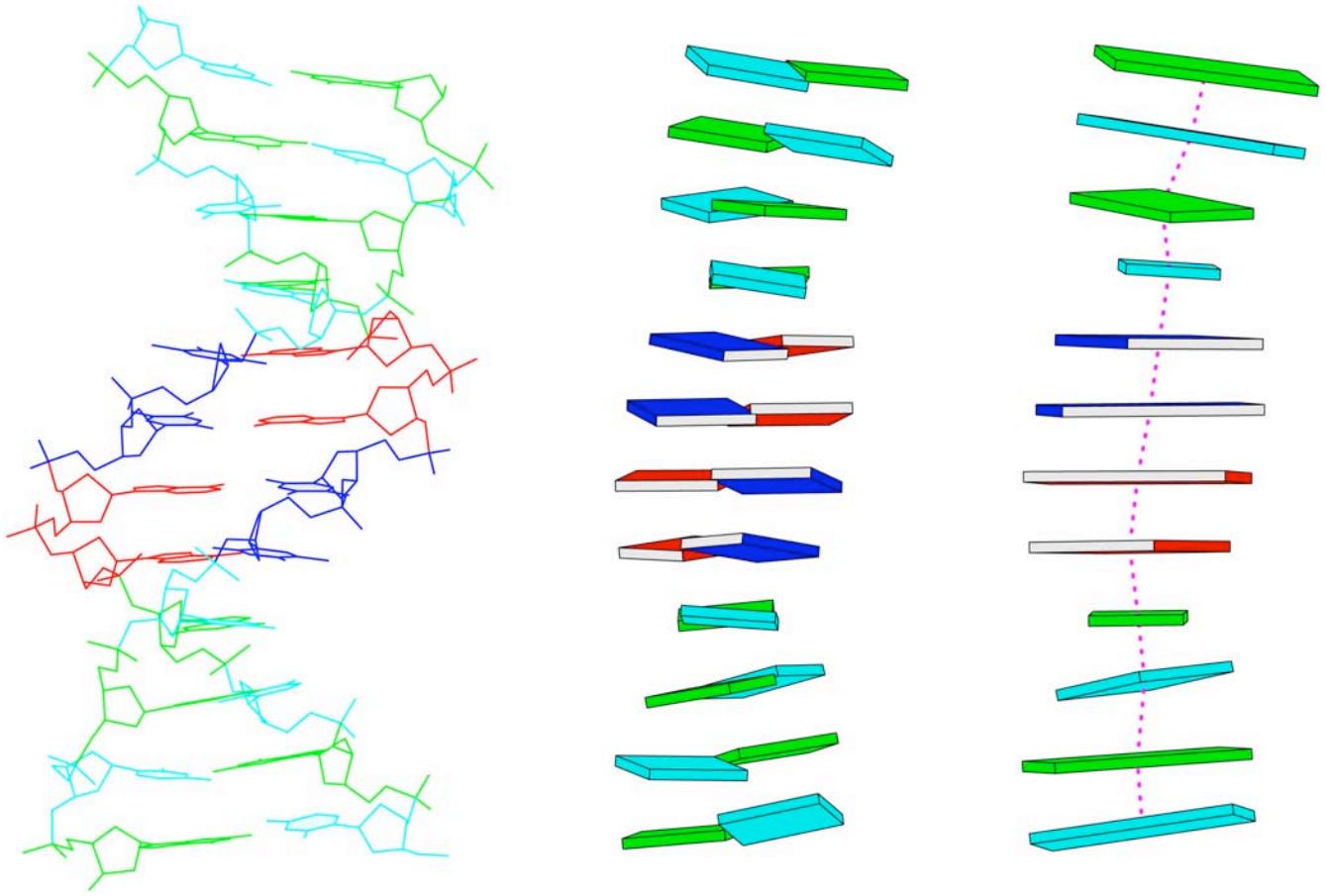




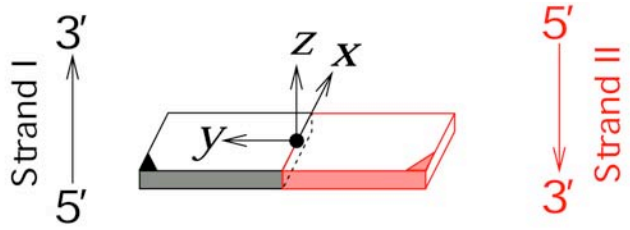
J. Zheng *et al*. (2009) From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461**, 74-77.

DNA as a collection of rigid-body parameters

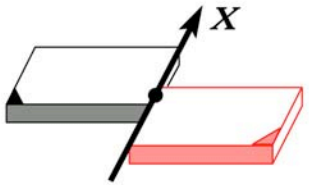
DNA sequence-dependent structure is easily understood at the base-pair level.
(bd1084; Shui et al., 1998)



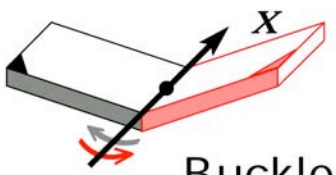
Complementary base-pair frame and parameters



