The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

2038-26

Conference: From DNA-Inspired Physics to Physics-Inspired Biology

1 - 5 June 2009

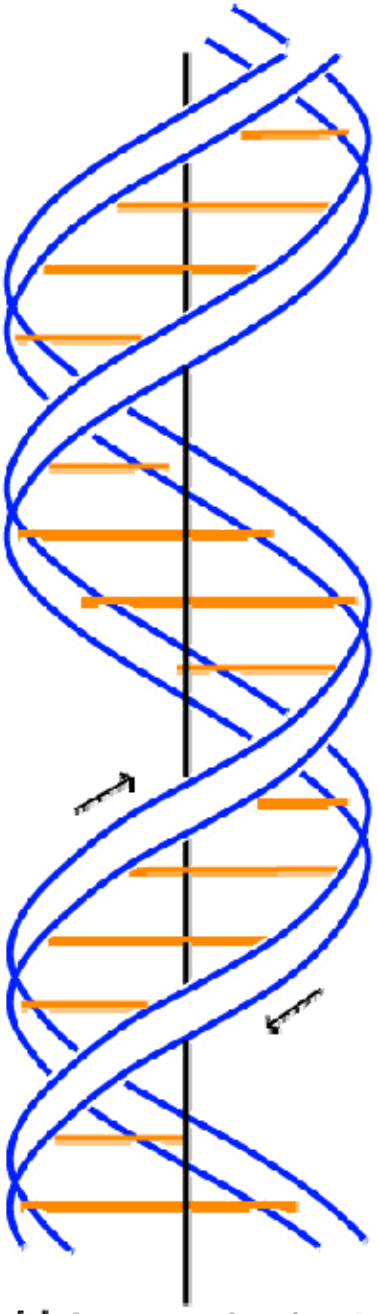
Recognition of DNA by Proteins: Direct and Indirect Readout

Loren WILLIAMS
School of Chemistry & Biochemistry
Georgia Institute of Technology
Atlanta GA 30332-0400
USA

Recognition of DNA by Proteins: Direct and Indirect Readout

[preceded by a regression into DNA (& RNA*) electrostatics]

Loren Williams
School of Chemistry and Biochemistry
Georgia Tech



We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). Watson & Crick, 1953.

Phase 1: Fiber Diffraction

27. Lacy, E., Hardison, R. C., Quon, D. & Maniatis, T. *Cell* **18**, 1273–1283 (1979).
28. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).

Crystal structure analysis of a complete turn of B-DNA

Richard Wing*, Horace Drew, Tsunehiro Takano, Chris Broka, Shoji Tanaka, Keiichi Itakura† & Richard E. Dickerson

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

DNA is probably the most discussed and least observed of all biological macromolecules. Although its role in biology is a central one, with many examples such as operators and restriction sites where specific base sequences have control functions or interact with specific enzymes, the structures that DNA can adopt have been based until now only on sequence-averaged fibre diffraction patterns. Recent improvements in triester synthesis methods have made possible the preparation of sufficient homogeneous DNA of predetermined sequence for crystallization and X-ray structure analysis. We report here the first single-crystal structure analysis of more than a complete turn of right-handed B-DNA, with the self-complementary dodecamer sequence d(CpGpCpGpApApTpTpCpGpCpG) or CGCGAATTCGCG.

Two of the first crystal structure analyses of DNA prepared by the triester method, CGCGCG (ref. 1) and high-salt CGCG (ref. 2), produced variants of a totally unexpected left-handed zigzag or Z helix, of considerable structural interest but uncertain biological significance. The dodecamer sequence reported here was originally chosen to incorporate an *Eco*RI restriction endonuclease site, GAATTC, within ends related to the CGCG tetramer that was already being investigated in our laboratory. The discovery of the Z helix made this sequence potentially even more interesting as an example of two Z-compatible segments,

* Present address: Department of Chemistry, University of California at Riverside, Riverside, California 92521.
† City of Hope National Medical Center, Duarte, California 91010.

Phase 2: Single Crystal

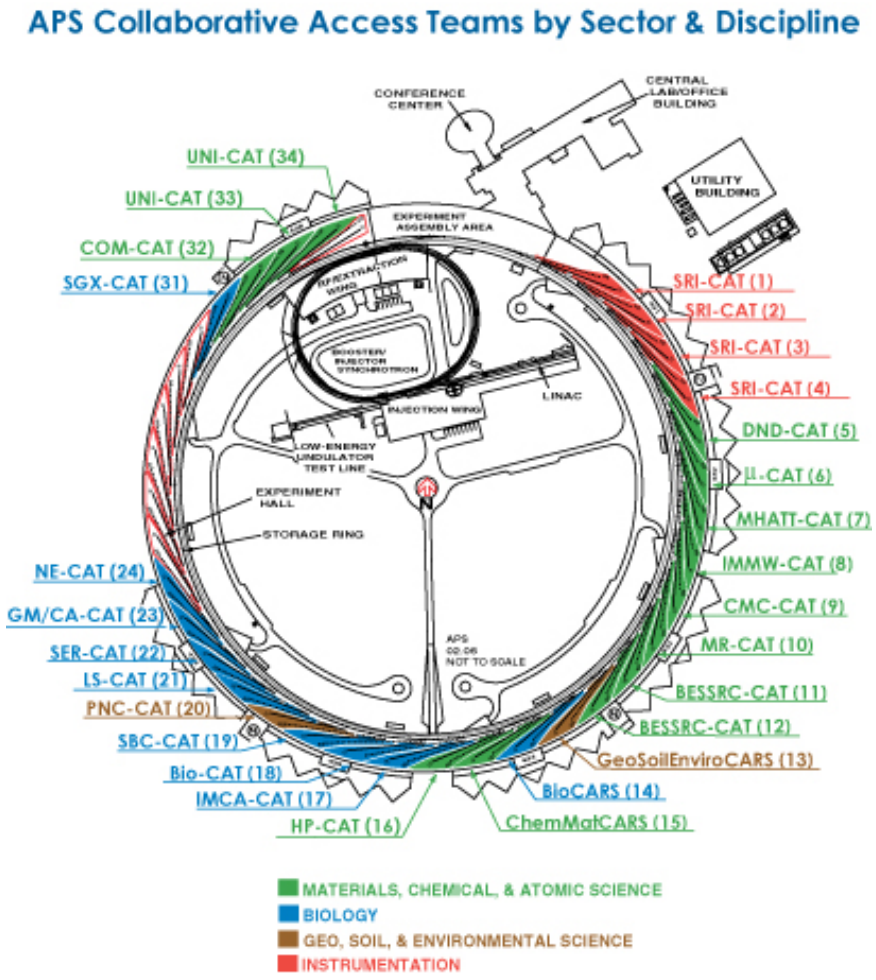
photographs suggests a cross pattern centred on the *c* axis and shows a cluster of very strong reflections at ~3.5 Å along the *c* axis, both implying that this is close to the direction of the DNA helix axis. Intensities remain strong in all directions out to 2.0 Å, and they exhibit rapid decline parallel essentially to that axis beyond 1.9 Å. Of the 4,691 possible reflections to 1.9 Å resolution, 2,818 were found to have an intensity greater than 2σ and were used in the analysis. Two isomorphous heavy atom derivatives were used: *cis*-dichlorodiamino platinum (II) obtained by diffusion, and a 3-Br derivative obtained by *de novo* synthesis of the dodecamer with 5-bromocytosine in the third position along each chain. The 1-Br derivative was crystallized but proved not to be isomorphous, and the 9-Br derivative was synthesized but not needed. Isomorphism in the *cis*-Pt derivative began to fail beyond 4-Å resolution, but the 3-Br derivative remains isomorphous to 2.7 Å.

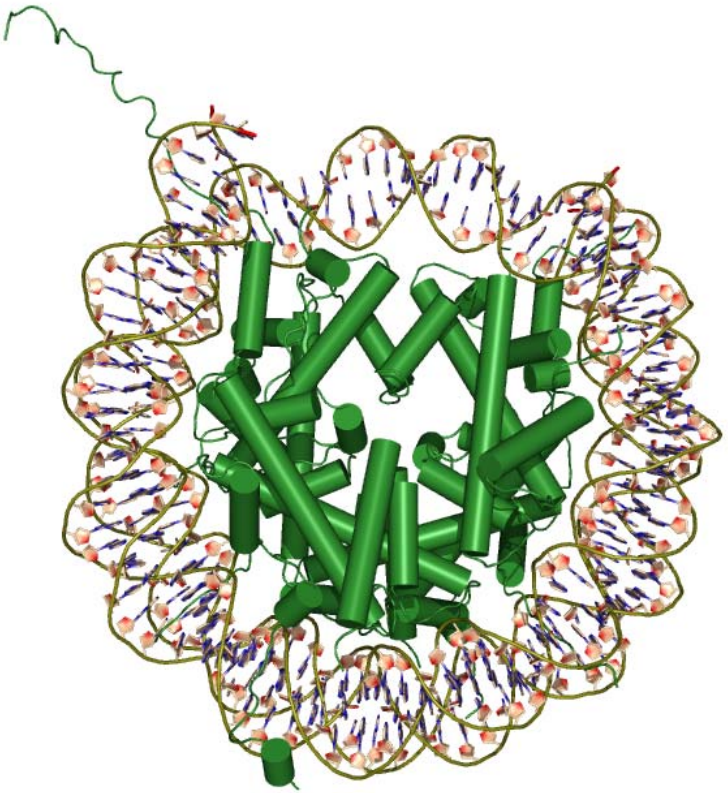
The present report describes the partially refined structure obtained from multiple isomorphous replacement (MIR) analysis at 2.7 Å (mean figure of merit 57%), followed by Jack–Levitt refinement procedures[3] using 2,725 2σ intensities between 8.0 and 1.9 Å. The current residual error or *R* factor is 24.8% for a DNA molecule of 486 atoms and 9 initial water molecules. The structure of the DNA itself is essentially correct and is reported now because of its general interest. Refinement will continue with the addition of more solvent and spermine atoms, and some improvement in local nucleotide conformations.

A skeletal drawing of CGCGAATTCGCG is presented in Fig. 1, and a space-filling version from the same orientation in Fig. 2. The molecule is a Watson–Crick B helix with an average rise per residue of 3.4 Å, and 10.1 base pairs per turn. This is somewhat less than the 10.4 ± 0.1 base pairs per turn as measured in solution by Wang[4] or calculated by Levitt[5] from energy considerations.

The deoxyribose rings in our structure have been fitted to electron density in the MIR map and then allowed to refine; no attempt has been made to impose a uniform C2′-*endo* or C3′-*exo* geometry on them. At this intermediate stage of refinement they are about equally distributed between C2′-*endo* and the O1′-*endo* configuration that is related by rocking about the glycosyl bond, but no conclusions about sugar puckering should be drawn until refinement is completed. No sequence-dependent variations in the uniform B helix have been seen, but these will be watched for as refinement continues.

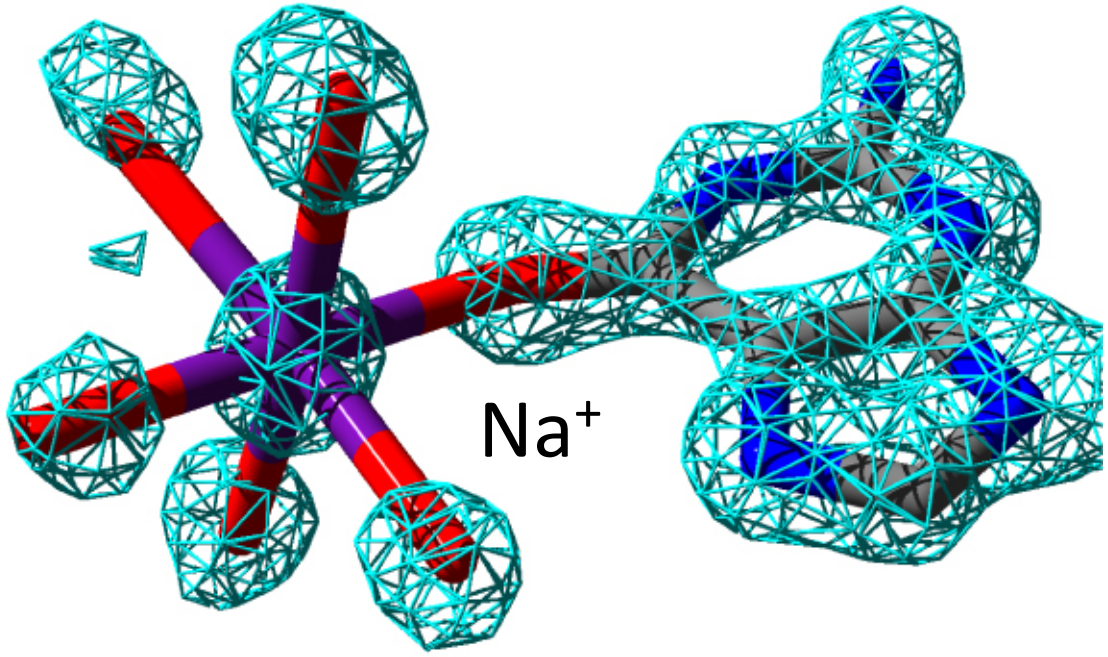
Phase 3: Synchrotron Radiation
A) Large Assemblies



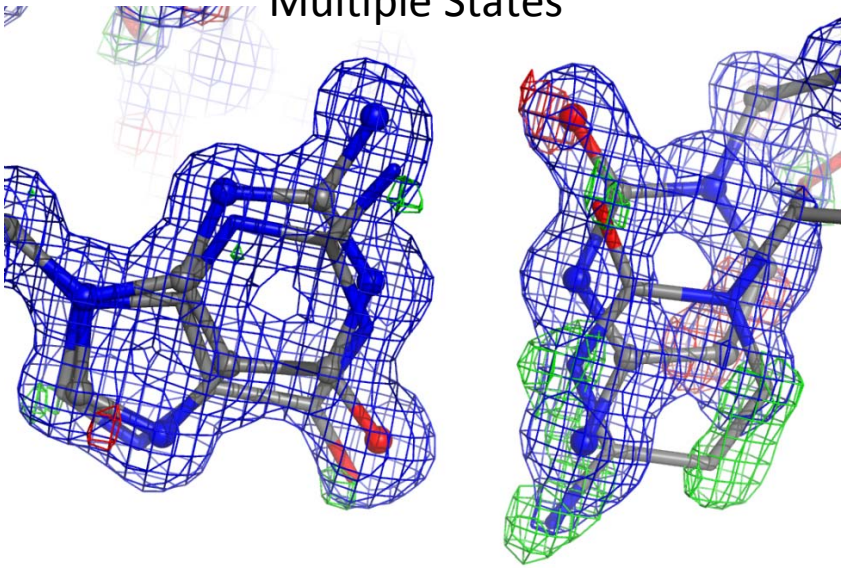


Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., & Richmond, T.J., Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389 (6648), 251-260 (1997).

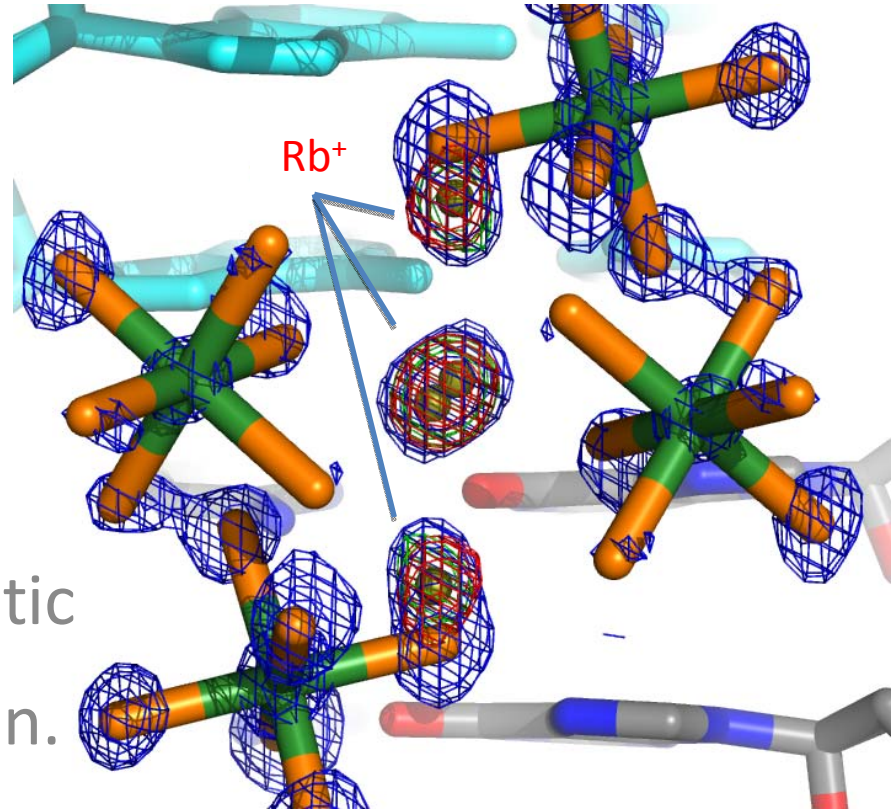
Phase 3: Synchrotron Radiation
B) Ultrahigh resolution/multiwavelength

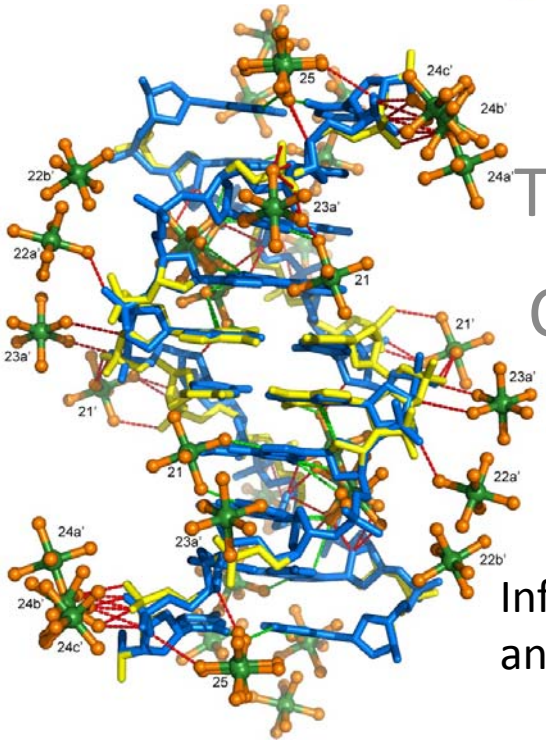


Multiple States



Identification of Group I ions
by anomalous methods





Goal:

Total Electrostatic

Characterization.