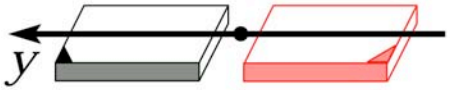
Base-Pair Reference Frame



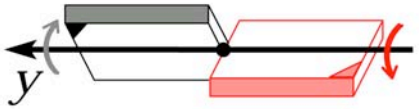
Shear



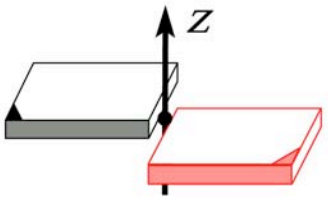
Buckle



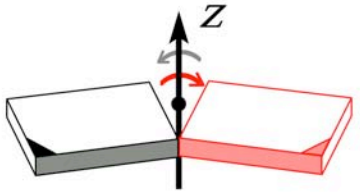
Stretch



Propeller

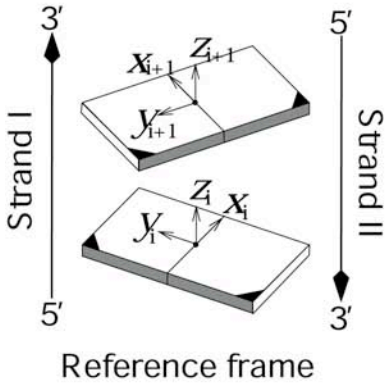


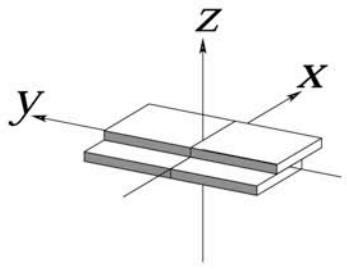
Stagger



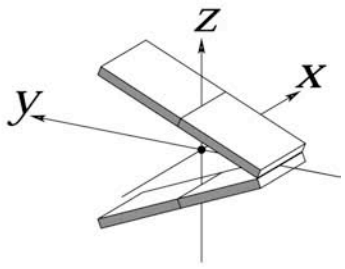
Opening

Nucleic acid base-pair "step" parameters

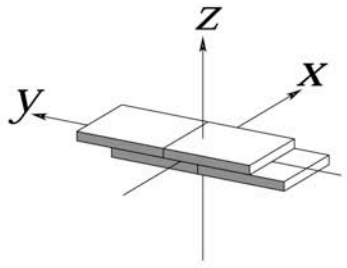