Inference of correlations between ion positions
and DNA conformation

General Rules of Engagement between B-DNA and Cations

- **General Rules**
  EVERYTHING* competes for common sites in the grooves of DNA.
  a) For CG base pairs, cations localize predominantly in the major groove.
  b) For AT base pairs, cations localize predominantly in the minor groove.
  c) Group I readily dehydrates. Group II does not.
  d) Interaction preference :O > :N > H_2N-.
  (i.e., keto O >imine>amino)

- [Group I (Na^+, K^+ etc), Group II hydrates [$Mg^{2+}(H_2O)_6$, $Ca^{2+}(H_2O)_6$], polyamines, and cationic amino acid sidechains]

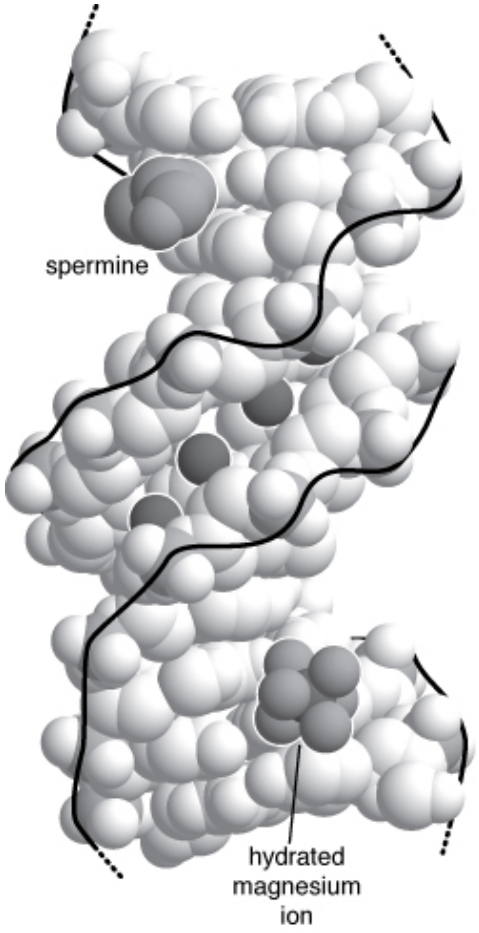
Exceptions to the General Rules of Engagement between B-DNA and Cations

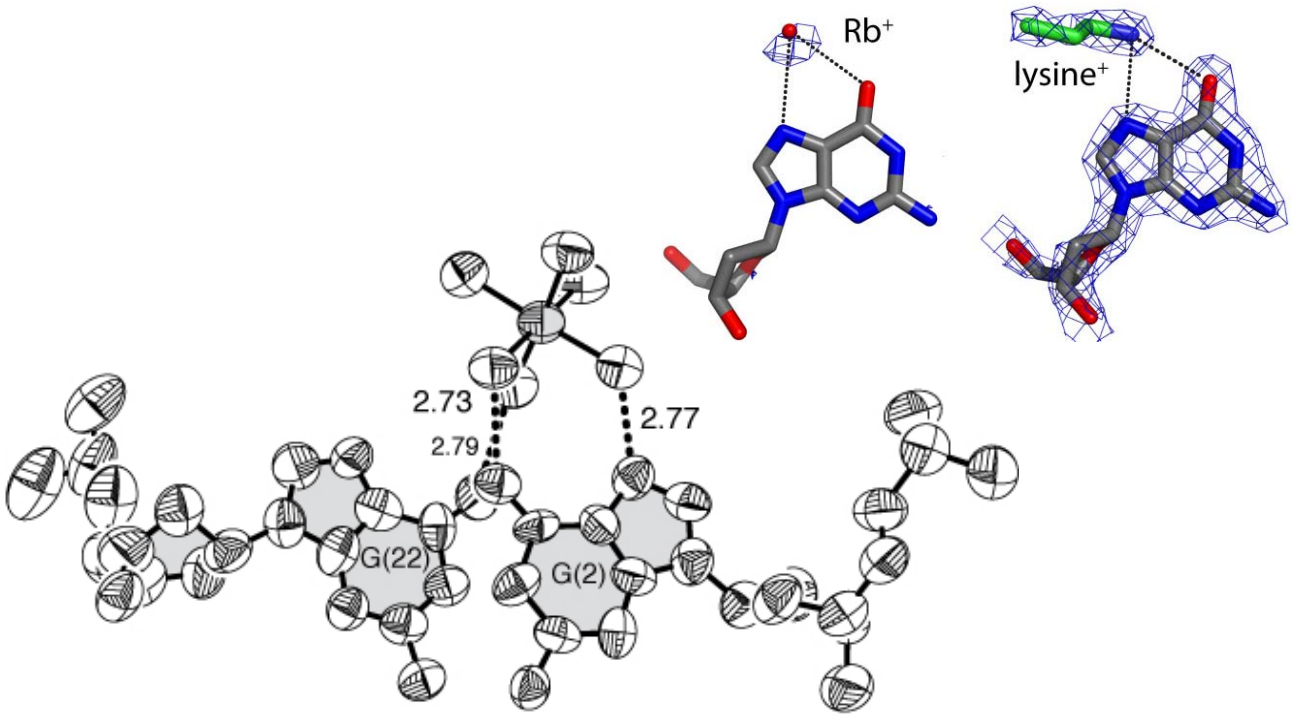
a) When the minor groove is especially narrow (i.e., A-tracts), $Mg^{2+}(H_2O)_6$ is not found inside the minor groove.

b) When the minor groove is especially narrow, $Mg^{2+}(H_2O)_6$ localizes in the lip of the groove.

c) $Mg^{2+}(H_2O)_6$ associates with phosphate groups (the other ions do not).

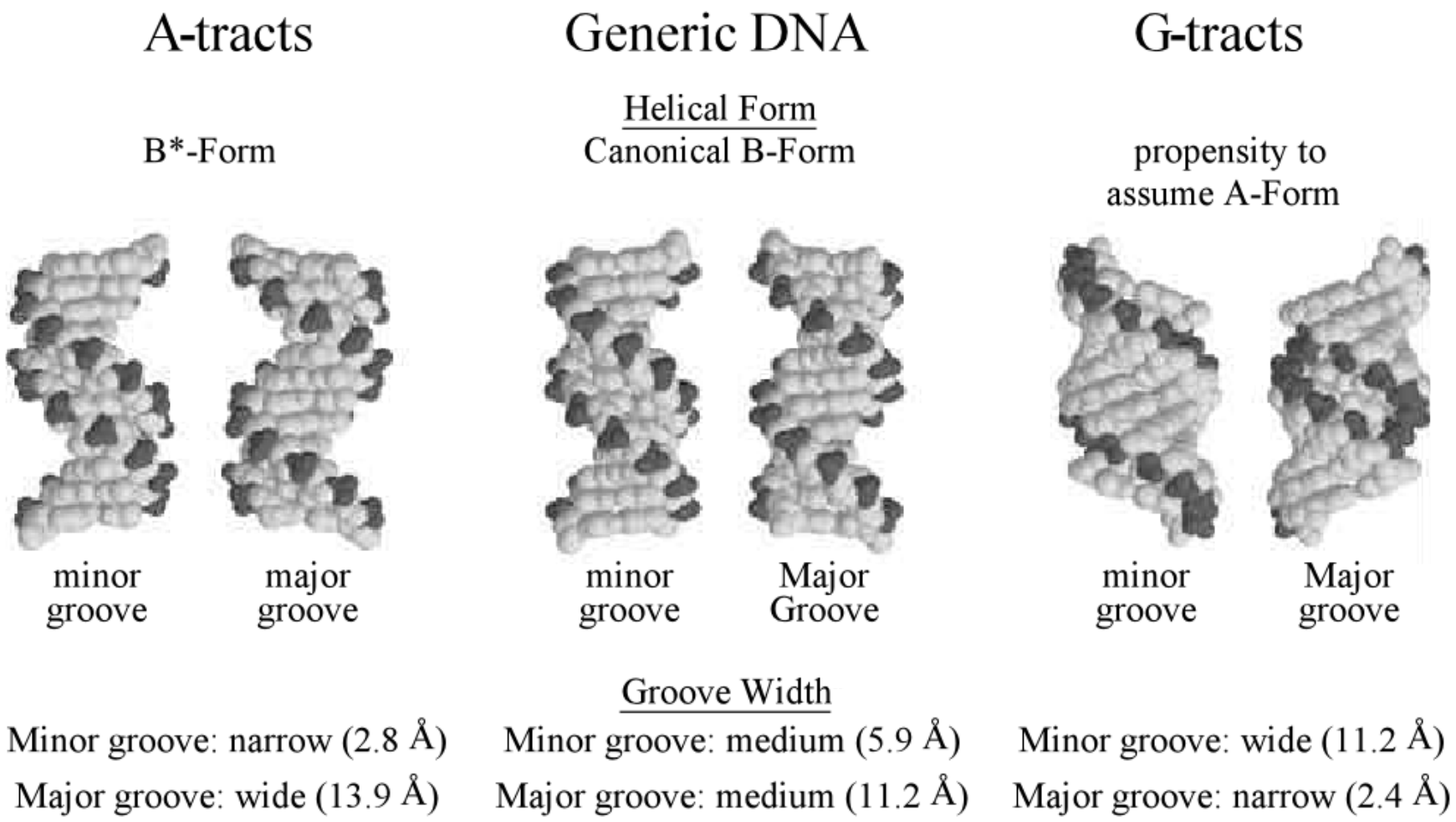
d) Polyamines appear to avoid the minor groove.





Hsiao, C. et al., Complexes of Nucleic Acids with Group I and II Cations in Nucleic Acid Metal Ion Interactions, edited by N. Hud (The Royal Society of Chemistry, London, 2008), pp. 1-35.
Gold, B., Marky, L.M., Stone, M.P., & Williams, L.D., A review of the role of the sequence-dependent electrostatic landscape in DNA alkylation patterns. Chem Res Toxicol. 19 (11), 1402-1414 (2006).
Moulaei, T. et al., Structure of B-DNA with cations tethered in the major groove. Biochemistry 44 (20), 7458-7468 (2005).
Howerton, S.B., Nagpal, A., & Williams, L.D., Surprising roles of electrostatic interactions in DNA-ligand complexes. Biopolymers 69 (1), 87-99 (2003).
Hamelberg, D., Williams, L.D., & Wilson, W.D., The Effect of Neutralized Phosphate Backbone on the Minor Groove Structure of B-DNA. Nucleic Acids Res. 30, 3615-3623 (2002).
Howerton, S.B., Sines, C.C., VanDerveer, D., & Williams, L.D., Locating Monovalent Cations in the Grooves of B-DNA. Biochemistry 40 (34), 10023-10031 (2001).
Woods, K. et al., Monovalent Cations Sequester within the A-tract Minor Groove of $[d(CGCGAATTCGCG)]_2$. J. Am. Chem. Soc. 122, 1546-1547 (2000).
Williams, L.D. & Maher, L.J., Electrostatic Mechanisms of DNA Deformation. Annu. Rev. Biophys. Biomol. Struct. 29, 497-521 (2000).
Sines, C.C., McFail-Isom, L., Howerton, S.B., VanDerveer, D., & Williams, L.D., Cations mediate B-DNA conformational heterogeneity. J. Am. Chem. Soc. 122, 11048-11056 (2000).
Shui, X. et al., Effects of Cationic Charge on Three-Dimensional Structures of Intercalative Complexes: Structure of a bis-Intercalated DNA Complex Solved by MAD Phasing. Current Med. Chem. 7, 59-71 (2000).
Hamelberg, D., McFail-Isom, L., Williams, L.D., & Wilson, W.D., The Flexible Structure of DNA: Ion Dependence of Minor-Groove Structure and Dynamics. J. Am. Chem. Soc. 122, 10513-10520 (2000).
Shui, X., Sines, C., McFail-Isom, L., VanDerveer, D., & Williams, L.D., Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. Biochemistry 37, 16877-16887 (1998).
Shui, X., McFail-Isom, L., Hu, G.G., & Williams, L.D., The B-DNA Dodecamer at High Resolution Reveals a Spine of Water on Sodium. Biochemistry 37, 8341-8355 (1998).
McFail-Isom, L., Shui, X., & Williams, L.D., Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base P systems. Biochemistry 37, 17105-17110 (1998).

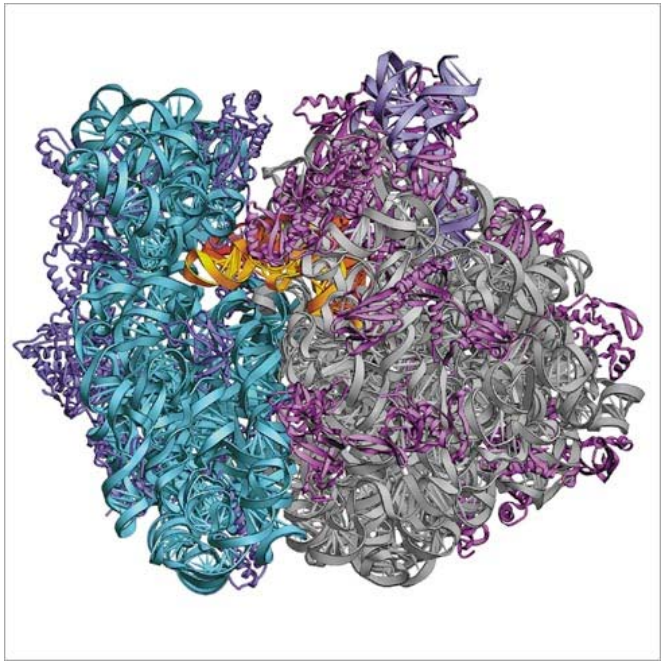
Ion Localization and Helical Form (B', B and A)

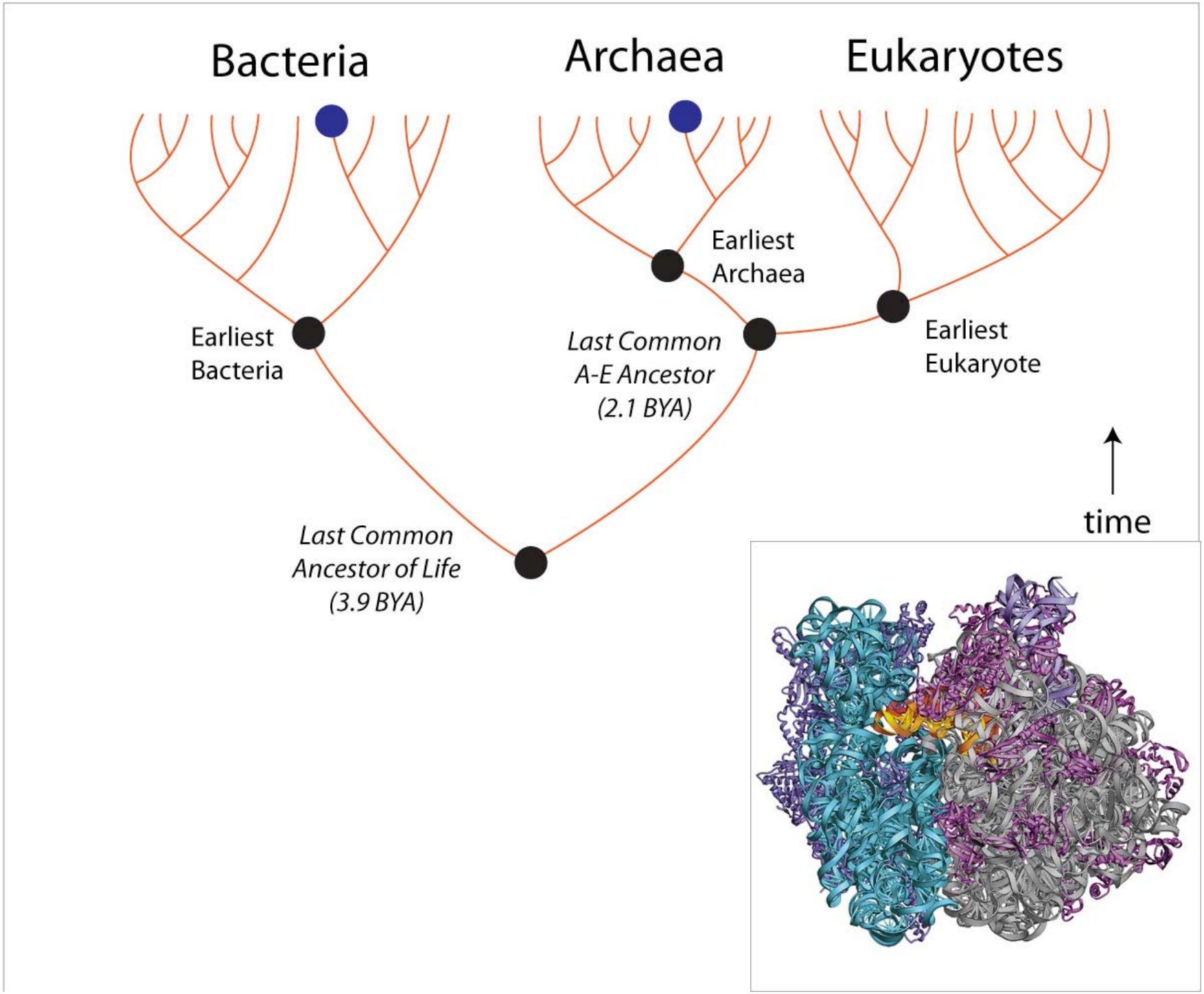


Hud, N.V. & Polak, M., DNA-cation interactions: The major and minor grooves are flexible ionophores. Curr. Opin. Struct. Biol. 11 (3), 293-301 (2001).

The Electrostatics, Cations and Molecular Interactions in the Ribosome

- Wimberly, B.T. et al., Structure of the 30S ribosomal subunit. Nature 407 (6802), 327-339 (2000).
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., & Steitz, T.A., The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science 289 (5481), 905-920 (2000).
- Harms, J. et al., High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell 107 (5), 679-688 (2001).
- Korostelev, A., Trakhanov, S., Laurberg, M., & Noller, H.F., Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. Cell 126 (6), 1065-1077 (2006).
- Selmer, M. et al., Structure of the 70S ribosome complexed with mRNA and tRNA. Science 313 (5795), 1935-1942 (2006).




Bacteria
Archaea
Eukaryotes
Earliest
Archaea
Earliest
Bacteria
Last Common
A-E Ancestor
(2.1 BYA)
Earliest
Eukaryote
time
Last Common
Ancestor of Life
(3.9 BYA)

Here we illustrate the futility of further effort and quickly change the topic.

Nucleic Acids Research Advance Access published March 11, 2009

Nucleic Acids Research, 2009, 1–9
doi:10.1093/nar/gkp119

A recurrent magnesium-binding motif provides a framework for the ribosomal peptidyl transferase center

Chiaolong Hsiao and Loren Dean Williams*

School of Chemistry and Biochemistry, Parker H. Petit Institute of Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332-0400, USA

Received October 23, 2008; Revised February 7, 2009; Accepted February 11, 2009

ABSTRACT

The ribosome is an ancient macromolecular machine responsible for the synthesis of all proteins in all living organisms. Here we demonstrate that the ribosomal peptidyl transferase center (PTC) is supported by a framework of magnesium microclusters (Mg^{2+}-μc's). Common features of Mg^{2+}-μc's include two paired Mg^{2+} ions that are chelated by a common bridging phosphate group in the form $Mg^{2+}_{(a)}$–(O1P-P-O2P)–$Mg^{2+}_{(b)}$. This bridging phosphate is part of a 10-membered chelation ring in the form $Mg^{2+}_{(a)}$–(OP-P-O5′-C5′-C4′-C3′-O3′-P-OP)–$Mg^{2+}_{(a)}$. The two phosphate groups of this 10-membered ring are contributed by adjacent residues along the RNA backbone. Both Mg^{2+} ions are octahedrally coordinated, but are substantially dehydrated by interactions with additional RNA phosphate groups. The Mg^{2+}-μc's in the LSU (large subunit) appear to be highly conserved over evolution, since they are unchanged in bacteria (*Thermus thermophilus*, PDB entry 2J01) and archaea (*Haloarcula marismortui*, PDB entry 1JJ2). The 2D elements of the 23S rRNA that are linked by Mg^{2+}-μc's are conserved between the rRNAs of bacteria, archaea and eukarya and in mitochondrial rRNA, and in a proposed minimal 23S-rRNA. We observe Mg^{2+}-μc's in other rRNAs including the bacterial 16S rRNA, and the P4–P6 domain of the tetrahymena Group I intron ribozyme. It appears that Mg^{2+}-μc's are a primeval motif, with pivotal roles in RNA folding, function and evolution.

INTRODUCTION

Recent structures of large RNAs, such as the P4–P6 domain of the tetrahymena ribozyme (1–5), and larger RNAs such as ribosomes (6–14) continue to reveal important new principles of RNA structure and folding. Globular cores of large RNAs are characterized by base–base tertiary interactions and 'buried' phosphate groups (15,16). During RNA-folding cations are sequestered from bulk solvent, and held in close proximity to the polymer. Folding increases proximities of phosphate groups, and the electrostatic repulsion among them. Thus phosphate–phosphate repulsion must be offset by attraction between phosphates and cations.

Mg^{2+}, since the beginning of life, has been closely associated with some of the central players in biological systems—phosphates and phosphate esters (17). Mg^{2+} shares a special geometric, electrostatic, thermodynamic relationship with phosphates (18). The ionic radius of Mg^{2+} is small (0.65 Å), the charge density is high, the coordination geometry is octahedral (the AOCN—average observed coordination number—is 5.98), the preferred ligands are charged or neutral oxygens, and the hydration enthalpy is large (−458 kcal/mol) (19–22). Ligand–ligand crowding is one of the hallmarks of Mg^{2+} complexes, leading to highly restrained ligand–Mg^{2+}–ligand geometry, and strong ligand–ligand repulsive forces. In comparison with group I ions, Ca^{2+} or polyamines, Mg^{2+} has a much greater affinity for phosphate oxygens, and binds to them with well-defined geometry. Unlike other cations, Mg^{2+} brings phosphate oxygens in its first shell into direct contact with each other.

Chelation effects and topology influence interactions of nucleic acids with ions. For example magnesium forms a *mononuclear* motif with ADP/ATP in which one Mg^{2+} ion is chelated by a six-membered ring consisting of atoms $Mg^{2+}_{(a)}$–O^{β}P-P-O-P-O^{β}P–$Mg^{2+}_{(a)}$ (18). Here we observe that a framework of *dinuclear* magnesium complexes $Mg^{2+}_{(a)}$–RNA–$Mg^{2+}_{(b)}$ flanks the peptidyl transferase center (PTC) in LSU ribosomal structures (Figure 1). In these dinuclear clusters, two Mg^{2+} ions are chelated by a common bridging phosphate group. Additional

*To whom correspondence should be addressed. Tel: +1 404 894 9752; Fax: +1 404 894 7452; Email: loren.williams@chemistry.gatech.edu

© 2009 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

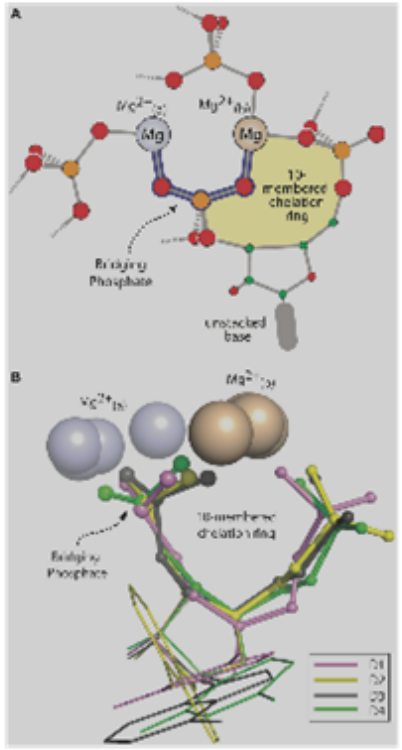


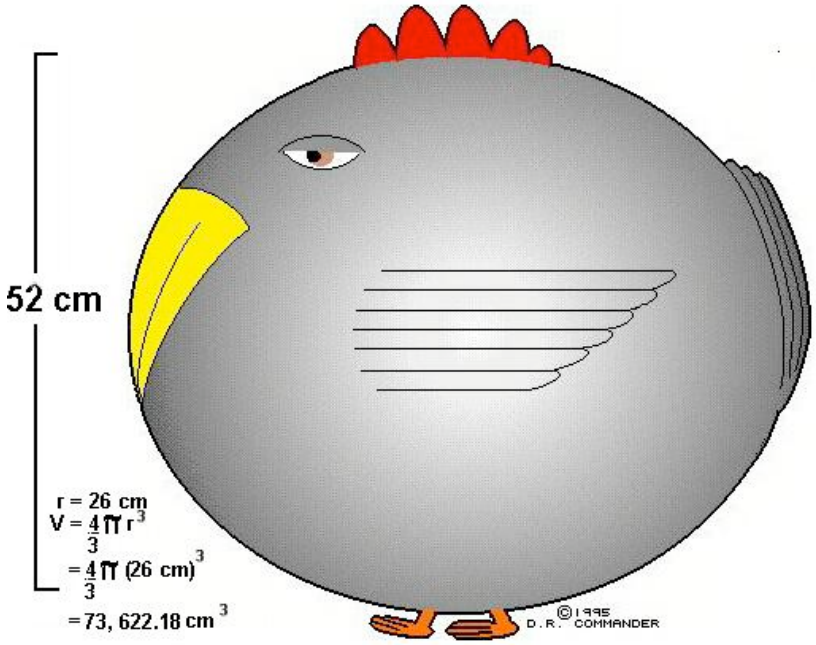
Figure 2. The Mg^{2+}-μc motif. (**A**) A schematic diagram illustrating the $Mg^{2+}_{(a)}$–(O1P-P-O2P)–$Mg^{2+}_{(b)}$ bridge (outlined in blue), the 10-membered chelation ring (yellow), an unstacked base and the additional RNA phosphate groups that enter the Mg^{2+} first shell at variable positions. Carbon is green, oxygen is red and phosphorous is orange. Magnesium (a) is cyan while magnesium (b) is brown. (**B**) Superimposition of four Mg^{2+}-μc's (D1 purple, D2 yellow, D3 gray and D4 green) from the *H. marismortui* LSU. All of the atoms shown here were used in the superimposition of the clusters except for the RNA bases.

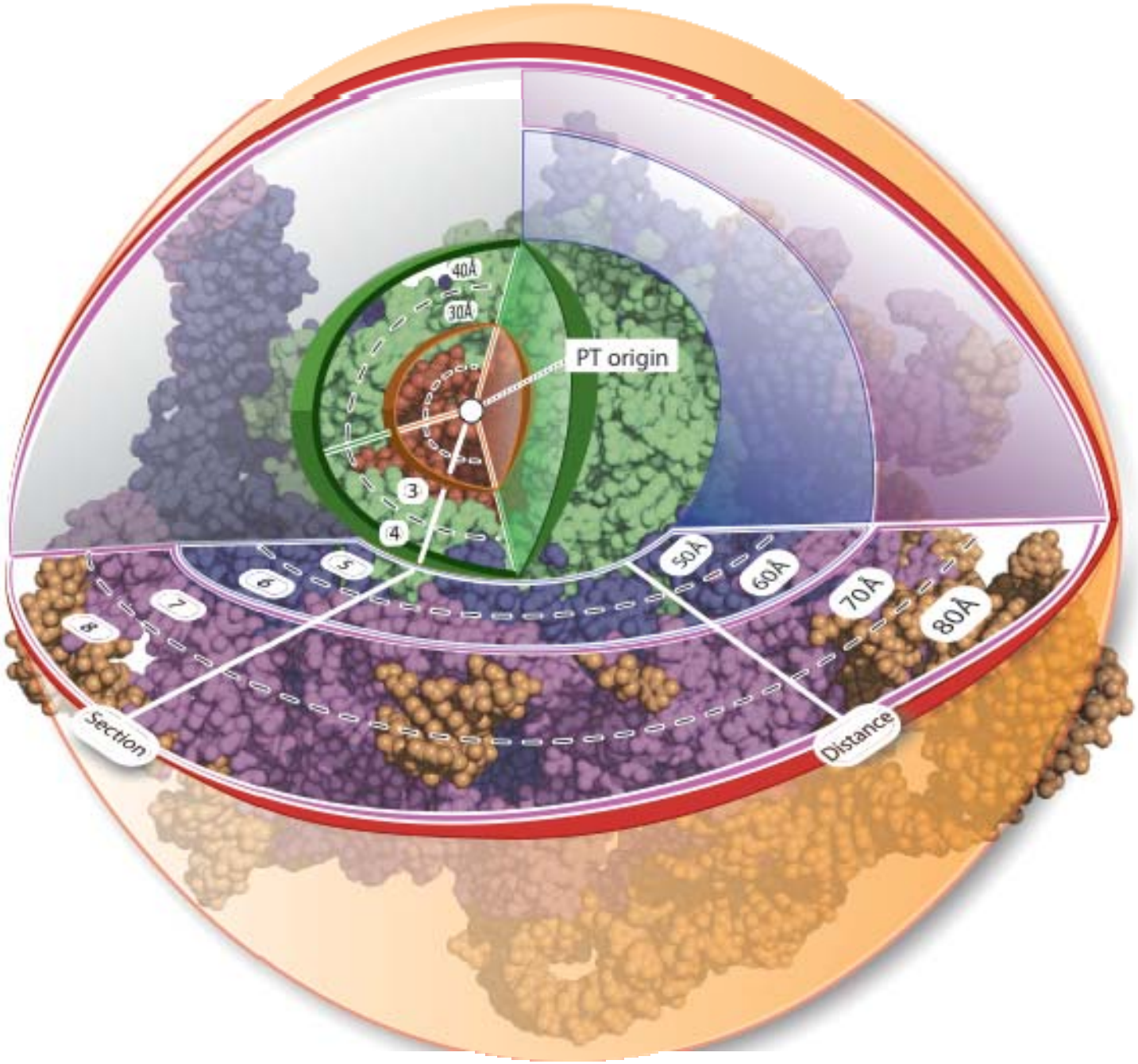
Figure 2B. The RMSD of the atomic positions of this superimposition is 0.59 Å.

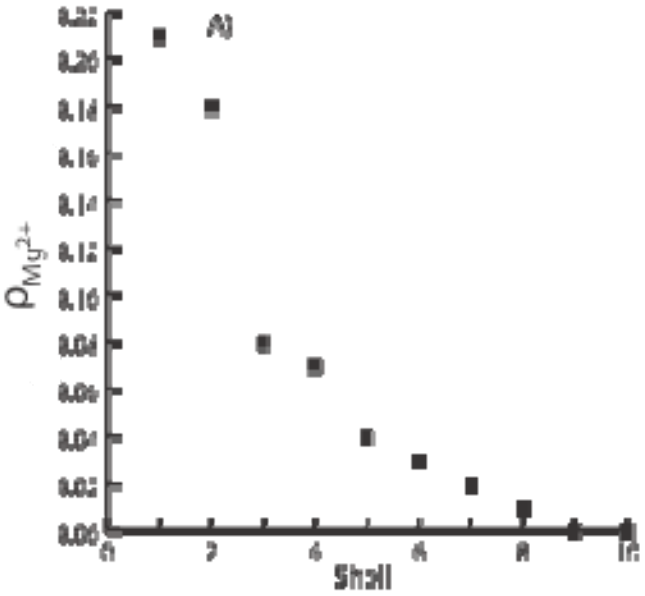
Although the basic motif is conserved in all four clusters, Mg^{2+}-μc's are elaborated by additional RNA ligands. RNA phosphate groups that are not part of the bridging phosphate or the 10-membered ring can enter the Mg^{2+} coordination sphere at various positions. Each Mg^{2+}-μc is defined *in toto* by two paired Mg^{2+} ions plus all RNA residues that engage in first shell interactions with the paired Mg^{2+} ions. Therefore each Mg^{2+}-μc has a unique shape. Specific descriptions of the four clusters are provided here.