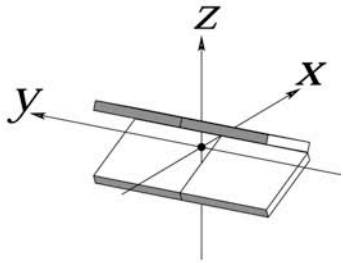
Shift



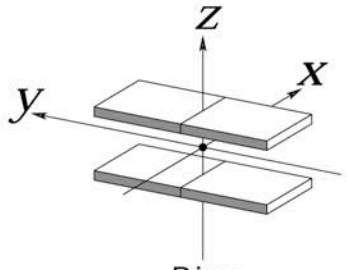
Tilt



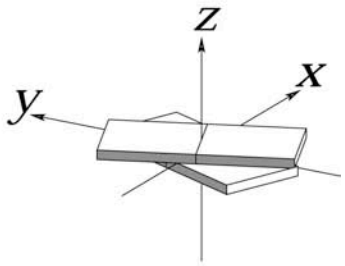
Slide



Roll

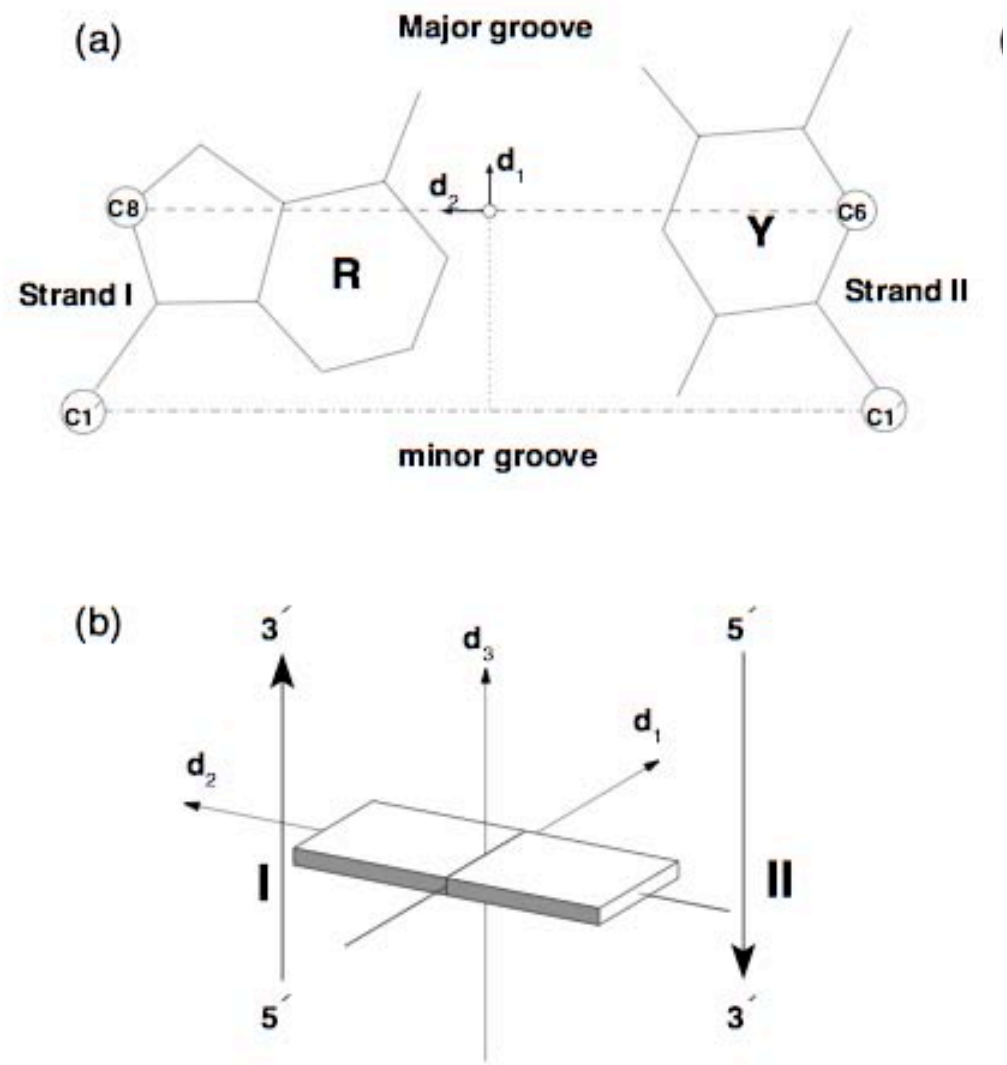


Rise

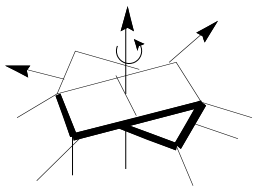


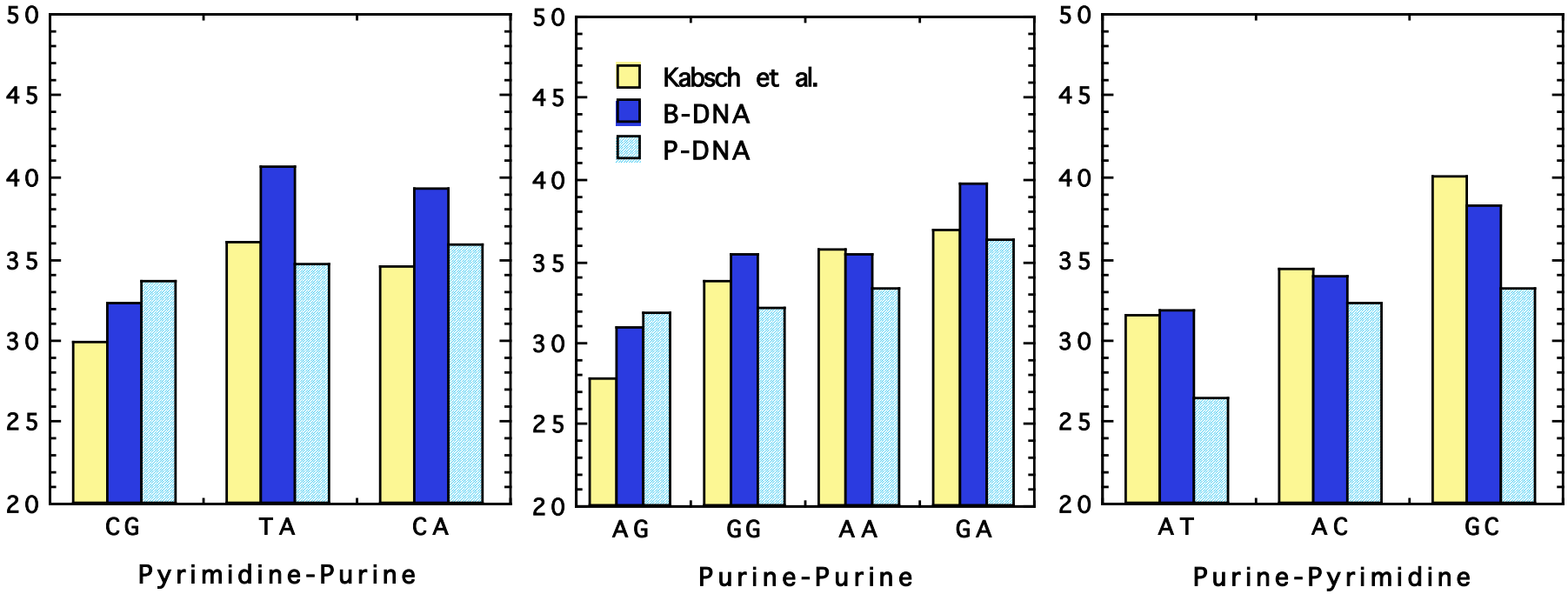
Twist

Standard base-pair coordinate frame



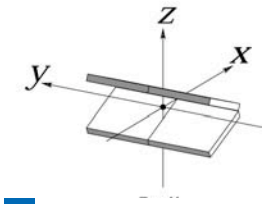
Comparative DNA Twist Angles
Crystal vs. Solution Averages