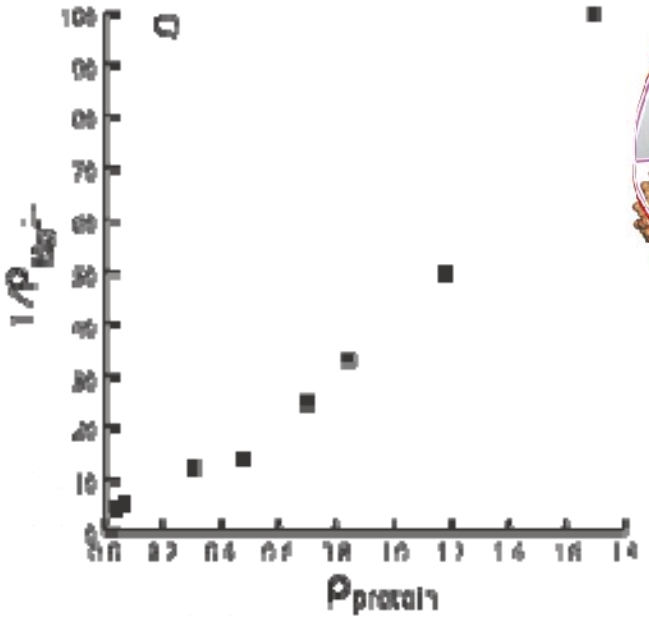
The Spherical Chicken

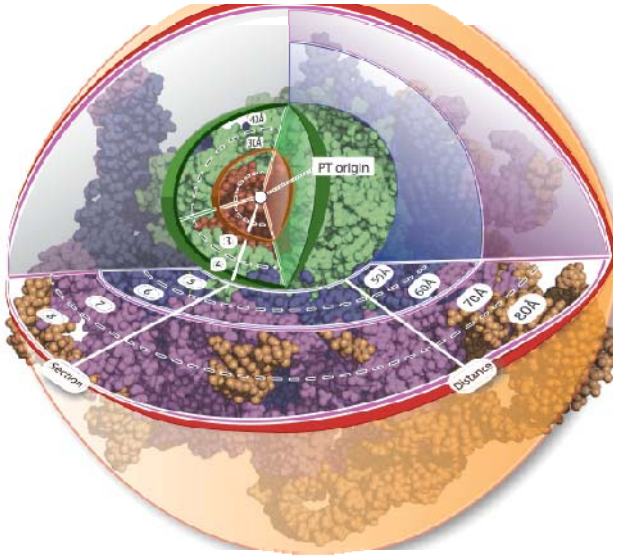
$\approx$

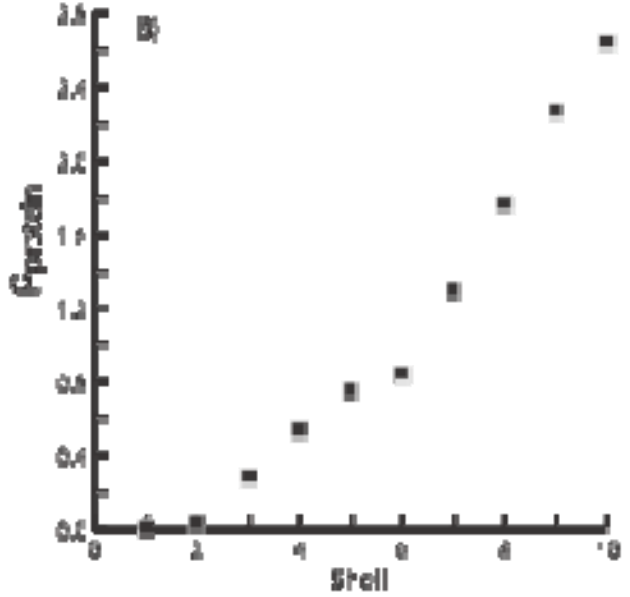


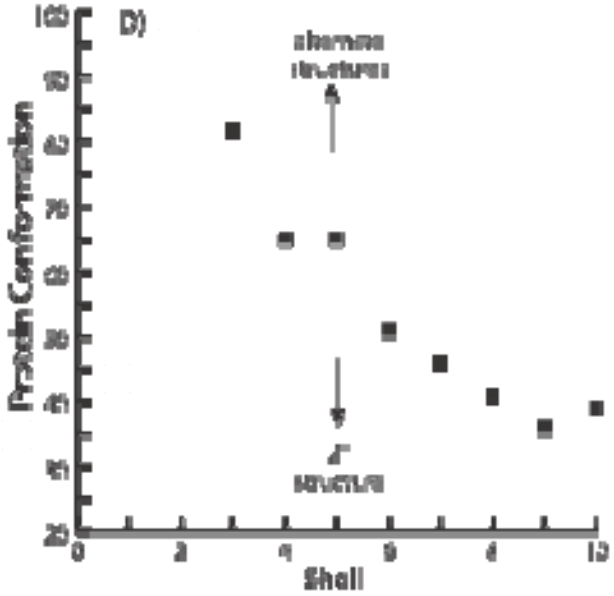

40Å
30Å
PT origin
3
4
5
6
7
8
50Å
60Å
70Å
80Å
Section
Distance

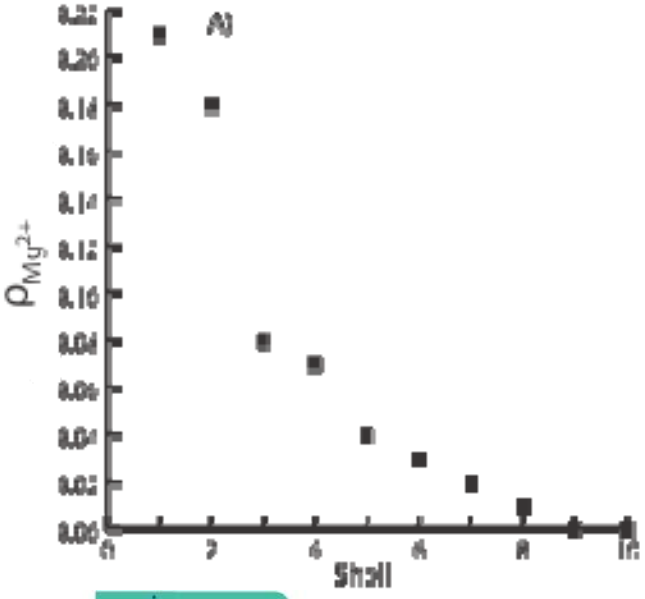

A)
$\rho_{Mg^{2+}}$
Shell

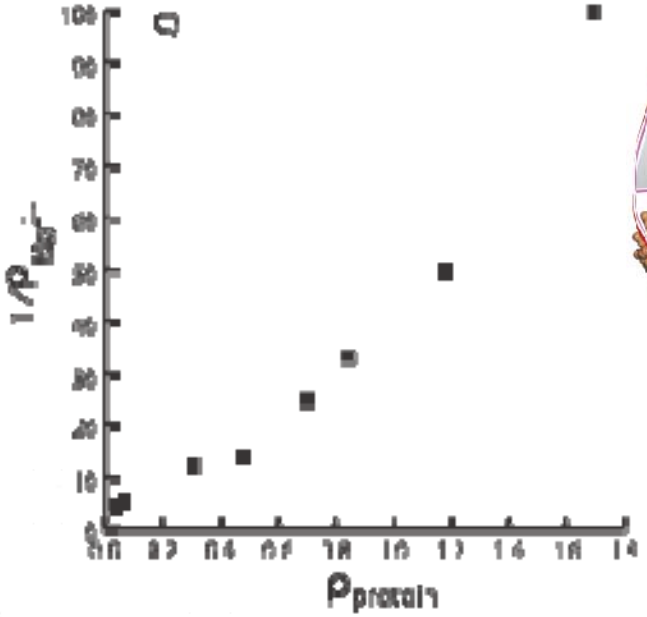

C)
$1/\rho_{Mg^{2+}}$
$\rho_{protein}$

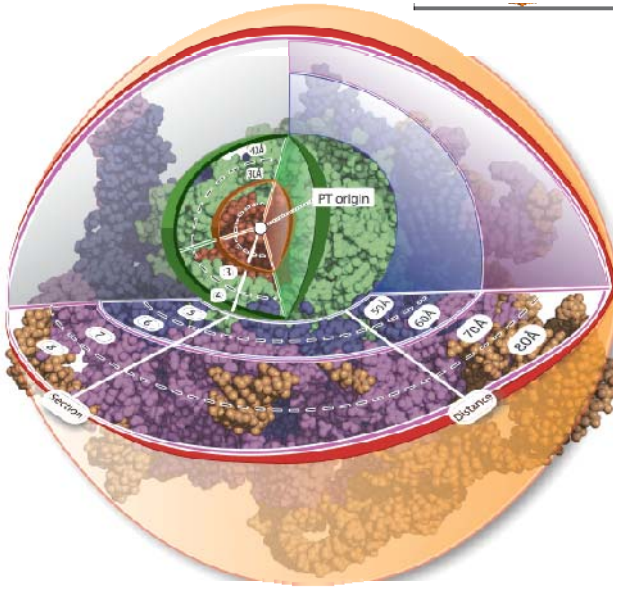

PT origin
Section
Distance

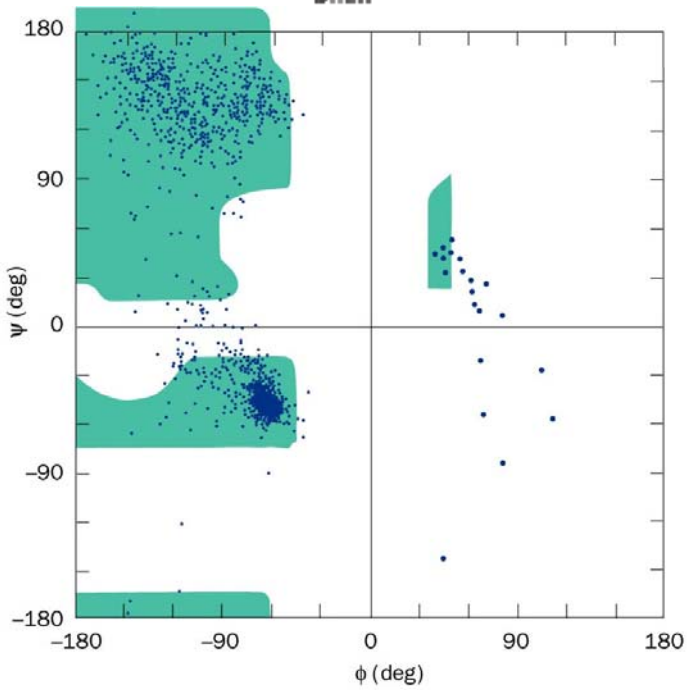

B)
$\rho_{protein}$
Shell

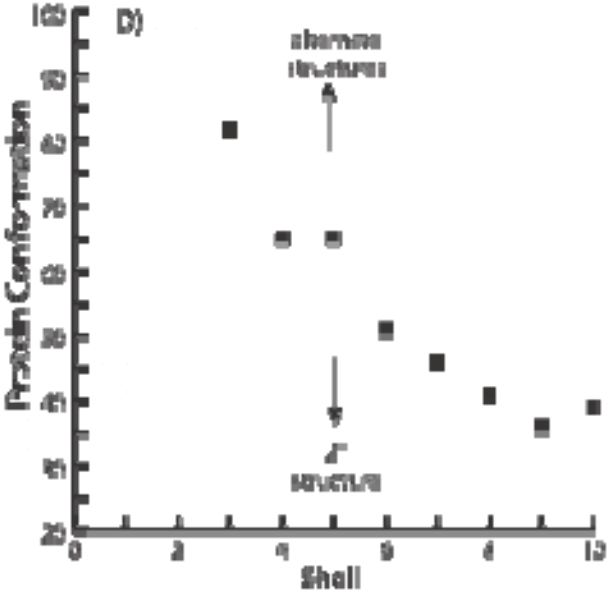

D)
Protein Conformation
Shell

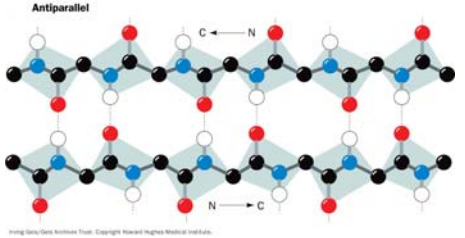

A)
Shell

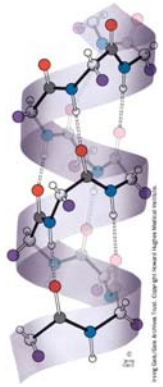

C)

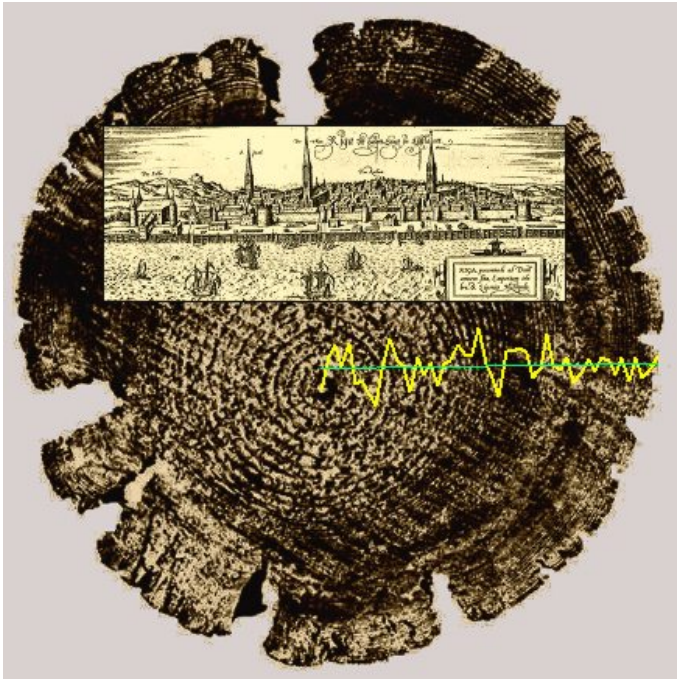

PT origin
Section
Distance

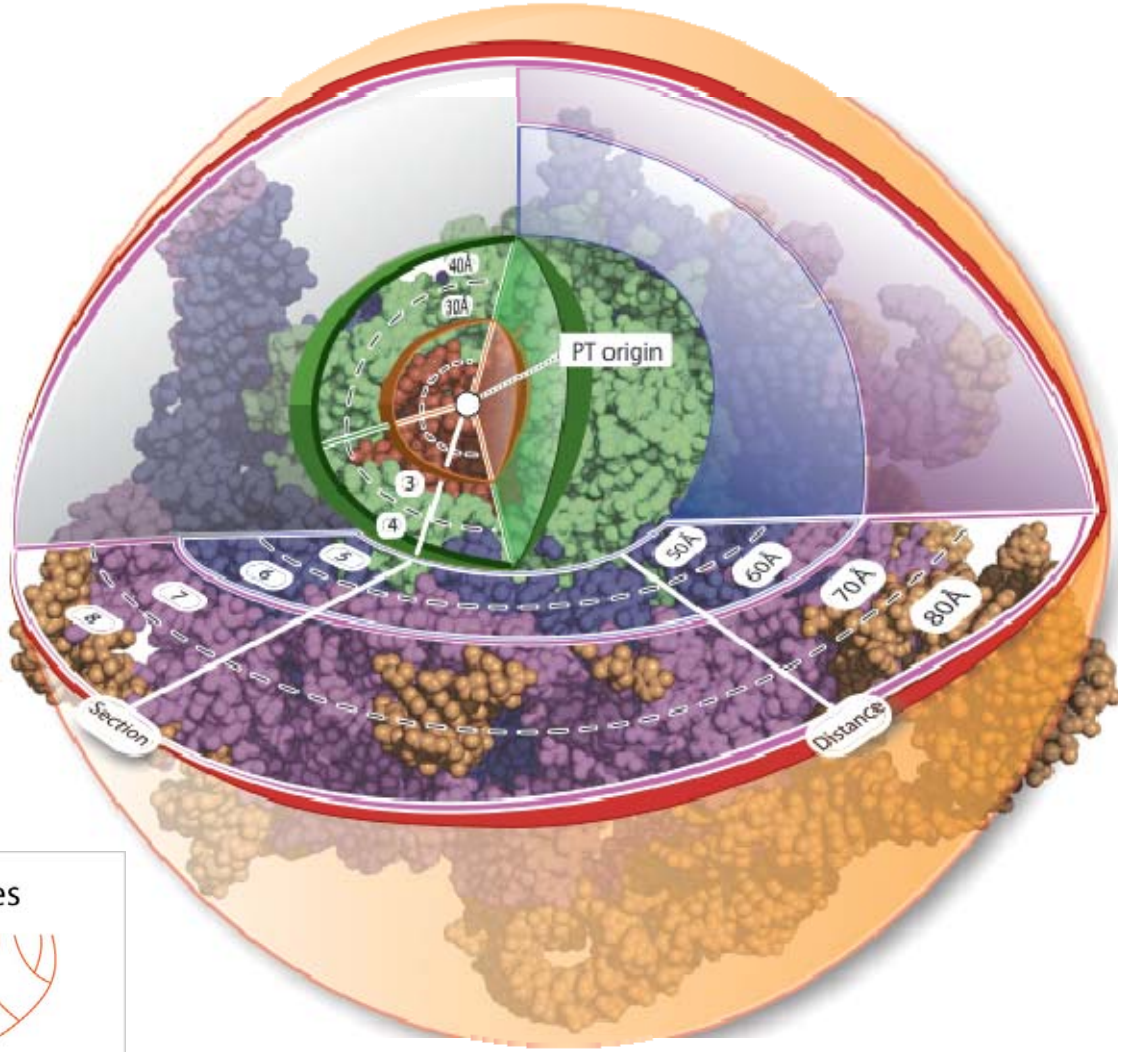

180
90
0
−90
−180
ψ (deg)
−180
−90
0
90
180
ϕ (deg)

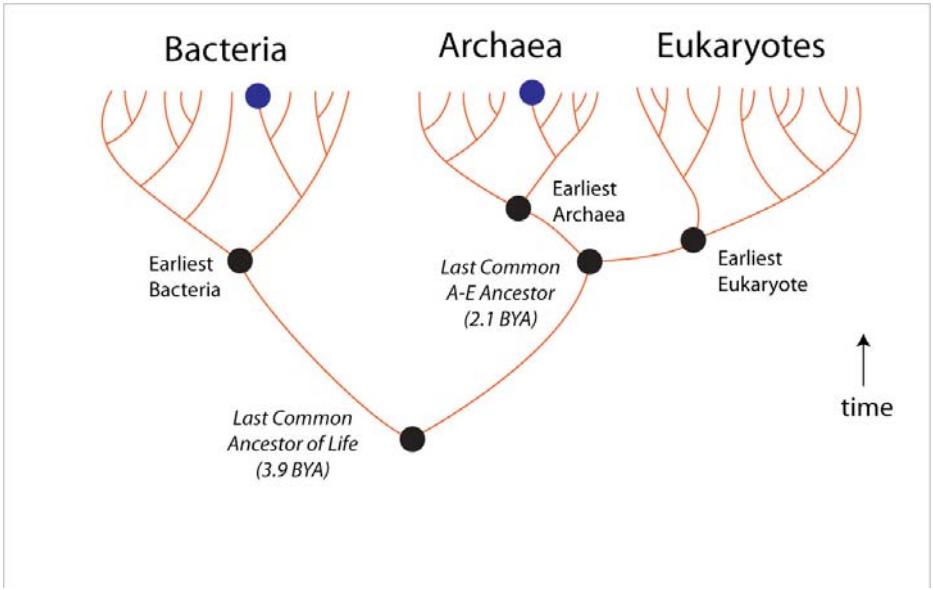

D)
Protein Conformation
Shell


Antiparallel
C
N
N
C






40Å
30Å
PT origin
3
4
5
6
7
8
50Å
60Å
70Å
80Å
Section
Distance


Bacteria
Archaea
Eukaryotes
Earliest Archaea
Earliest Bacteria
Last Common A-E Ancestor (2.1 BYA)
Earliest Eukaryote
Last Common Ancestor of Life (3.9 BYA)
time

Back to DNA-Protein Interaction…

The B’-state of DNA differs the B-state by

(i)Narrow minor groove,

(ii)Cooperative and monolithic minor groove hydration structure,

(iii)A requirement for sequence of at least three contiguous ApA or ApT steps (5’ to 3’),

(iv)TpA steps, G-C base pairs not allowed,

examples AAAA, AATT ATTT (yes)
TTAA, TATA, ATAT (no)
TCGA, CGCG (no, no)

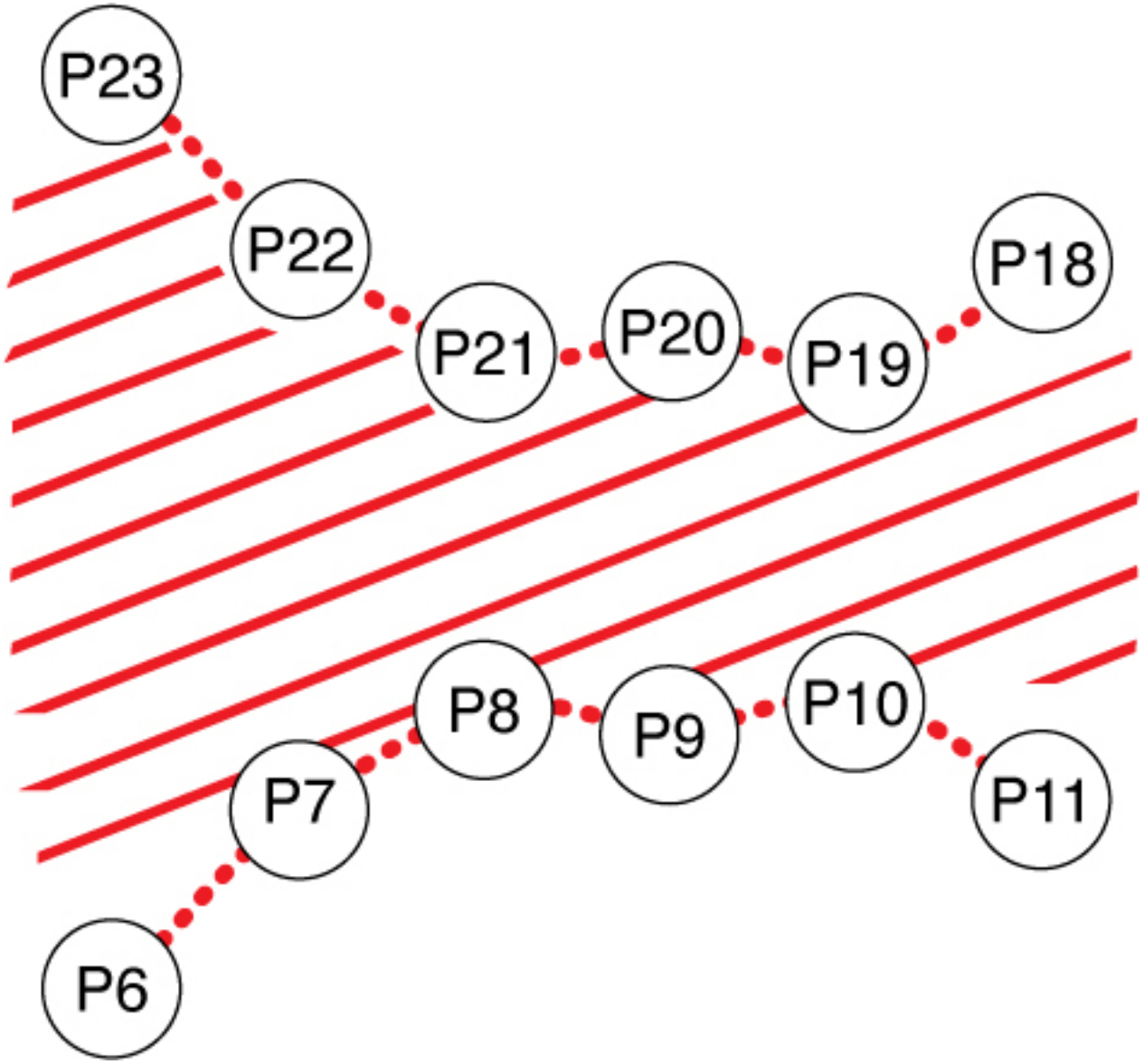
(v)DNA bending, when appropriately phased,

(vi)Monovalent ions within the minor groove,

(vii)Negative propeller twisting,

(viii)Rigidity,

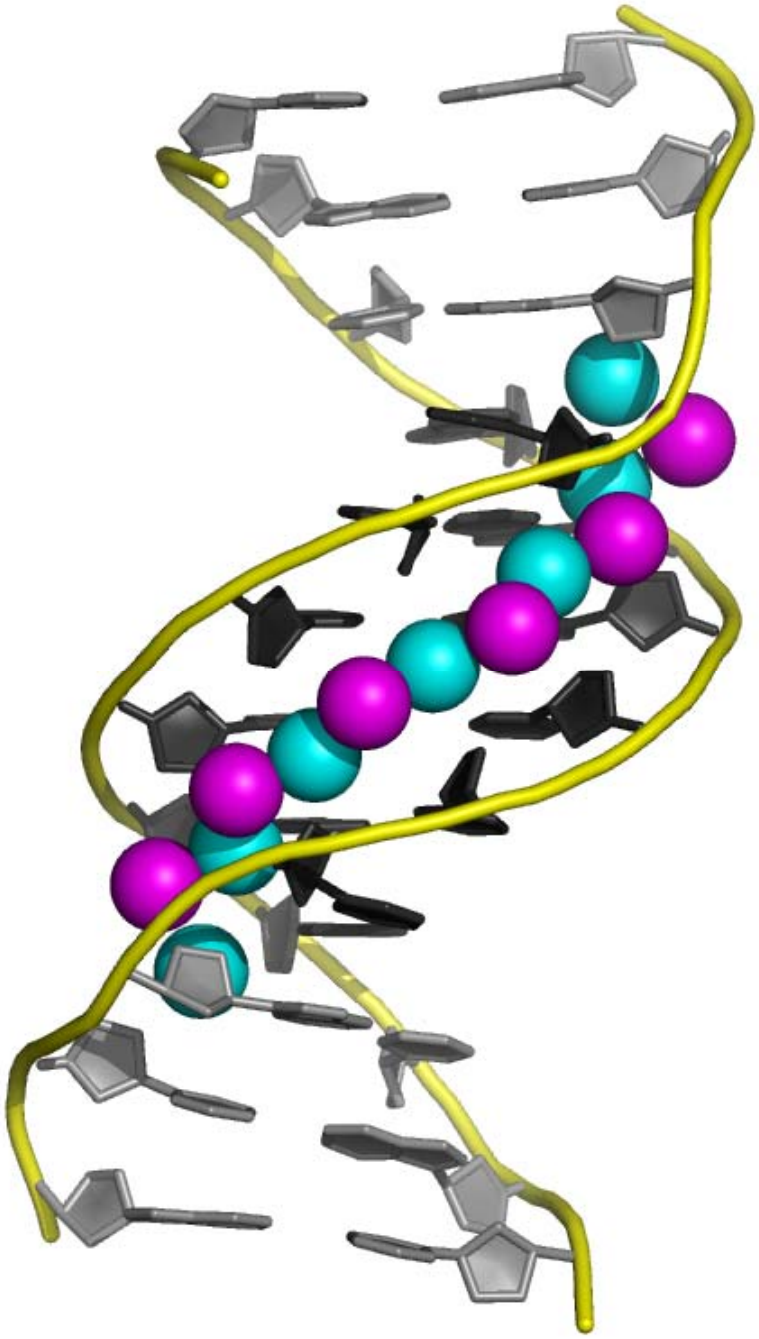
(ix)A‘pre-melting’ thermal transition that is distinct from duplex melting



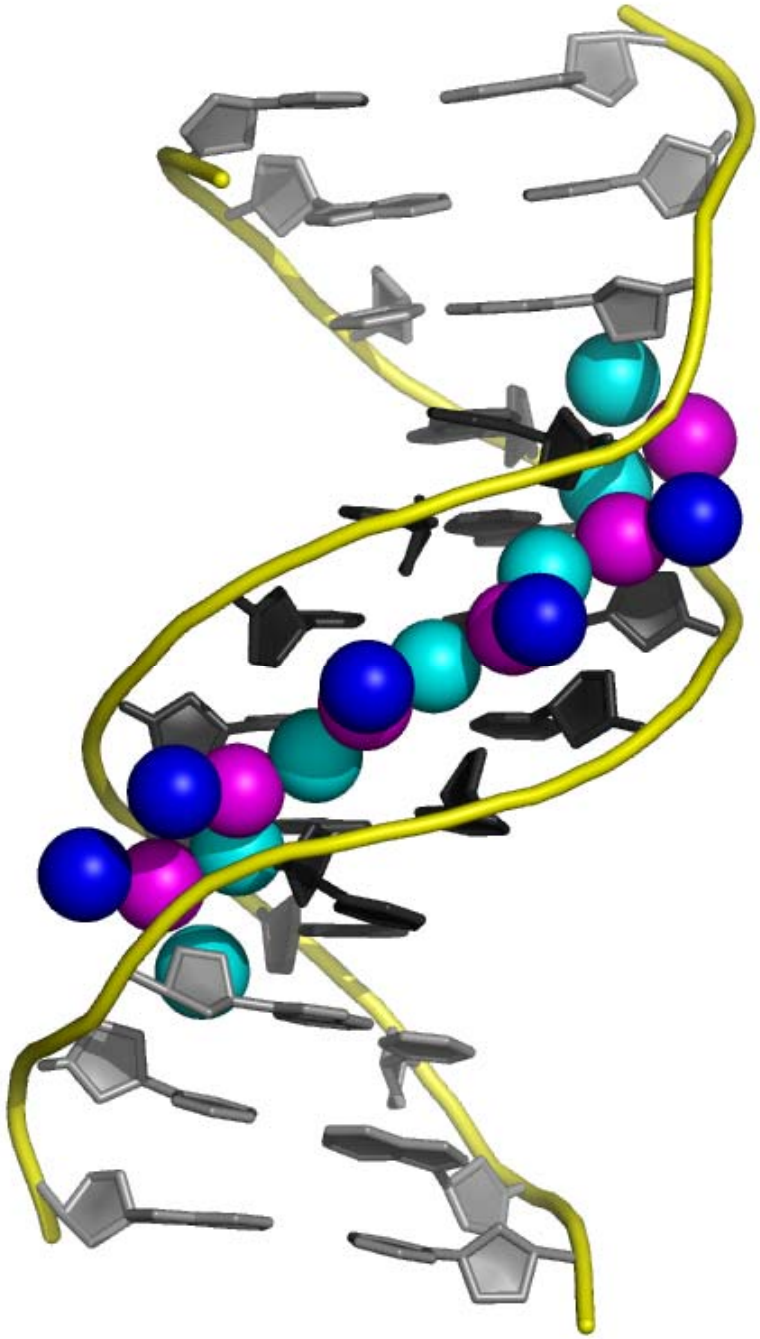
DNA Minor Groove Width

It is easy to determine the minor groove width of a DNA fragment, if you have the 3D structure.

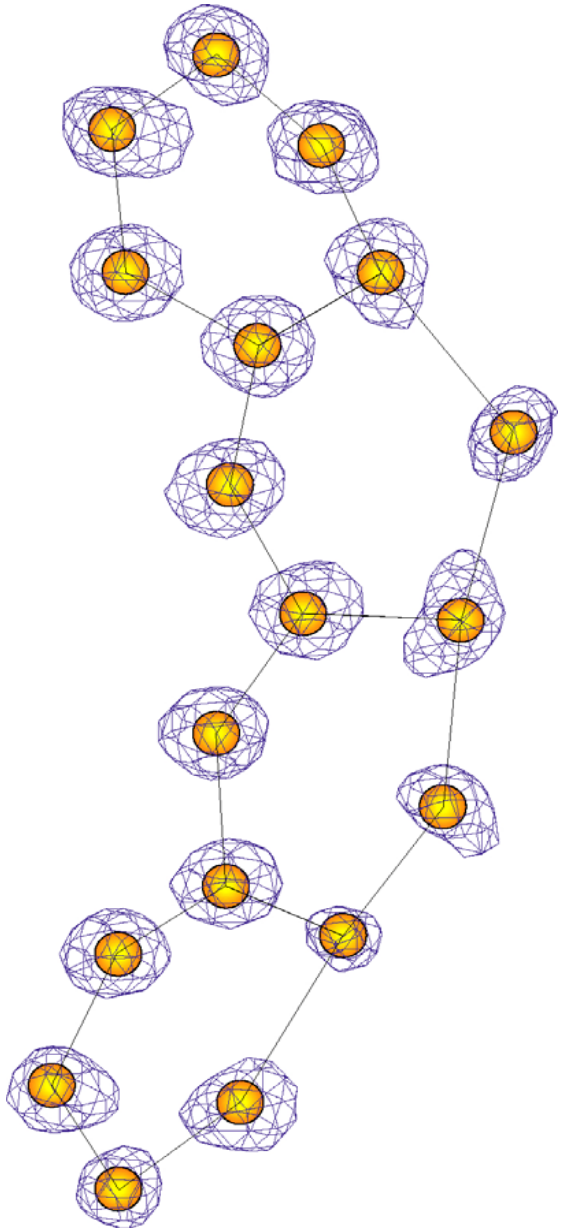
(thanks to programs by Wilma Olson and Richard Lavery)

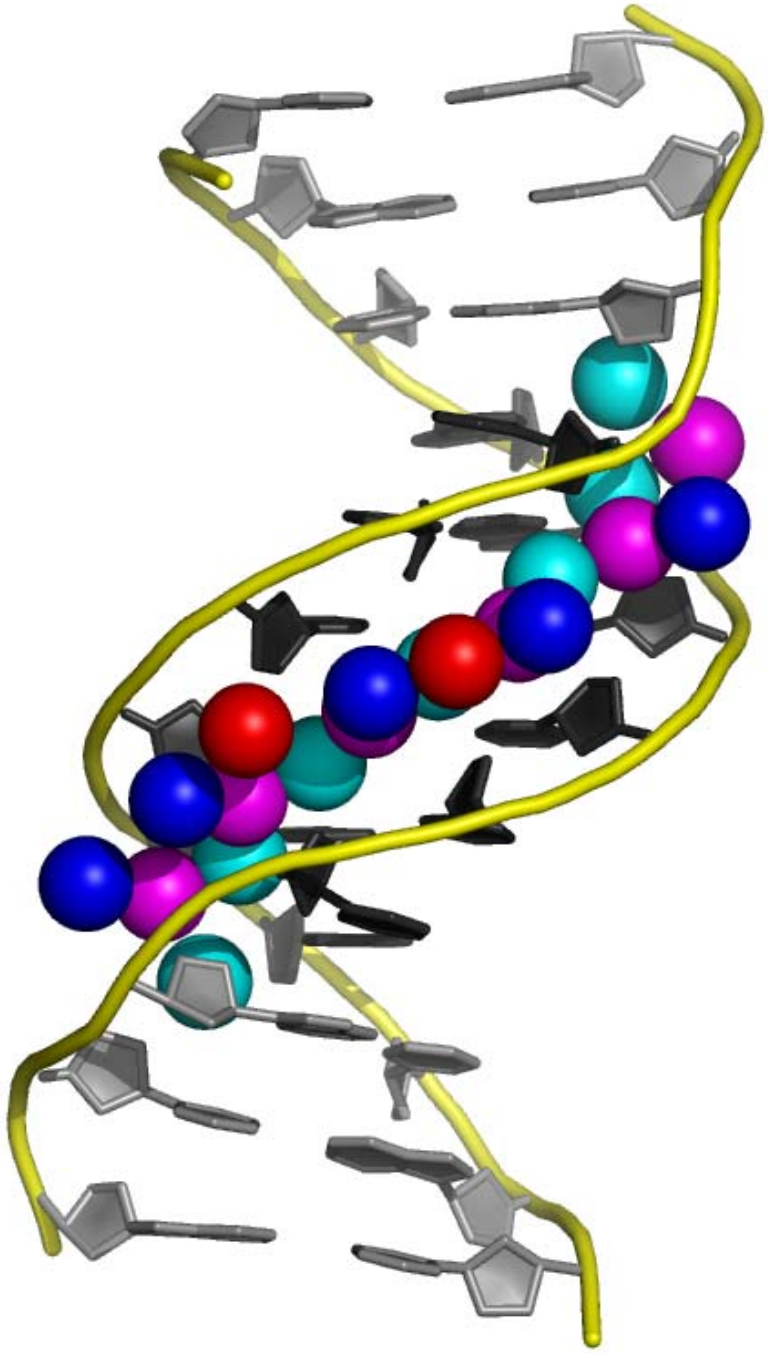


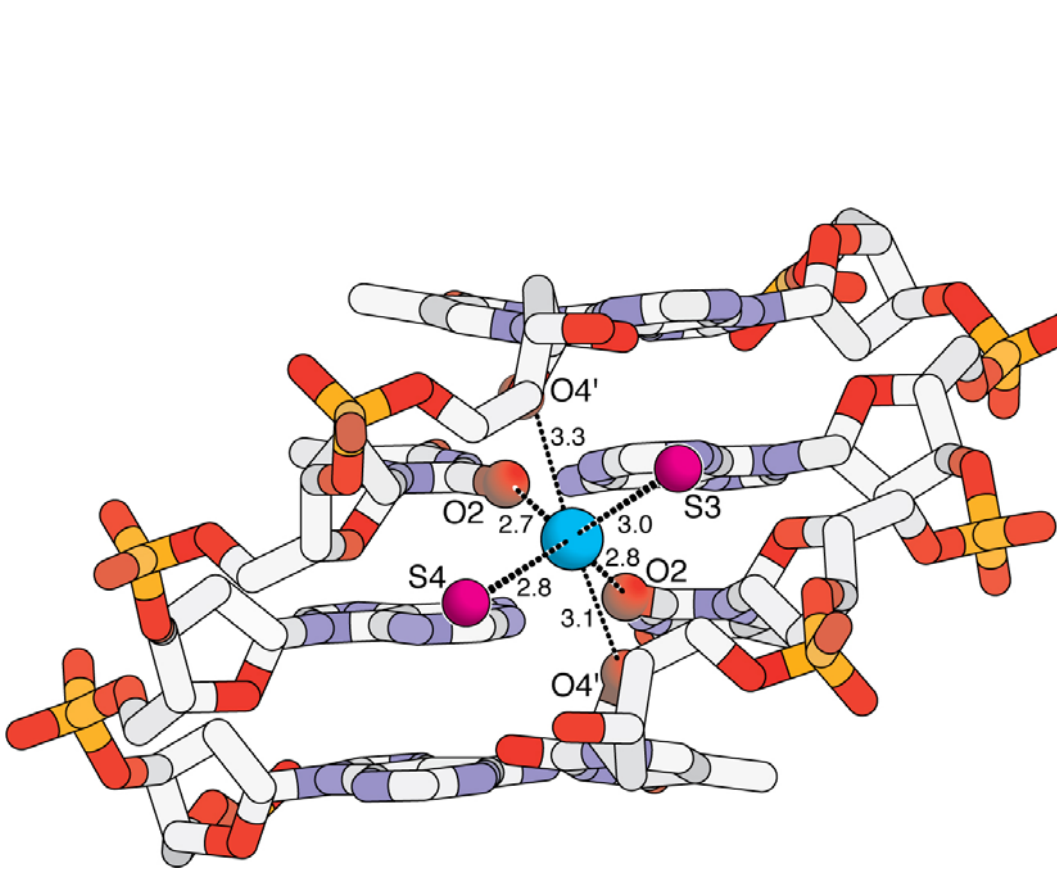
ZIG-ZAG Spine of Hydration
(B’-State)
Dickerson and Drew

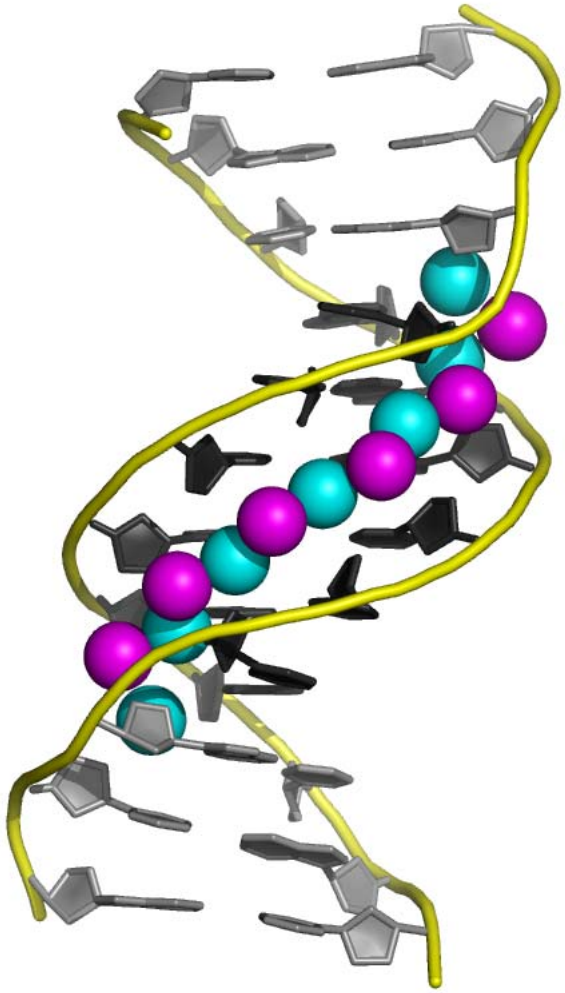


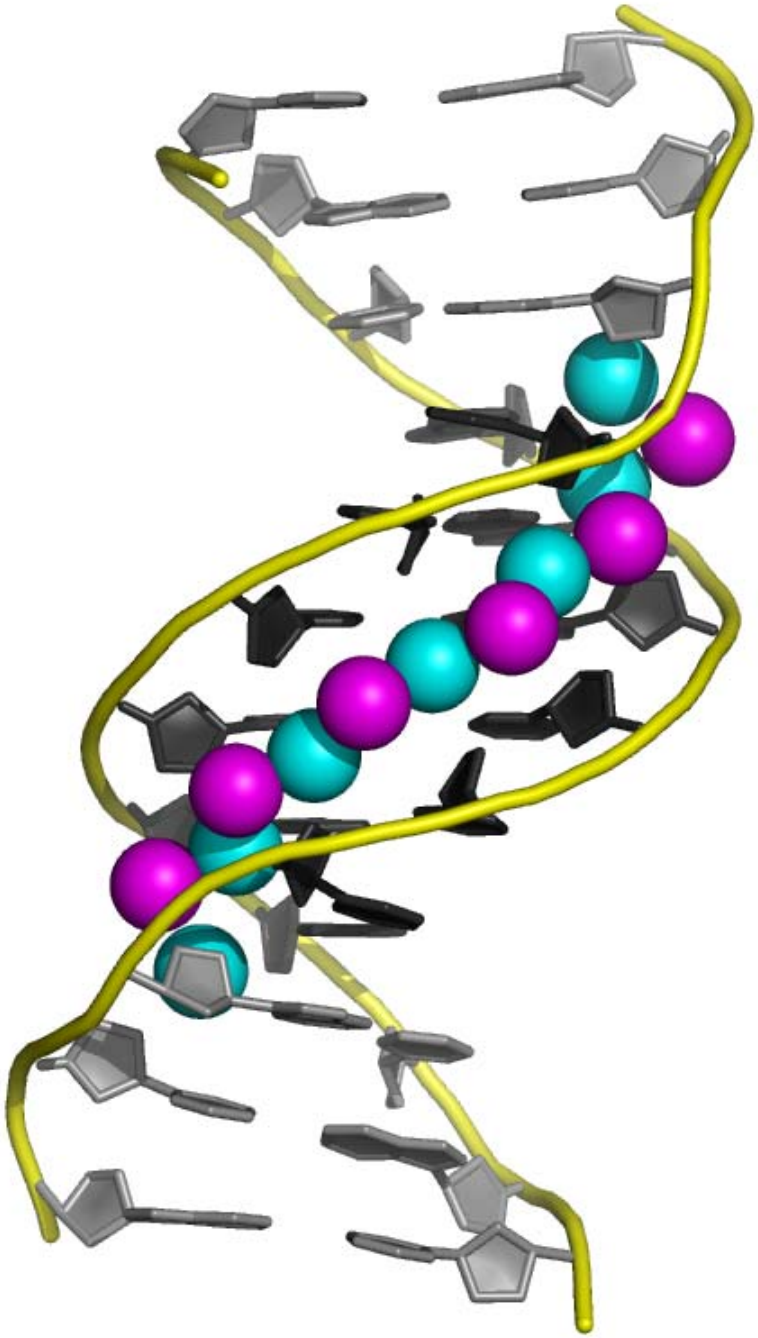
Shui, X., Sines, C., McFail-Isom, L., VanDerveer, D., & Williams, L.D., Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. Biochemistry 37, 16877-16887 (1998).