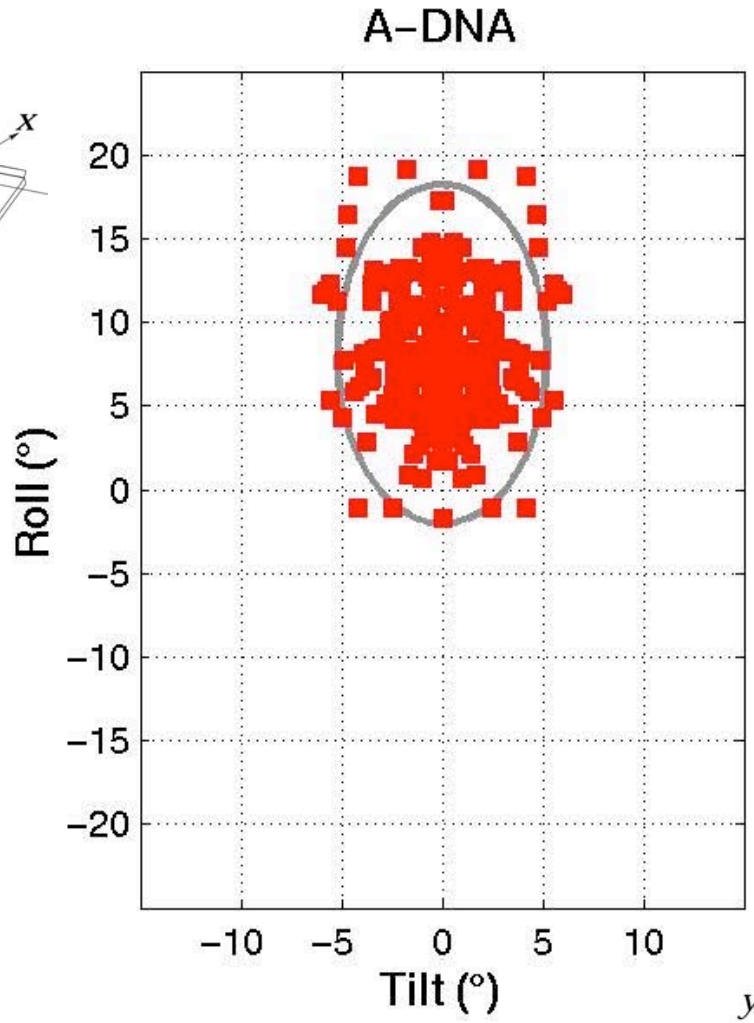


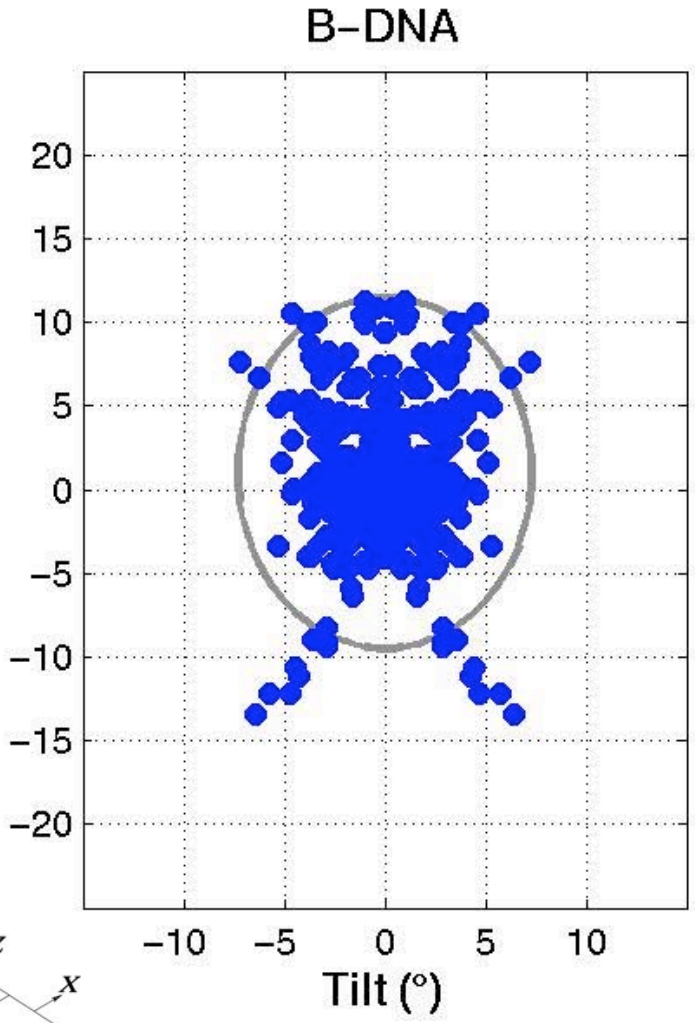


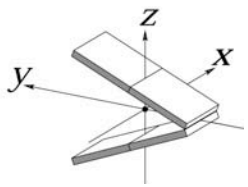
Gorin et al. (1995)