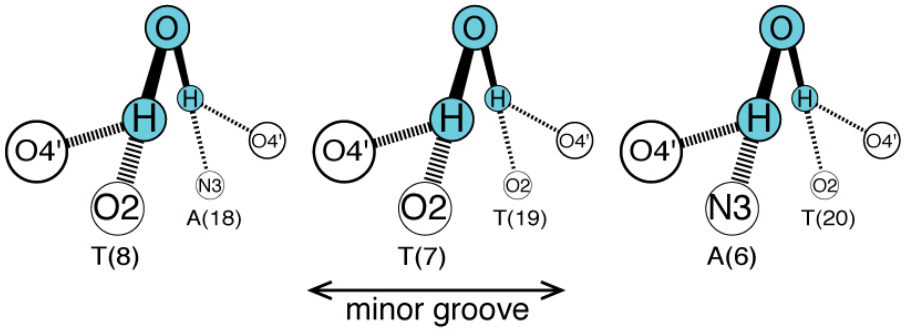


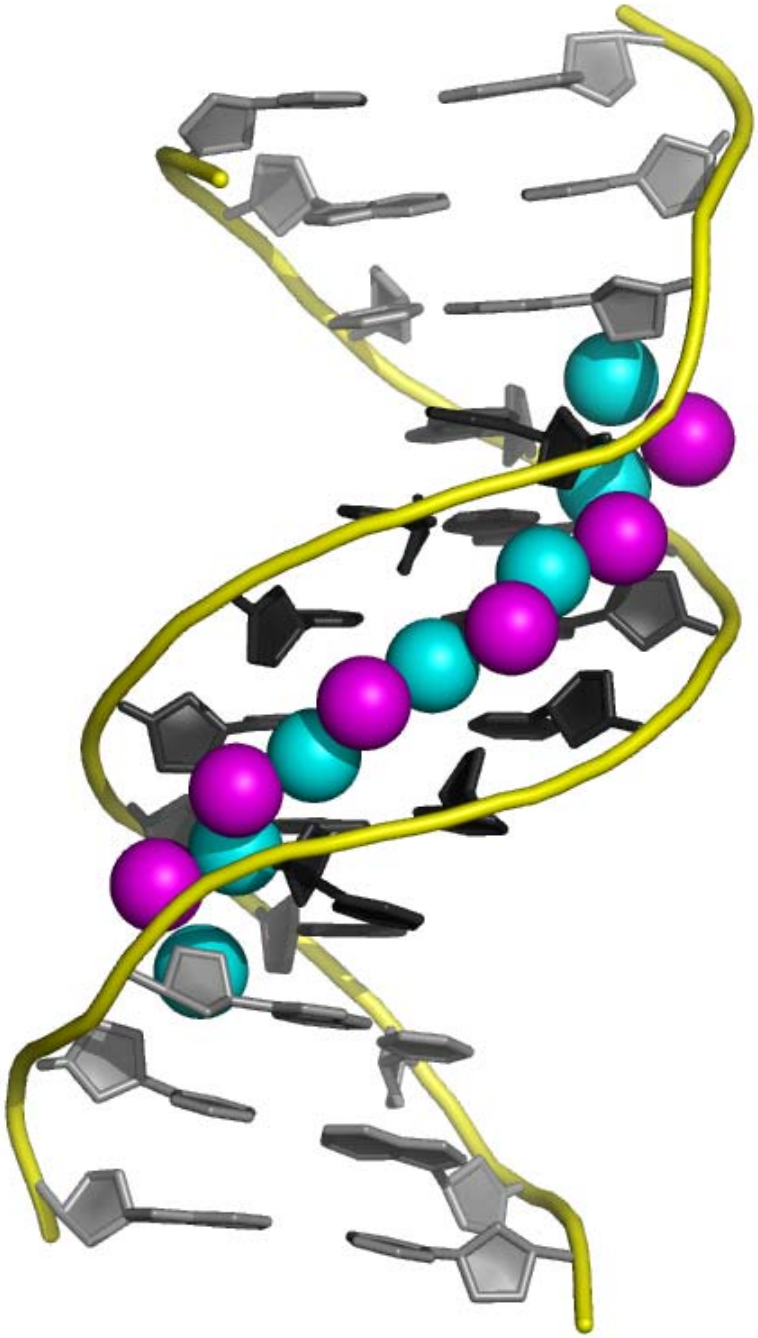


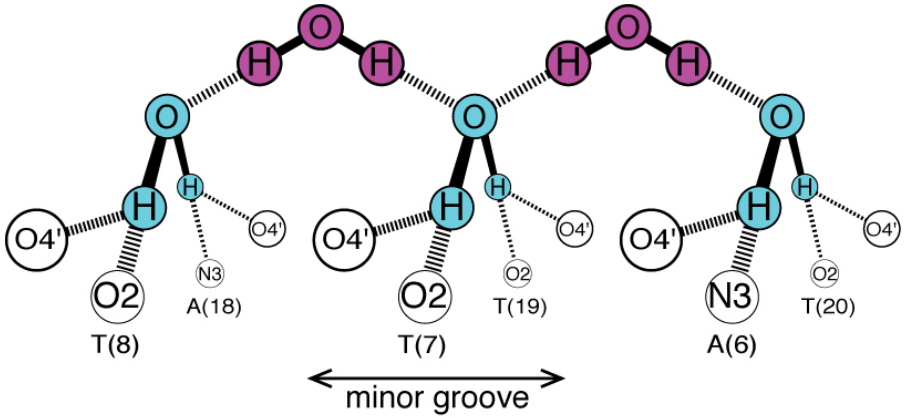

O4'
3.3
O2
2.7
3.0
S3
2.8
S4
2.8
O2
3.1
O4'

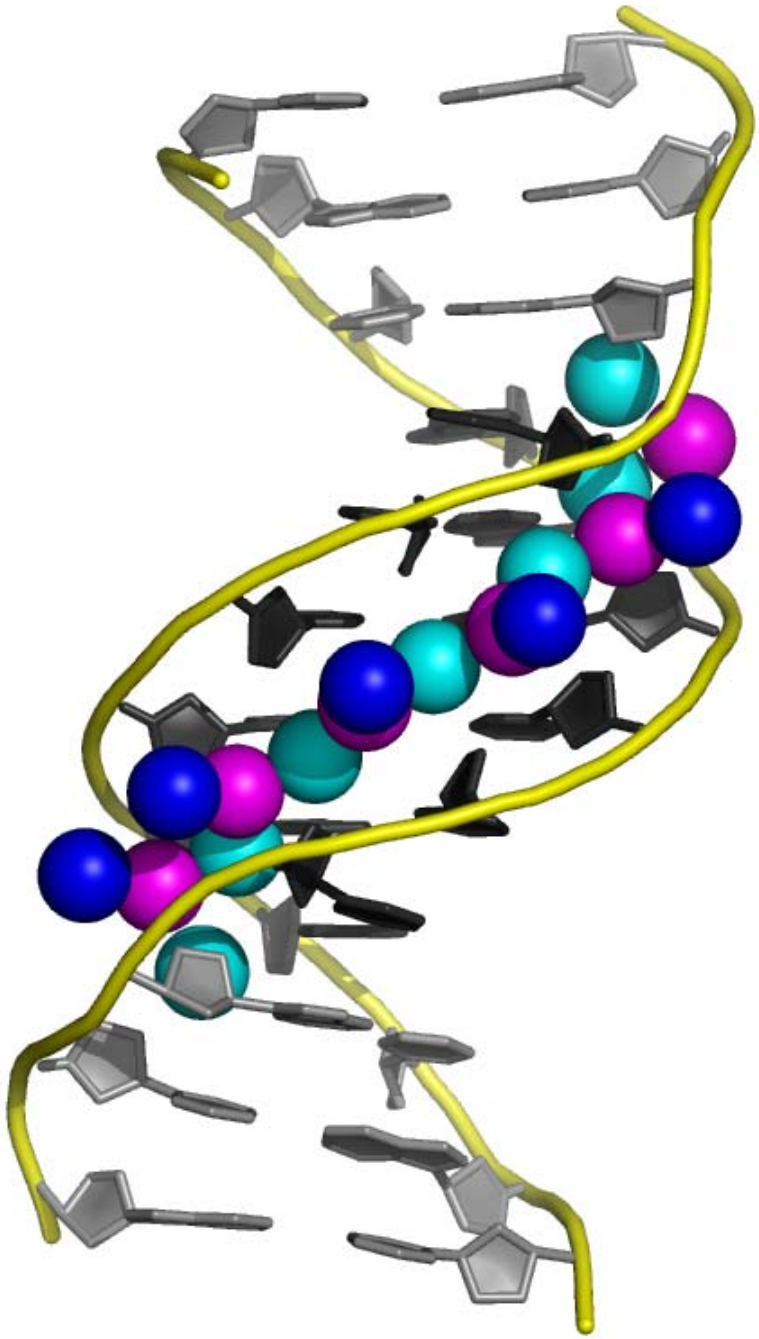


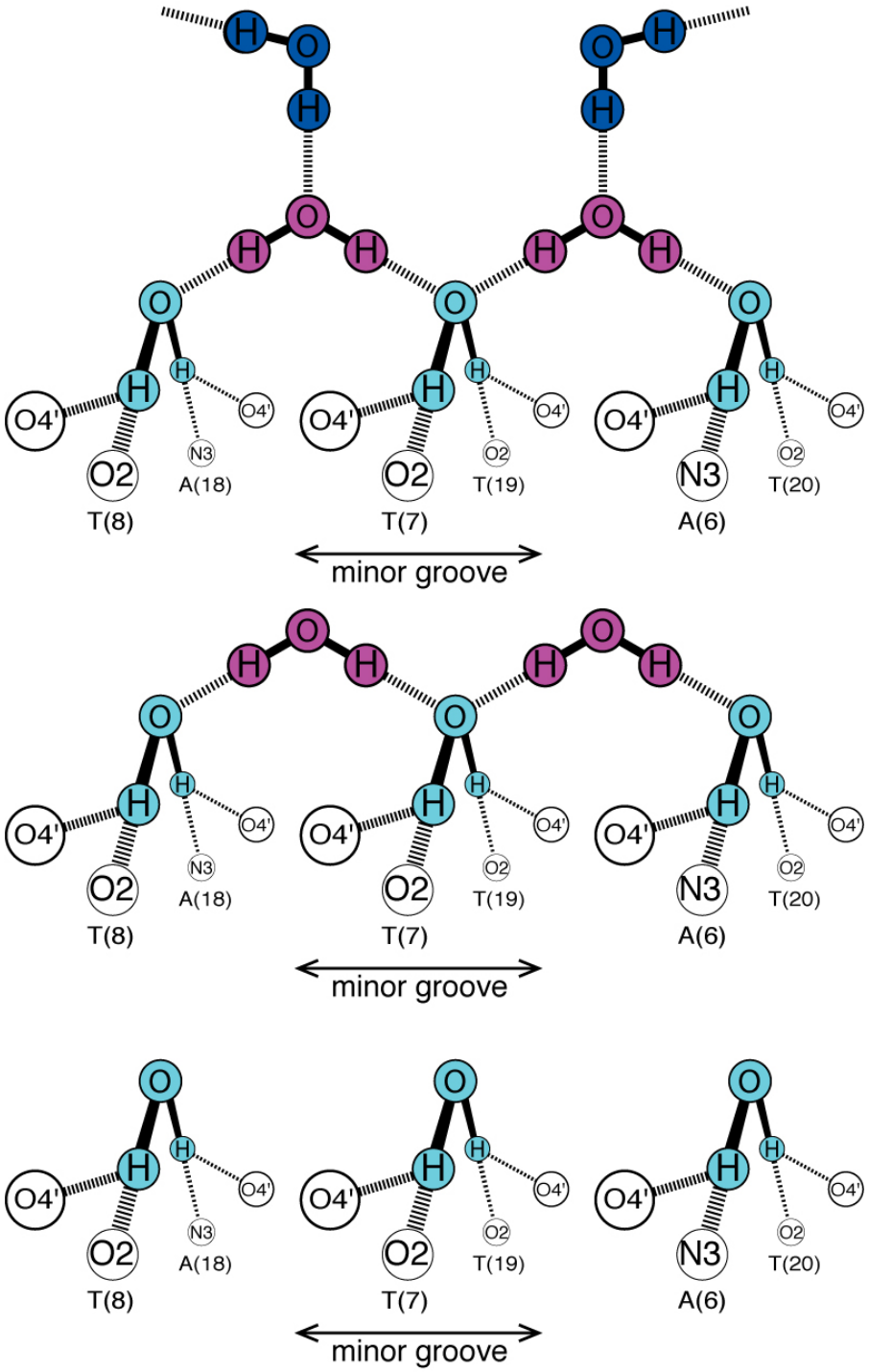


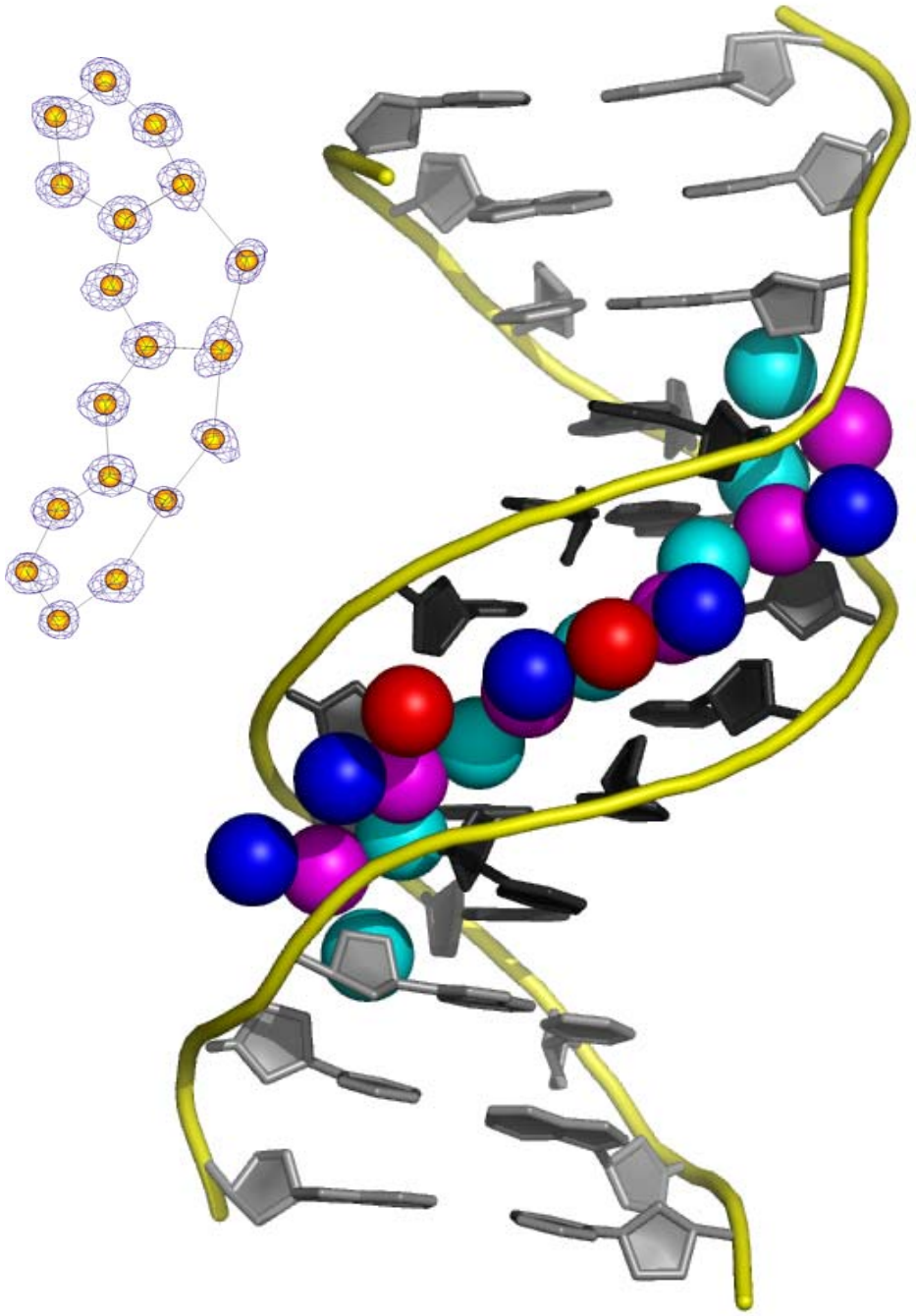

O
H
H
O4'
O4'
N3
O2
A(18)
T(8)
O
H
H
O4'
O4'
O2
O2
T(19)
T(7)
O
H
H
O4'
O4'
N3
O2
T(20)
A(6)
minor groove