Bending angles of base-pair "steps" in DNA crystal structures

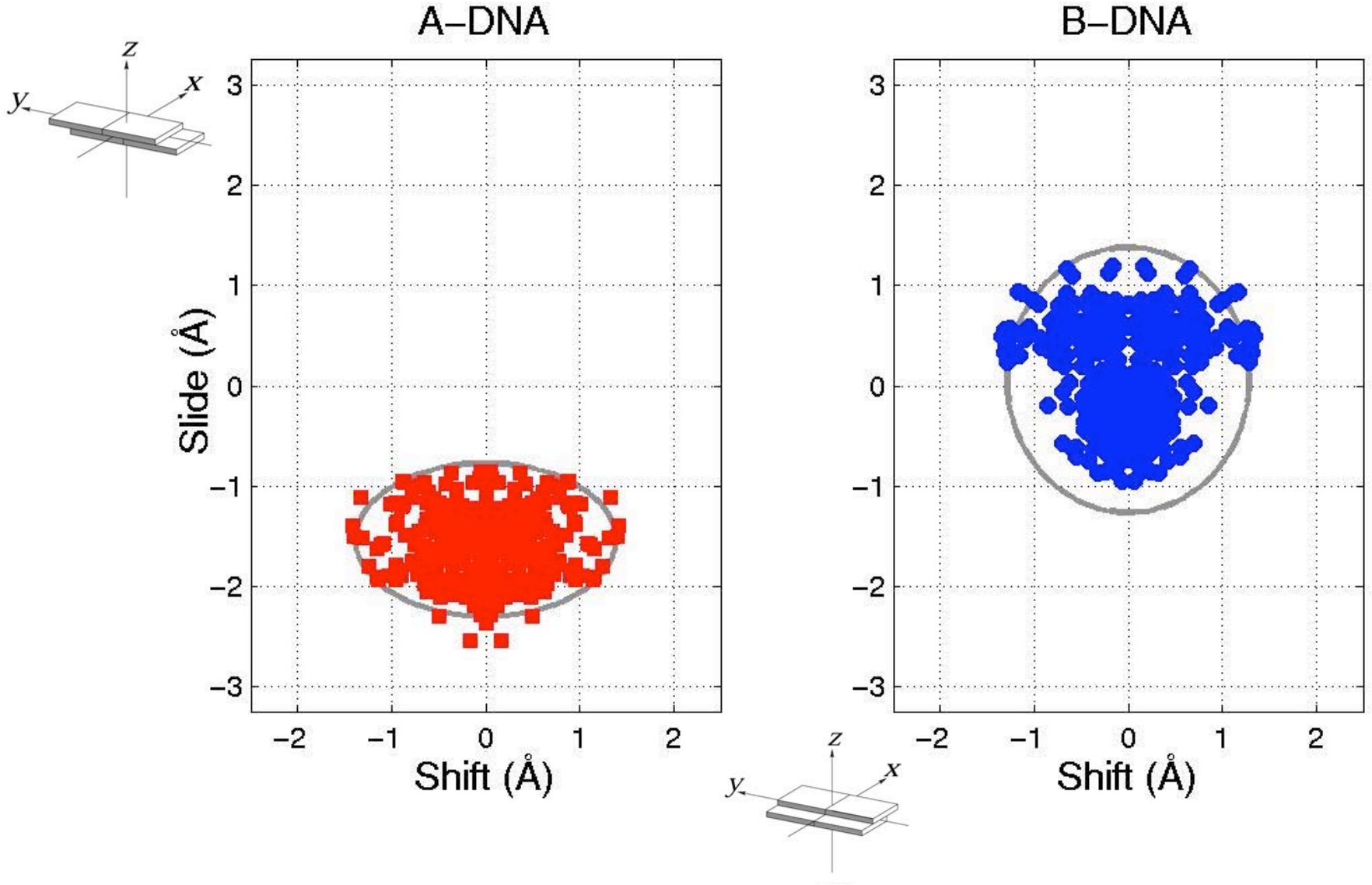




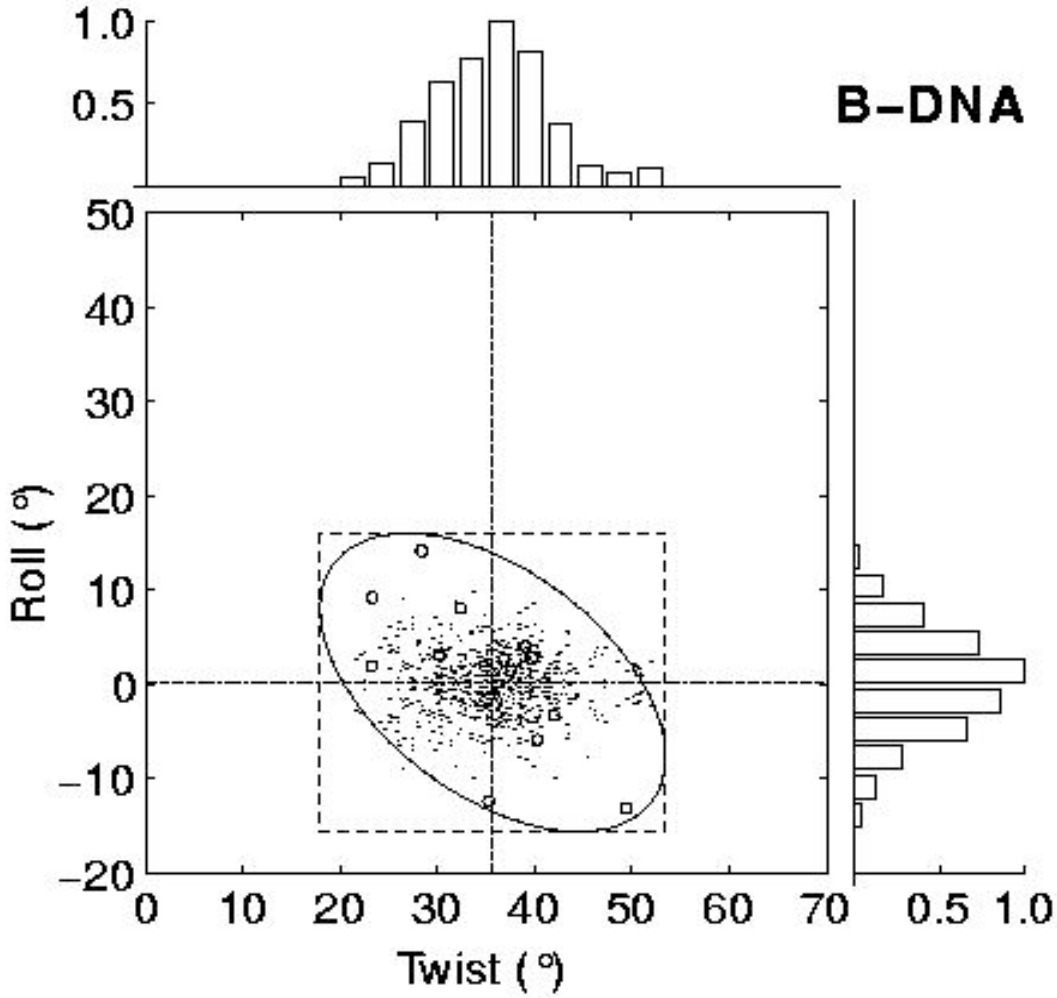




Shear displacement of dimers in DNA crystal structures



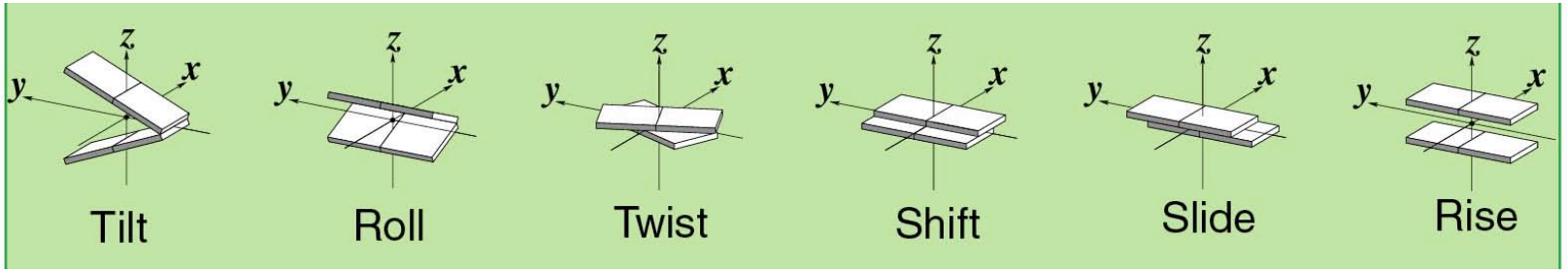
Intrinsic coupling of Roll and Twist angles in DNA structures



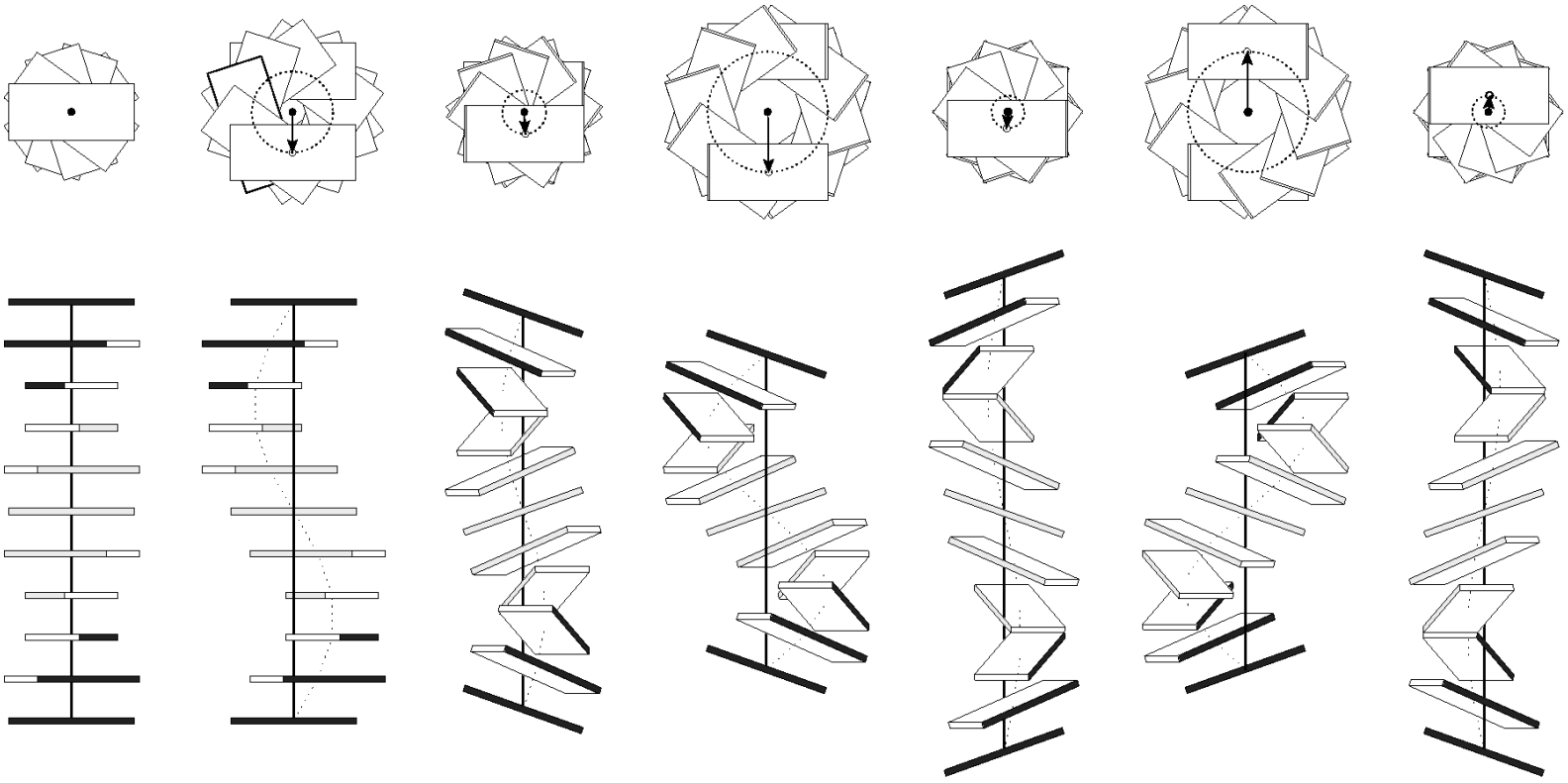
Olson *et al.* (1998)

The canonical A- and B-DNA structures exhibit differences in three of the six base-pair step parameters.