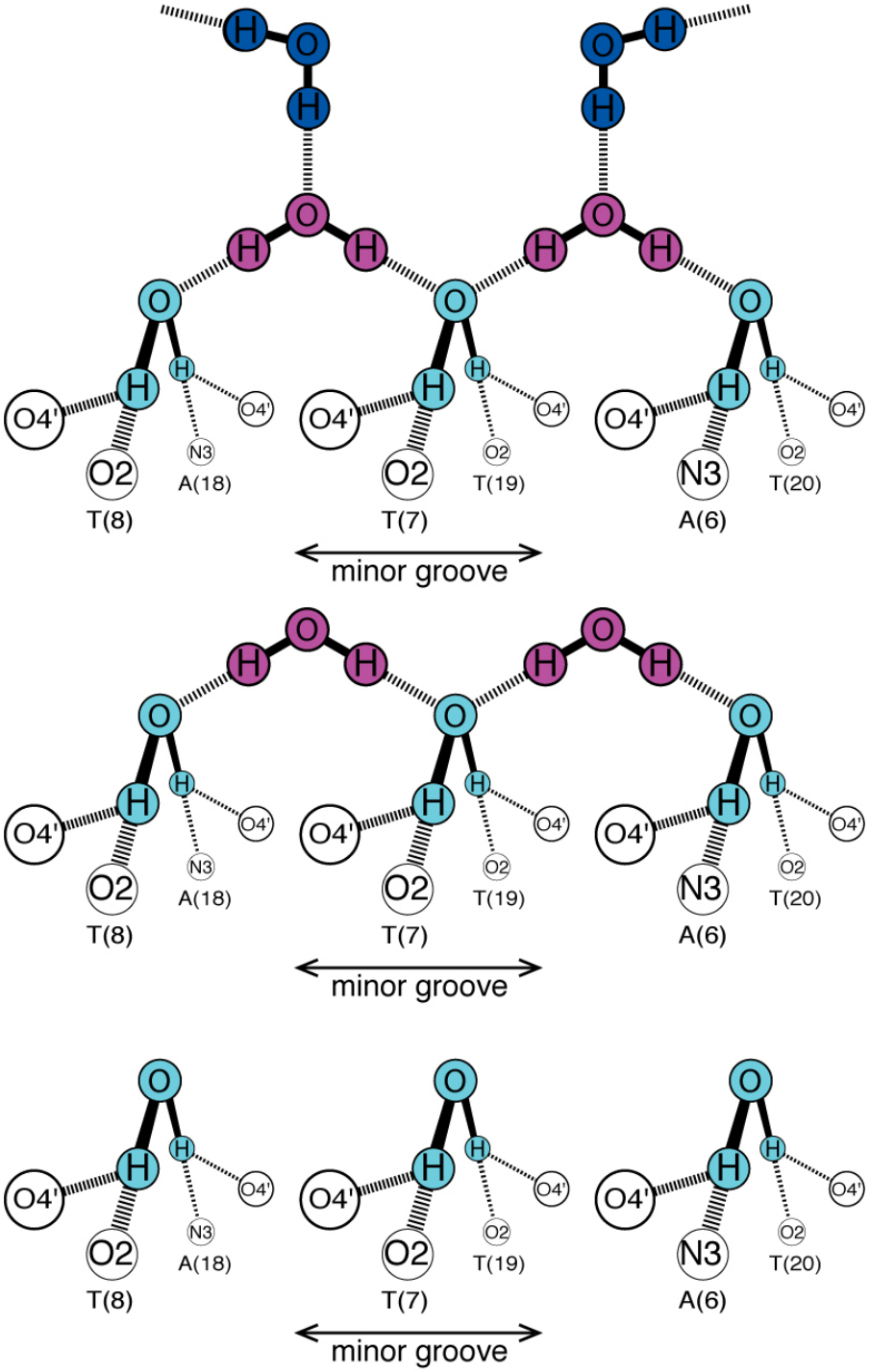



O
H
O4'
O2
N3
T(8)
A(18)
T(7)
T(19)
A(6)
T(20)
minor groove

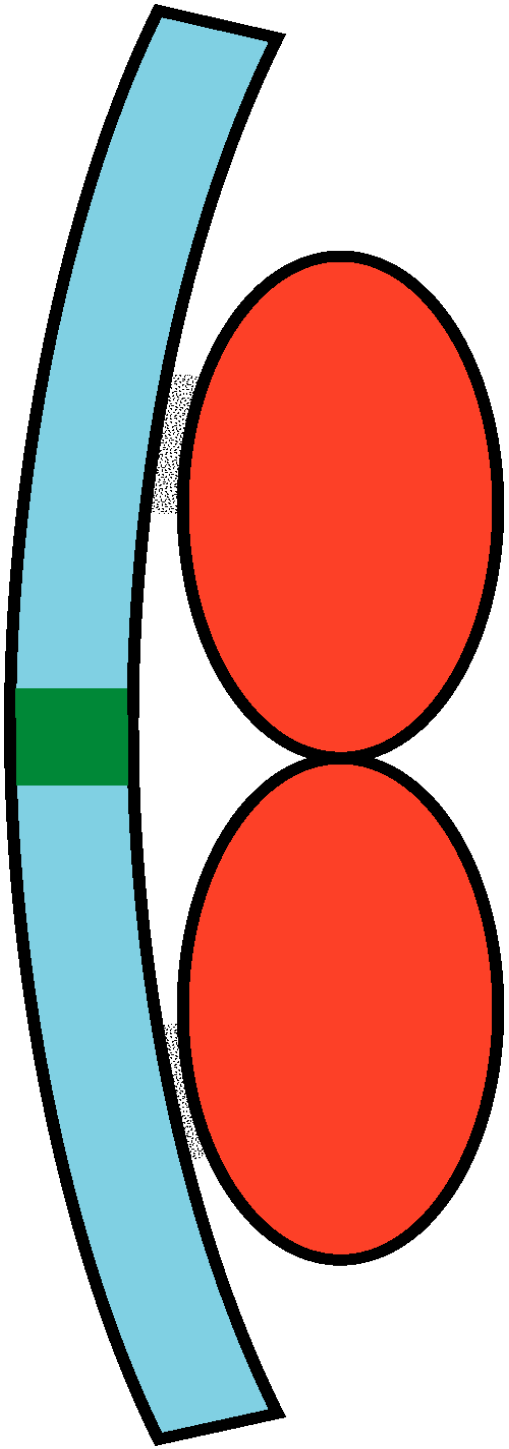



O4'
O2
N3
A(18)
T(8)
T(19)
T(7)
T(20)
A(6)
minor groove




O4'
O2
N3
A(18)
T(8)
O2
T(19)
T(7)
N3
O2
T(20)
A(6)
minor groove

DNA-Protein Recognition



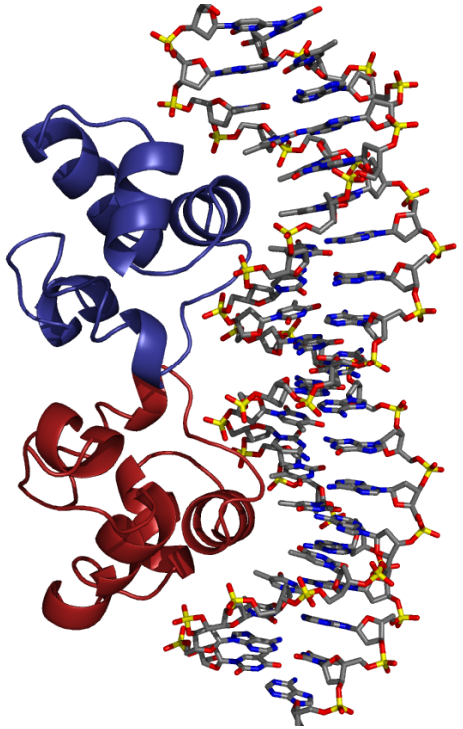
INDIRECT READOUT

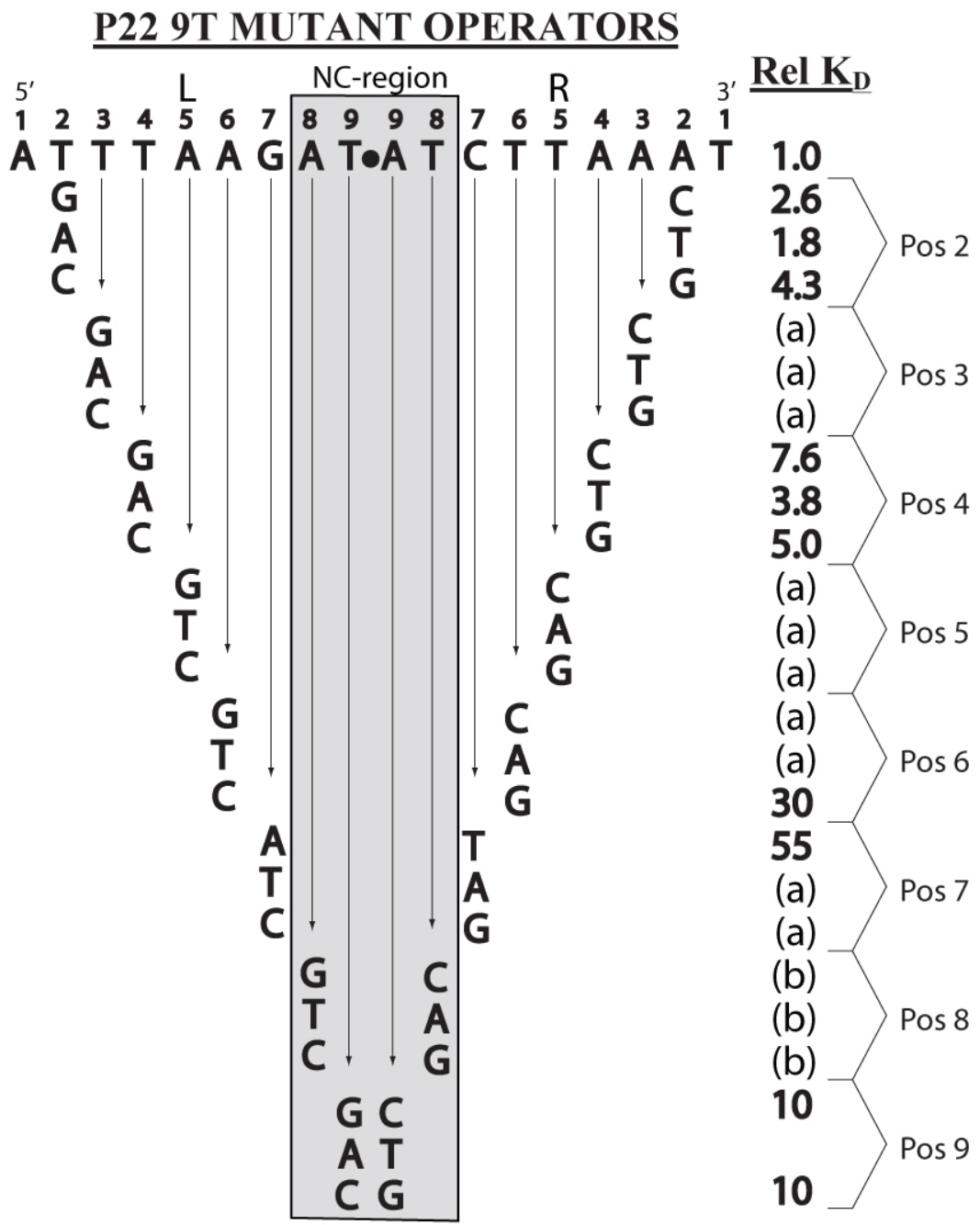
Effect of Non-Contacted Bases on the Affinity of 434 Operator for 434 Repressor and Cro

Gerald Koudelka, Stephen Harrison and Mark Ptashne

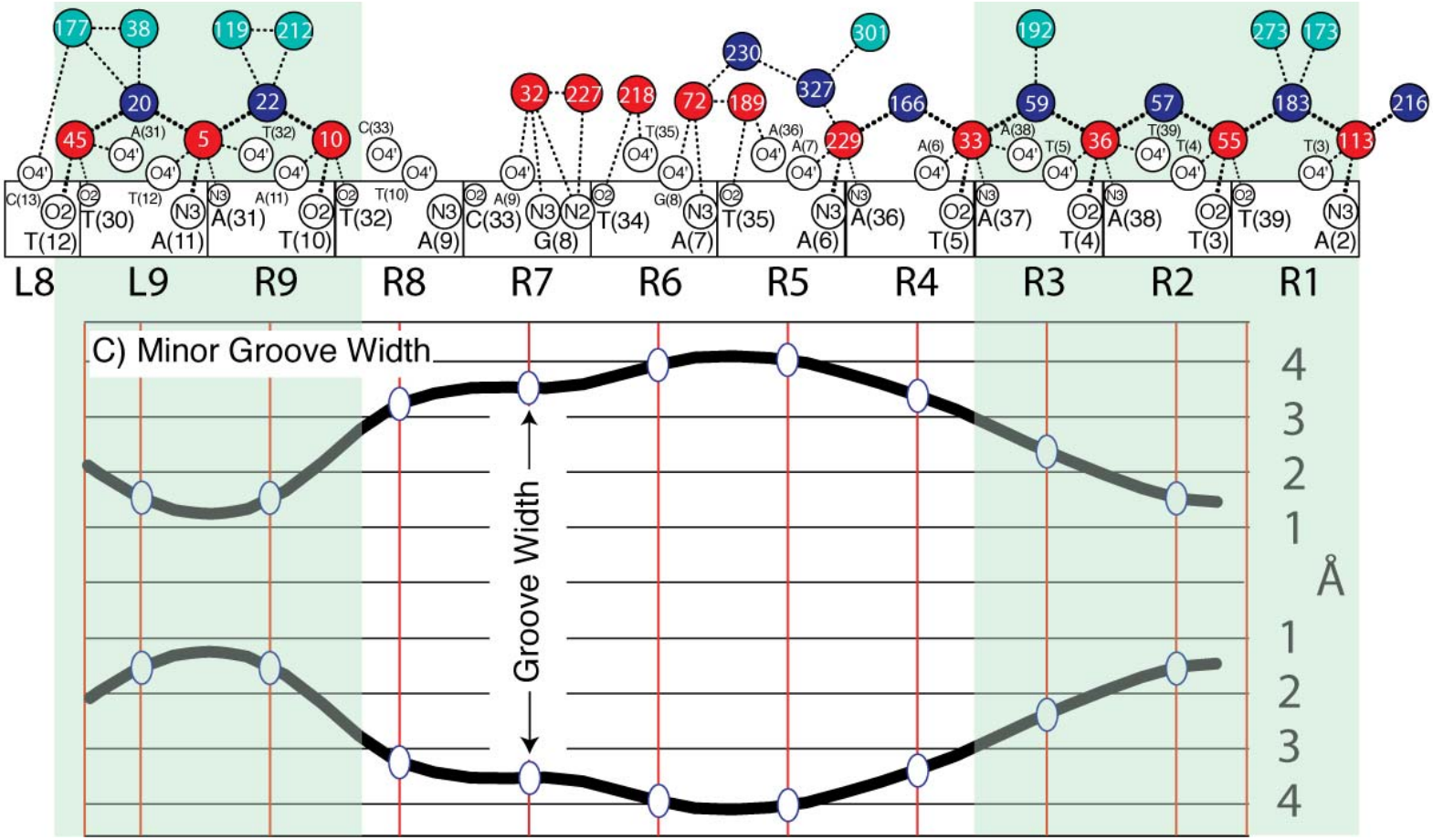
Nature (1987) 326 pp 886-888

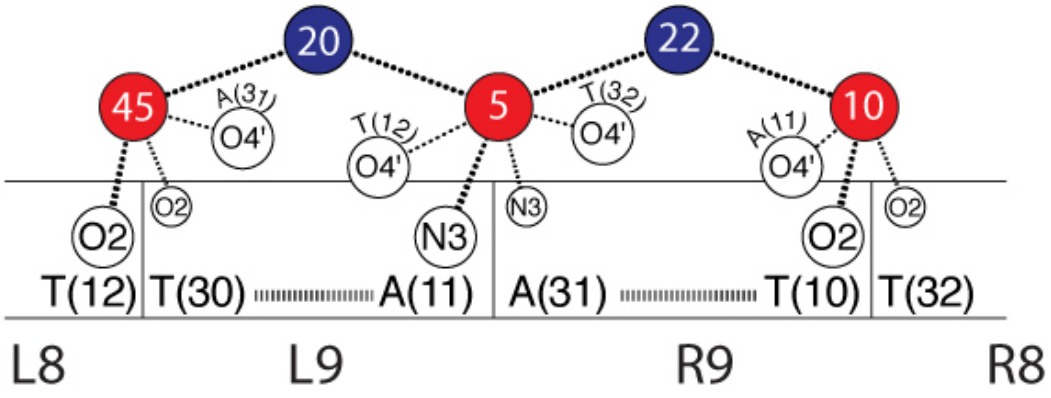
Changing the central (non-contacted) base pairs from TAAT to GCCG decreases the affinity 50-fold.

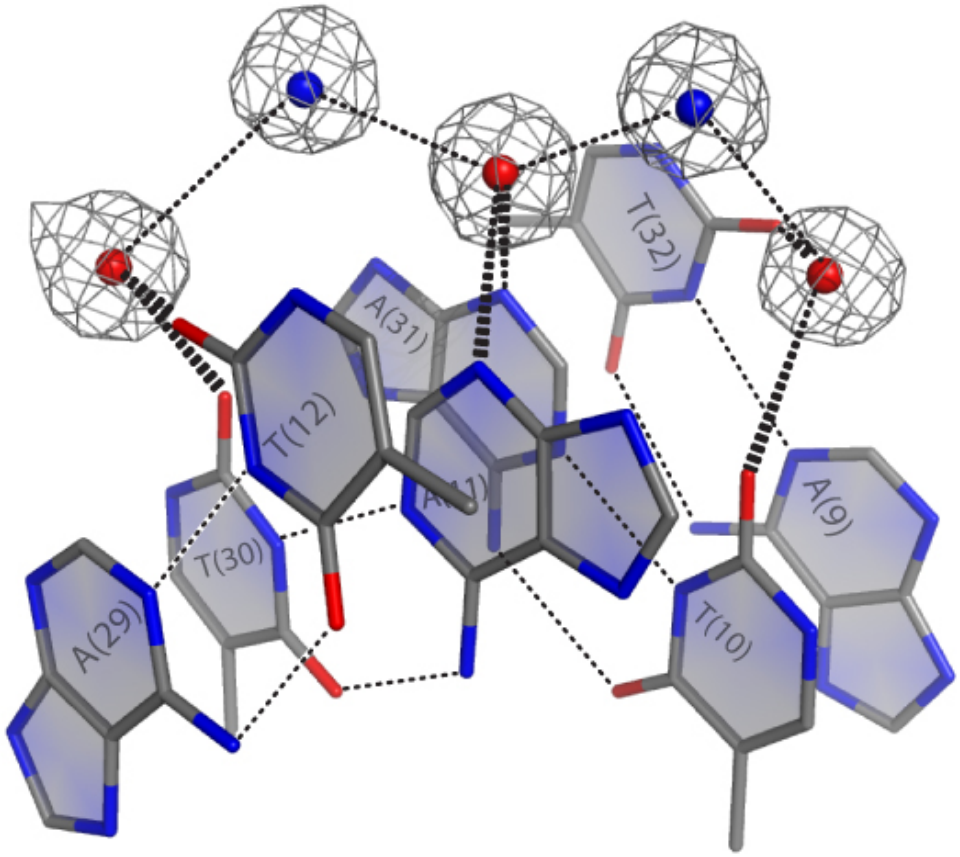


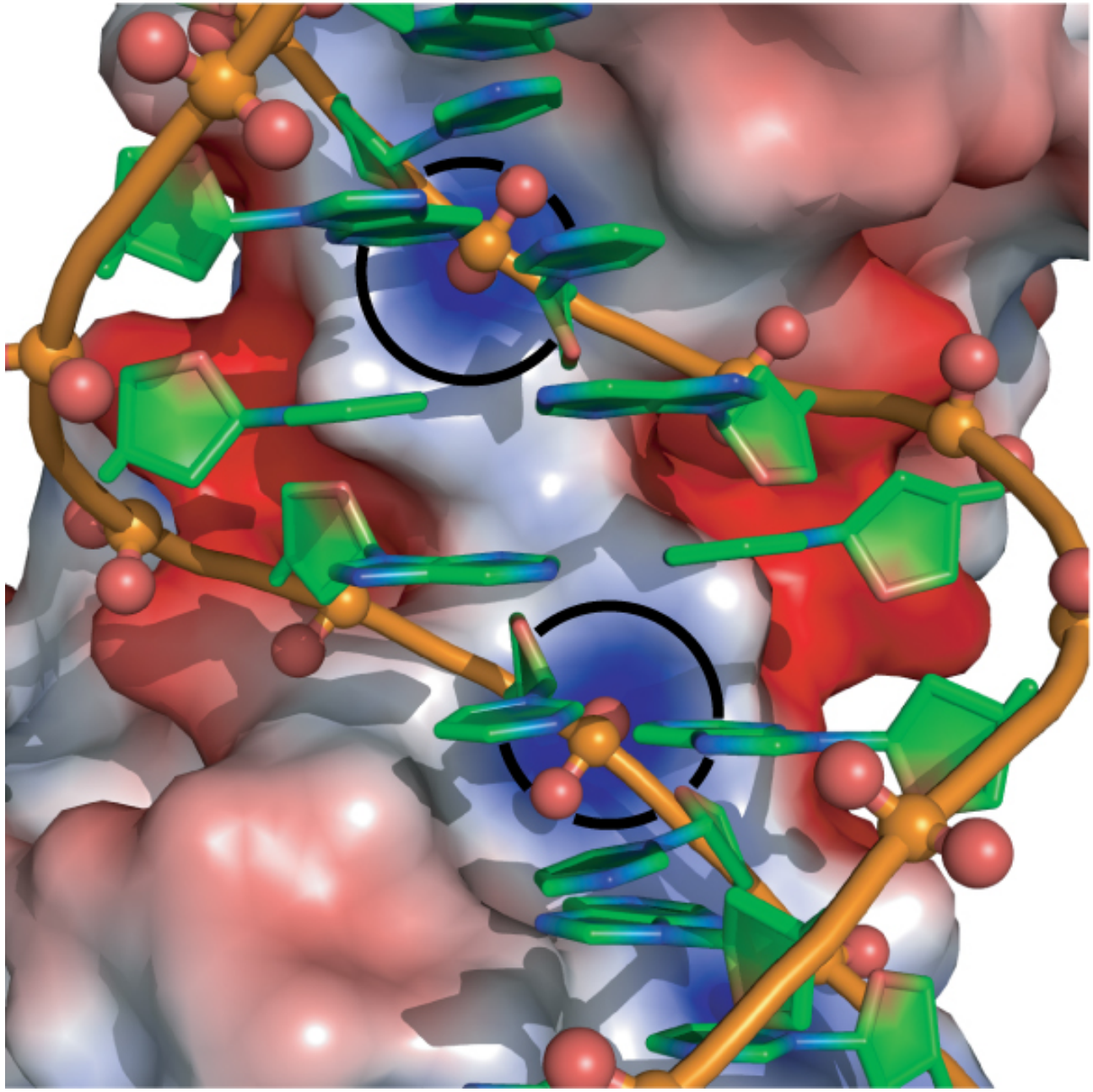

P22 9T MUTANT OPERATORS
NC-region
L
R
5′
3′
Rel K_D
1.0
2.6
1.8
4.3
(a)
(a)
(a)
7.6
3.8
5.0
(a)
(a)
(a)
(a)
(a)
30
55
(a)
(a)
(b)
(b)
(b)
10
10
Pos 2
Pos 3
Pos 4
Pos 5
Pos 6
Pos 7
Pos 8
Pos 9

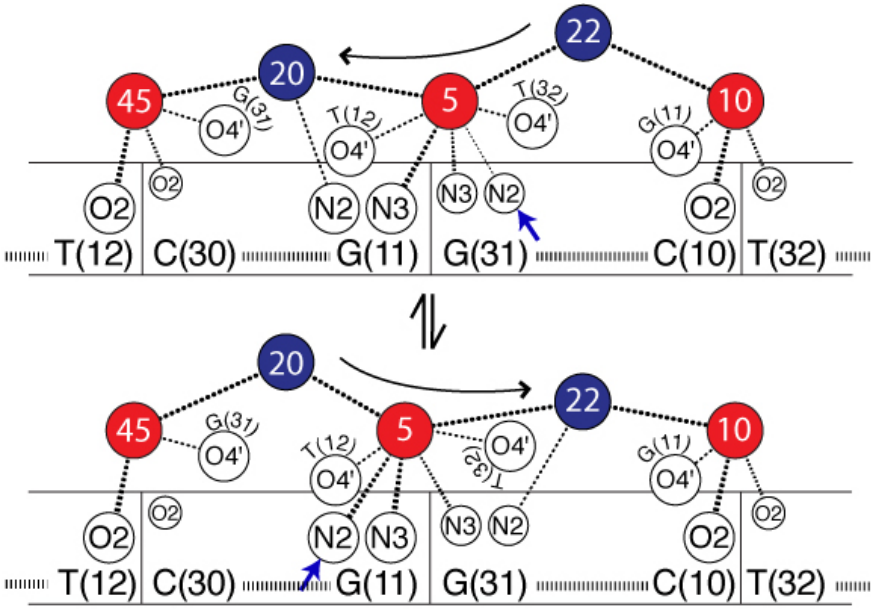
The protein (somehow) forces non-A-tract DNA into the B’ State

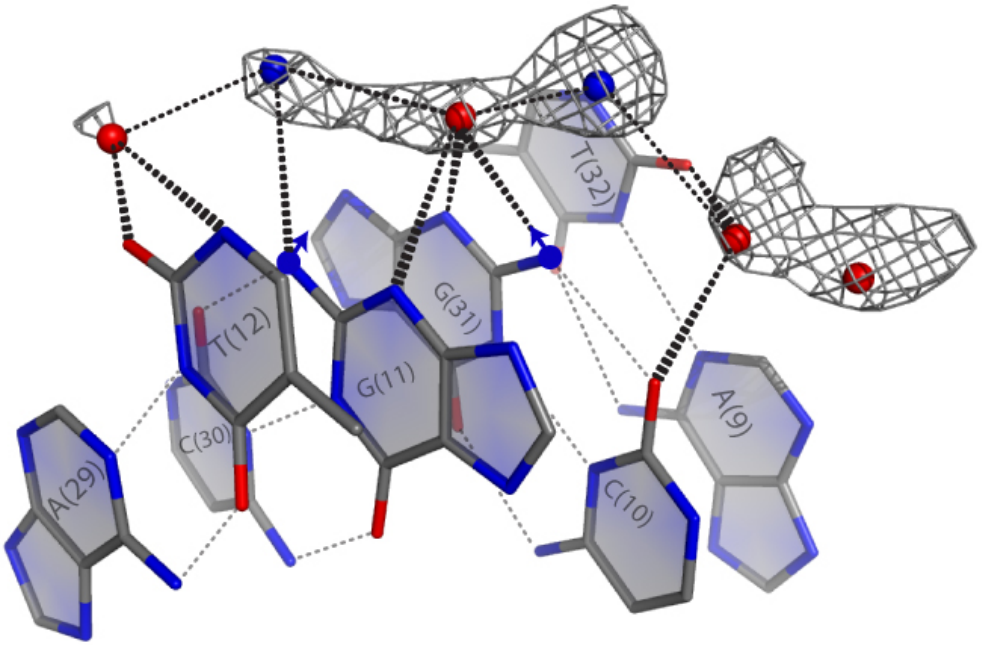



20
22
45
5
10
A(31)
O4'
T(12)
O4'
T(32)
O4'
A(11)
O4'
O2
O2
N3
N3
O2
O2
T(12)
T(30)
A(11)
A(31)
T(10)
T(32)
L8
L9
R9
R8

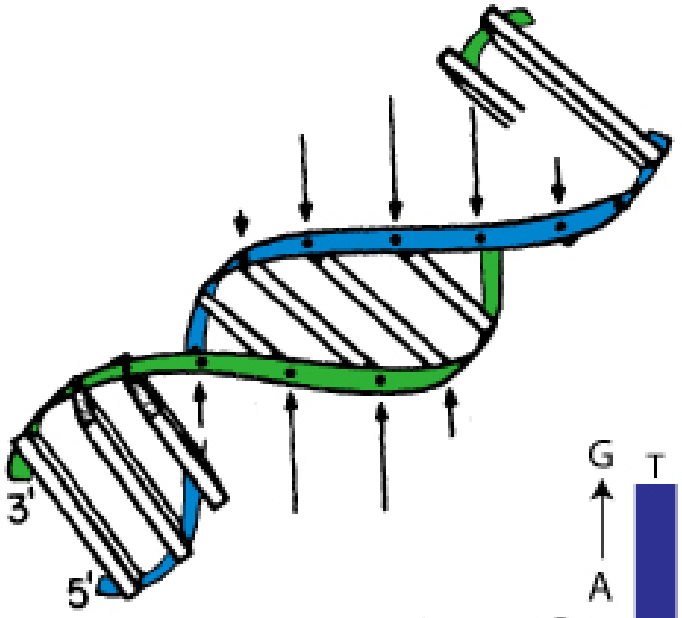

T(32)
A(31)
T(12)
T(30)
A(29)
A(9)
T(10)

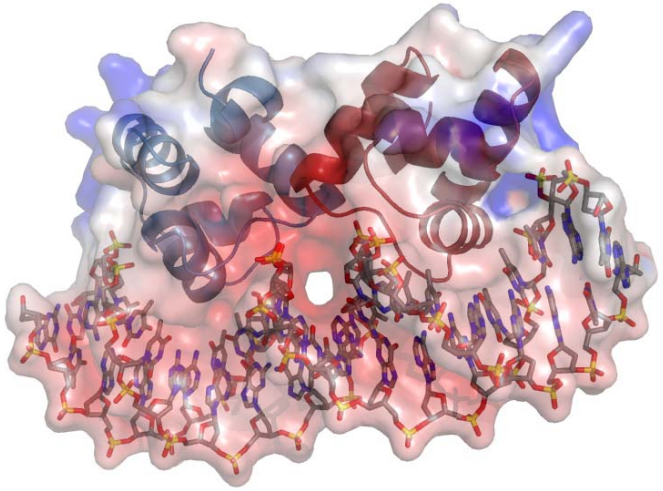


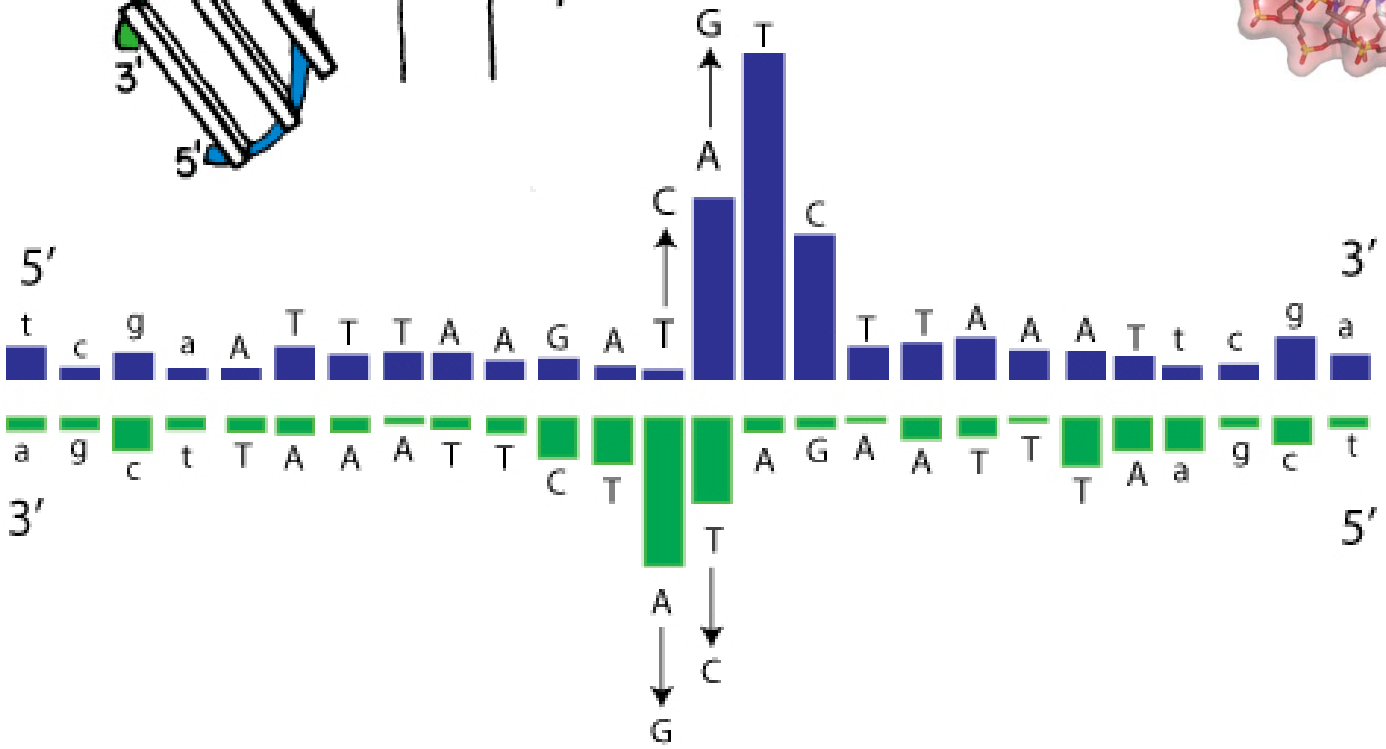

22
20
45
5
10
G(31)
T(12)
T(32)
G(11)
O4'
O2
N2
N3
T(12) C(30) G(11) G(31) C(10) T(32)
20
22
45
5
10
T(12) C(30) G(11) G(31) C(10) T(32)


T(32)
G(31)
T(12)
G(11)
C(30)
A(29)
A(9)
C(10)

A river runs through it…



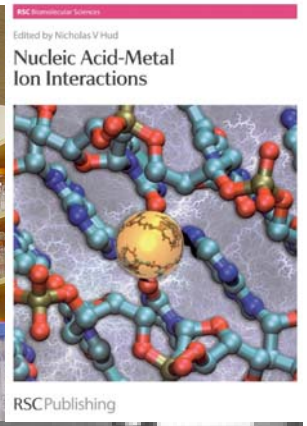


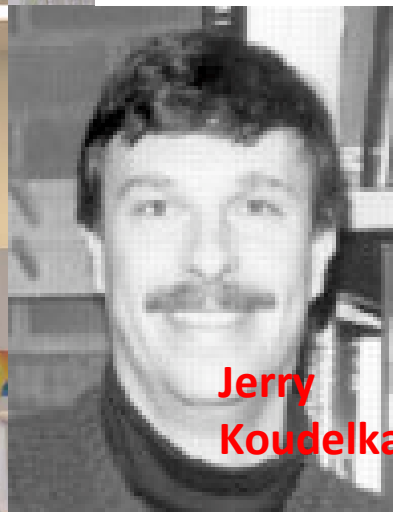


- Current Model

  A DNA sequence element can be ‘read’ by its free energy of transition from the B state to the B’ state.

- NSF
- NASA (Astrobiology Institute)





The following are omitted from the talk

Proc. Nat. Acad. Sci. USA
Vol. 73, No. 3, pp. 804–808, March 1976
Biophysics

Sequence-specific recognition of double helical nucleic acids by proteins

(base pairs/hydrogen bonding/recognition fidelity/ion binding)

NADRIAN C. SEEMAN, JOHN M. ROSENBERG*, AND ALEXANDER RICH

Department of Biology, Massachusetts Institute of Technology, Cambridge, Mass. 02139

Contributed by Alexander Rich, December 10, 1975

ABSTRACT **The base pairs in double helical nucleic acids have been compared to see how they can be recognized by proteins. We conclude that a single hydrogen bond is inadequate for uniquely identifying any particular base pair, as this leads to numerous degeneracies. However, using two hydrogen bonds, fidelity of base pair recognition may be achieved. We propose specific amino-acid side chain interactions involving two hydrogen bonds as a component of the recognition system for base pairs. In the major groove we suggest that asparagine or glutamine binds to adenine of the base pair, or arginine binds to guanine. In the minor groove, we suggest an interaction between asparagine or glutamine with guanine of the base pair. We also discuss the role that ions and other amino-acid side chains may play in recognition interactions.**

One of the important unsolved problems in molecular biology concerns the detailed molecular mechanism in the recognition of specific sequences of double helical nucleic acids by proteins. The problem which we are considering is the unique identification in a double helix of each of the four possible base pairs [A-U(T); U(T)-A; G-C; C-G] when compared with each of the other three while they are still in the double helical conformation.

There are only a few ways known by which polypeptides interact specifically with other molecules. These include hydrophobic interactions such as the base stacking interactions, and electrostatic interactions, of which the most important are hydrogen bonds. Because of the high specificity and directional character of hydrogen bonds, we believe they will play a major role in the recognition process. Here we indicate the similarities and differences of the four distinct Watson–Crick base pairs which could be probed by hydrogen bonding groups of proteins for specific recognition. It should be noted that base pair recognition is similar for both double helical DNA and RNA. The result of this analysis leads us to believe that a single hydrogen bond is unable to identify must make in order to distinguish between individual base pairs. We will assume throughout that a protein can use the double helical backbone of the nucleic acid in order to establish a frame of reference from which to probe the base pair. A pair of ribose residues in an RNA double helix is shown with two different types of base pairs superimposed. The major groove of the double helix is at the top of each figure while the minor groove is at the bottom. Fig. 1a shows the comparison of the A-U and the U-A base pair. A methyl group would be attached to the 5 position of uracil to illustrate thymine. Fig. 1b superimposes the G-C and C-G pairs, Fig. 1c the A-U and C-G pairs, while Fig. 1d shows the A-U and G-C pairs. The pair U-A superimposed on G-C is also represented in Fig. 1c by simply rotating the pair about a vertical axis, while U-A superimposed on C-G is similarly represented in reverse order in Fig. 1d.

Potential sites for discrimination are labeled in the figure, where W stands for possible recognition sites in the major or wide groove and S for sites in the minor or small groove. These sites have been selected in the figure if the atom or atomic grouping is accessible for hydrogen bonding to the double helix when the probe approaches it. The position arrows point to the heavier atoms. In the amino group the position of the nitrogen atom is taken rather than the two hydrogens which are attached to it, even though recognition at this point must occur through interactions involving the hydrogen atoms. Because the molecule is organized as an antiparallel double helix, the ribose sugars in Fig. 1 are related to each other by a vertical 2-fold axis. The primed recognition sites are related to the unprimed ones of the same number by this same 2-fold axis. These recognition sites are in the same place in all four parts of the diagram. The six possible pairs of base pairs have been illustrated in groups of two for ease of comparison.

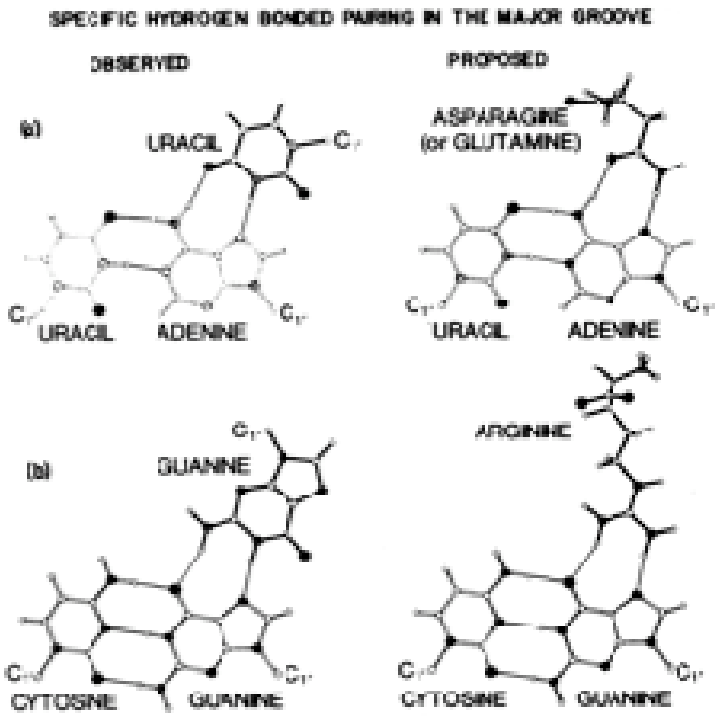


FIG. 2. Interactions between base pairs and other bases (observed) or amino-acid side chains (proposed). (a) The uracil binding to the U-A pair is seen in polynucleotides (11) as well as in single crystals of adenine and uracil derivatives (18). The conformation of the asparagine is taken from a neutron diffraction study (19). (b) The guanine binding to the C-G pair is seen in yeast phenylalanine tRNA (12, 13). The arginine conformation is taken from a neutron crystallographic analysis (20). Oxygen atoms have diagonal shading while nitrogen atoms are stippled.

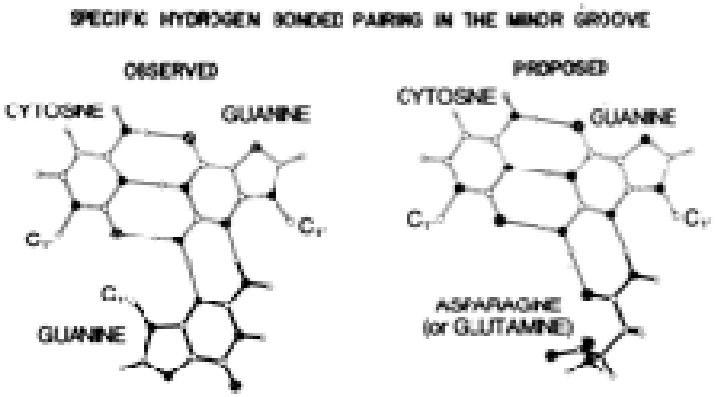


FIG. 3. Interactions between C-G base pairs and guanine (observed) or asparagine (proposed). The guanine interaction with the C-G pair is seen in the structure of the 9-ethyl guanine-1-methyl cytosine complex (16).

drogen bonds as the basis of specificity in the interaction (11). These analogies are useful in suggesting a system in which amino-acid side chains form similar specific pairs of hydrogen bonds with base pairs in the double helix. Examples are shown for the major groove in Fig. 2 and for the minor groove in Fig. 3. Fig. 2a shows the type of hydrogen bonding in which the uracil residue interacts specifically with the A-U pair in the triple-stranded complex (11). Sites W3′ and W1′ of adenine are used to form two specific hydrogen bonds with the C-2 carbonyl and N-3 of uracil. Site W2′ of adenine is used to form a hydrogen bond with the amide of asparagine. In a similar way in Fig. 2b an additional carboxyl group might further stabilize the interaction of arginine with the C-G base pair by having its oxygen atoms act as acceptors for the H atoms on cytosine N4 and the amino group of arginine.

Fig. 3 shows a way in which sites S1′ and S2 can be used for discriminating the C-G base pair in the minor groove. On the left we see how guanine interacts with the minor groove of the G-C pair as seen in the crystal structure of 9-ethyl guanine and 1-methyl cytosine (16). In this crystal structure the amino group N2 of guanine acts as a donor to site S1′ while N3 acts as an acceptor from site S2. An analog similar to this may be seen in the proposed interaction in which the amide of asparagine or glutamine forms a pair of hydrogen bonds with the same two sites of the guanine in the C-G base pair. An amino group might further stabilize this interaction in a manner analogous to that described for Fig. 2a.

There is discrimination in the specificity of the interaction of actinomycin D with double helical G-C pairs in DNA. In the structure of the complex between actinomycin D and deoxyguanosine, the specificity is determined by two hydrogen bonds from the peptide backbone of the antibiotic to guanine N3 and N2 (17). Even though these hydrogen bonds