	Tilt	Roll	Twist	Shift	Slide	Rise
A DNA	0	12	30	0	-1.4	3.3
B DNA	0	2	36	0	0	3.4

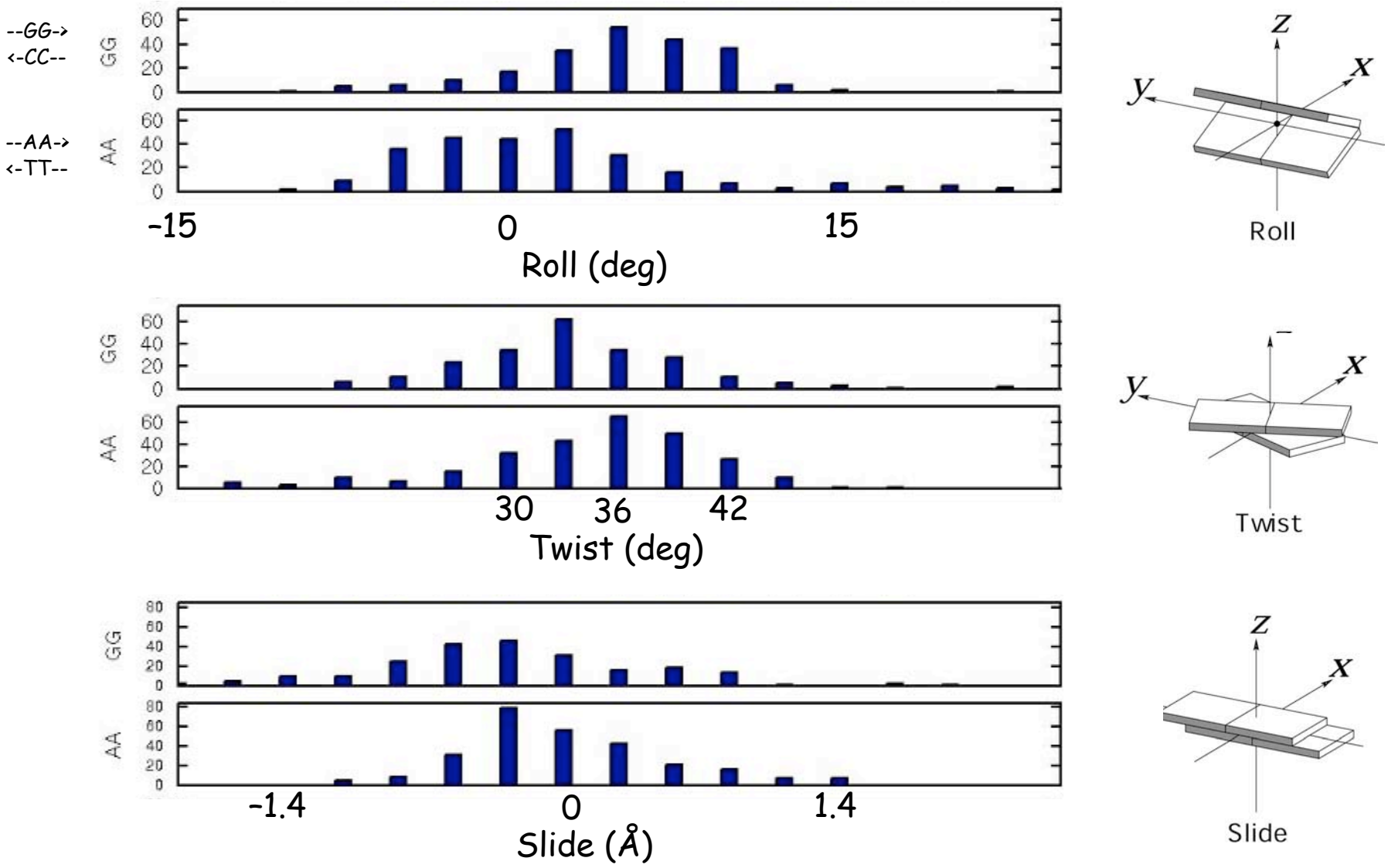


The differences in Roll, Twist, and Slide in A and B DNA account for the observed differences in global helical structure.



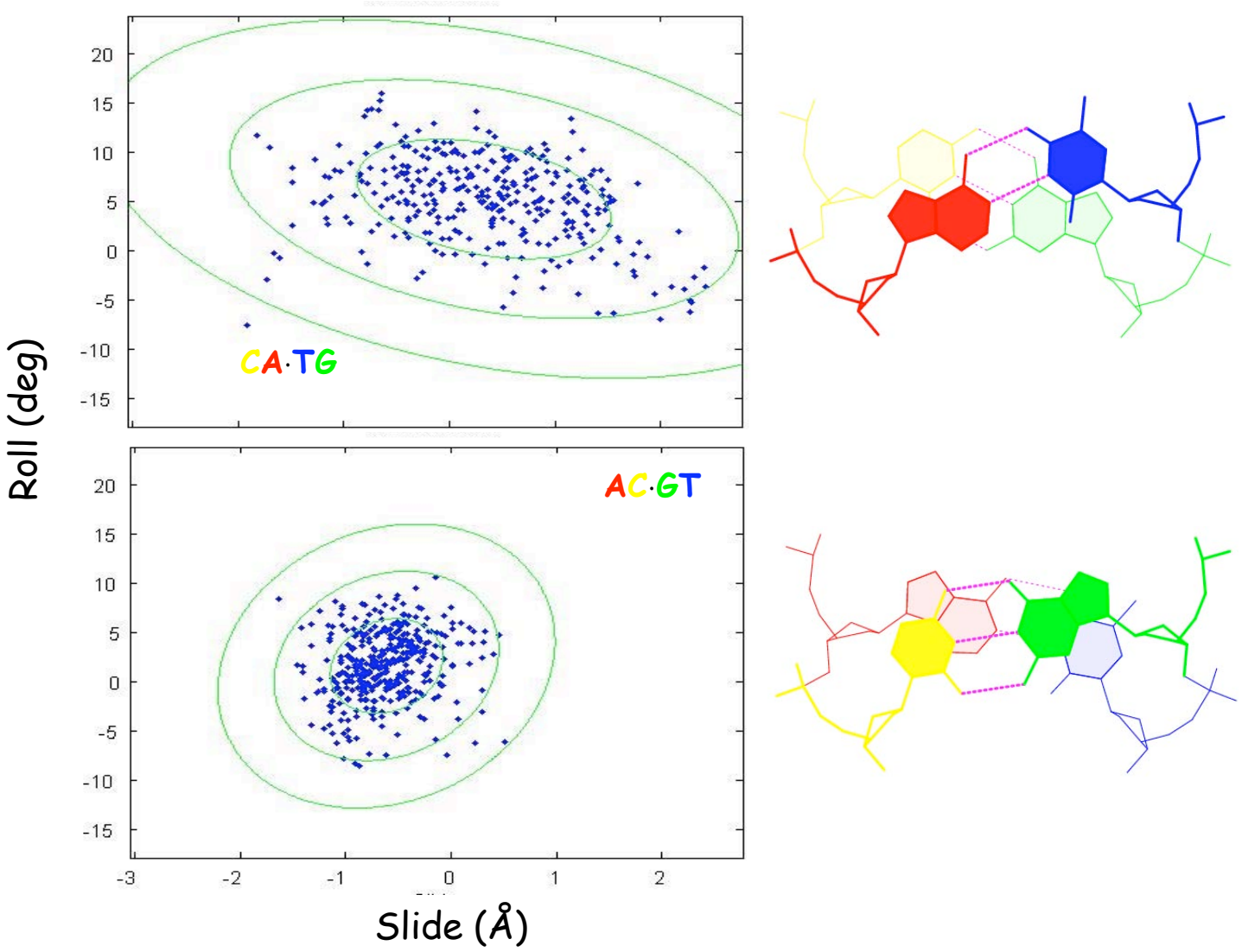
(a) B-DNA (b) −Slide (c) +Roll (d) −Slide, +Roll (A-DNA) (e) −Slide, −Roll (f) +Slide, −Roll (g) +Slide, +Roll

Roll, Slide, Twist exhibit subtle, sequence-dependent behavior.



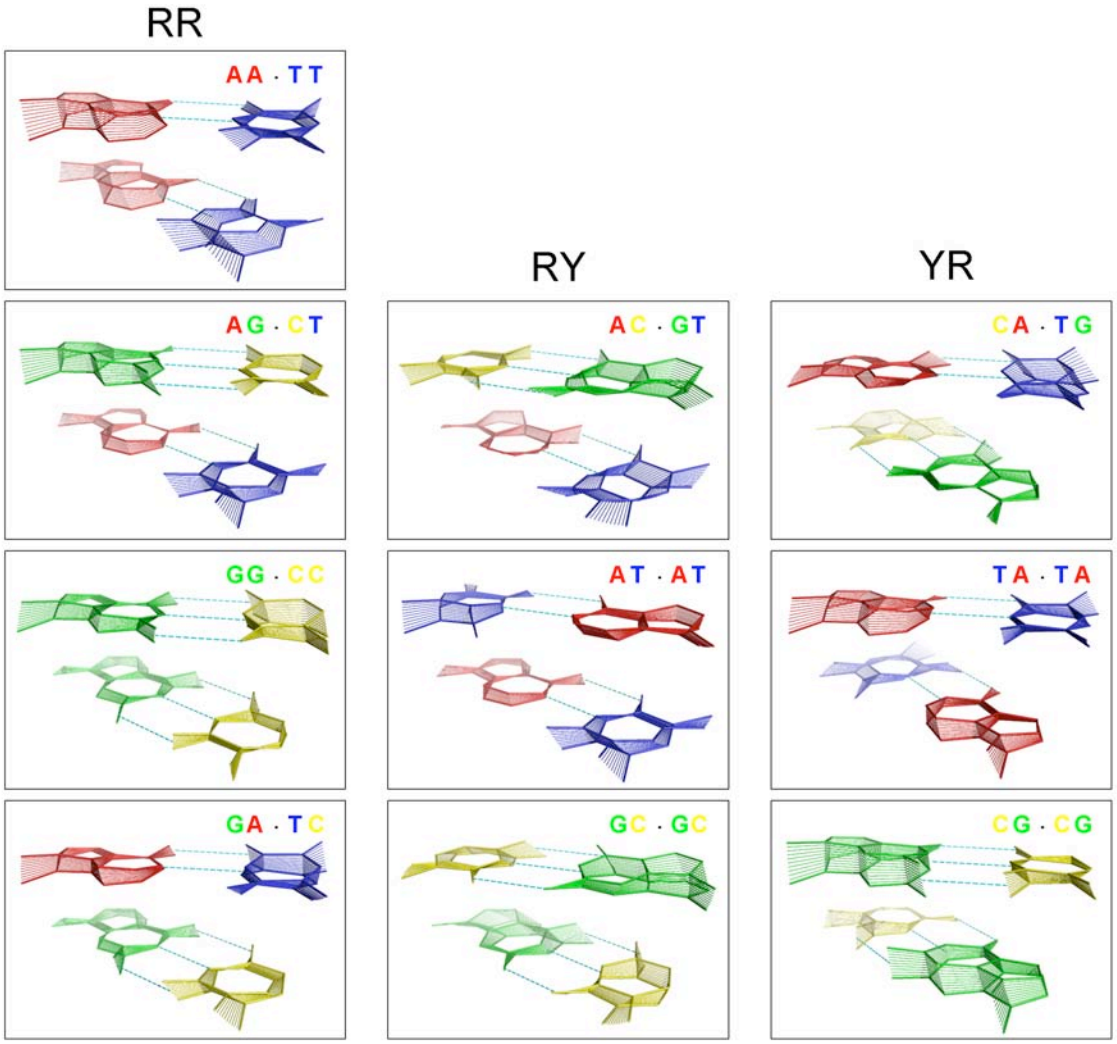
Sequence-dependent variation of the three base-pair 'step' parameters, which dominate the conformational variability in high-resolution protein-DNA structures.

CA·TG steps are naturally 'soft', allowing them to take up the deformations of structure responsible for the superhelical DNA path in the nucleosome.



341 CA·TG and 418 AC·GT steps from 239 protein-DNA crystal complexes of 2.5 Å or better resolution

Schematic of the structural and deformational code embedded in DNA sequence.



Assignment (due Tuesday, October 13, 2009):

1. Compare the overlap of bases at YR and RY base-pair steps in A, B, vs. Z DNA helices.

2. How much global bending is induced by the insertion of A DNA helical fragments of 1, 5, 6, and 11 base-pair steps within a B-DNA helix?