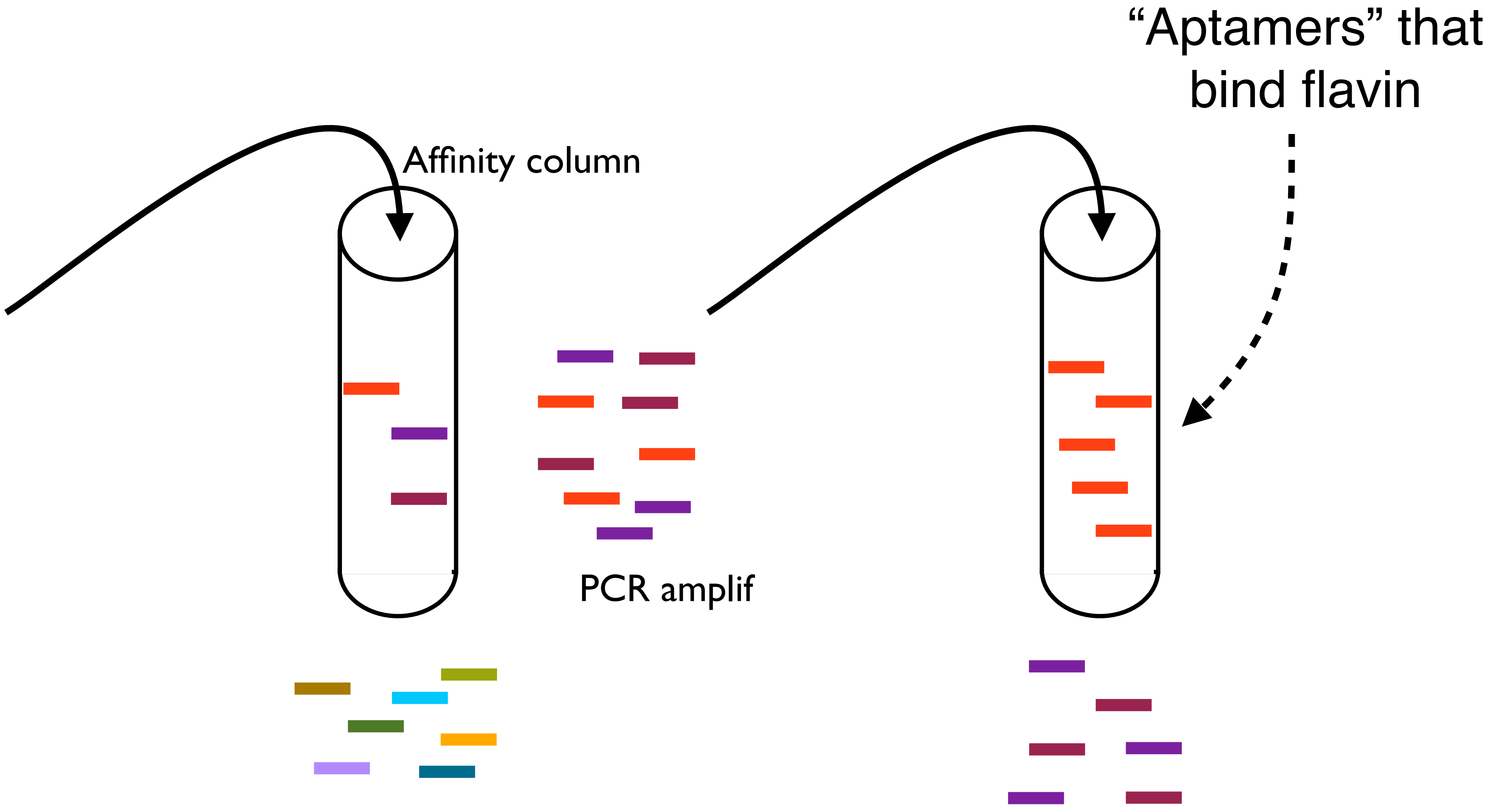
Nucleic Acid Structure & Biology

Goal: a ssDNA that binds flavin

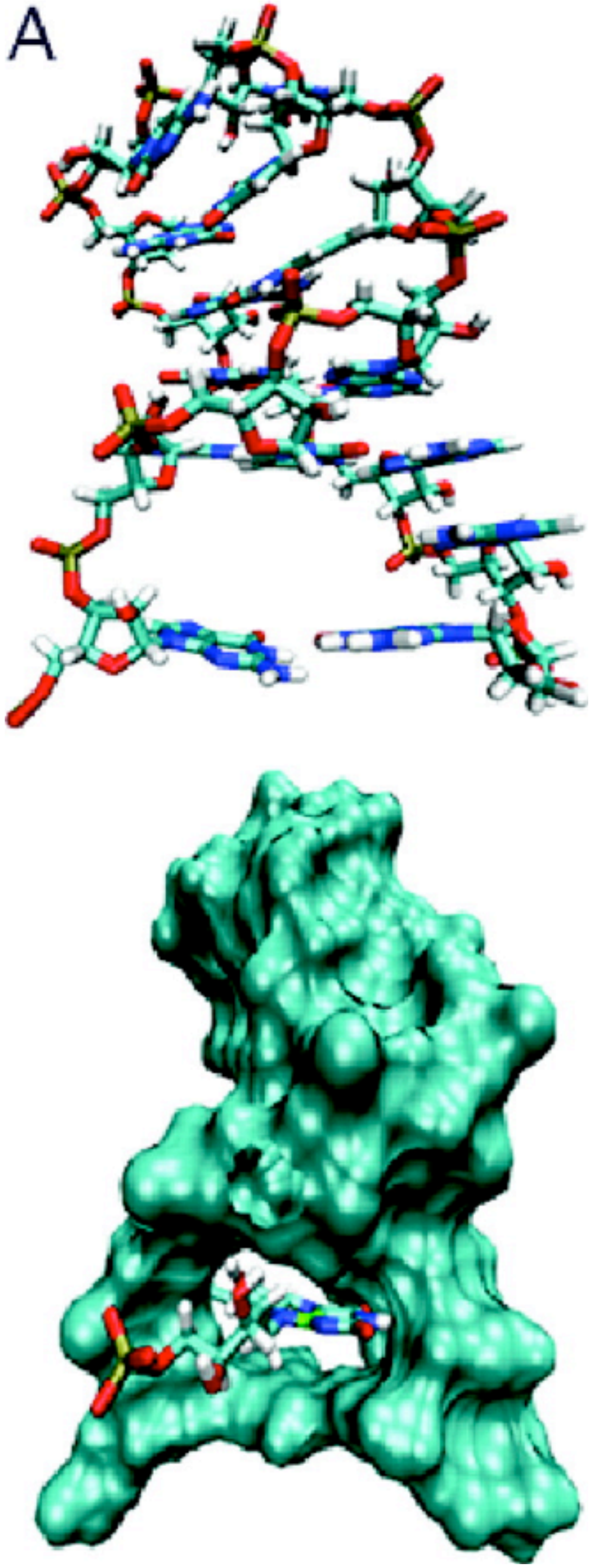
Use an affinity column with covalently tethered flavin

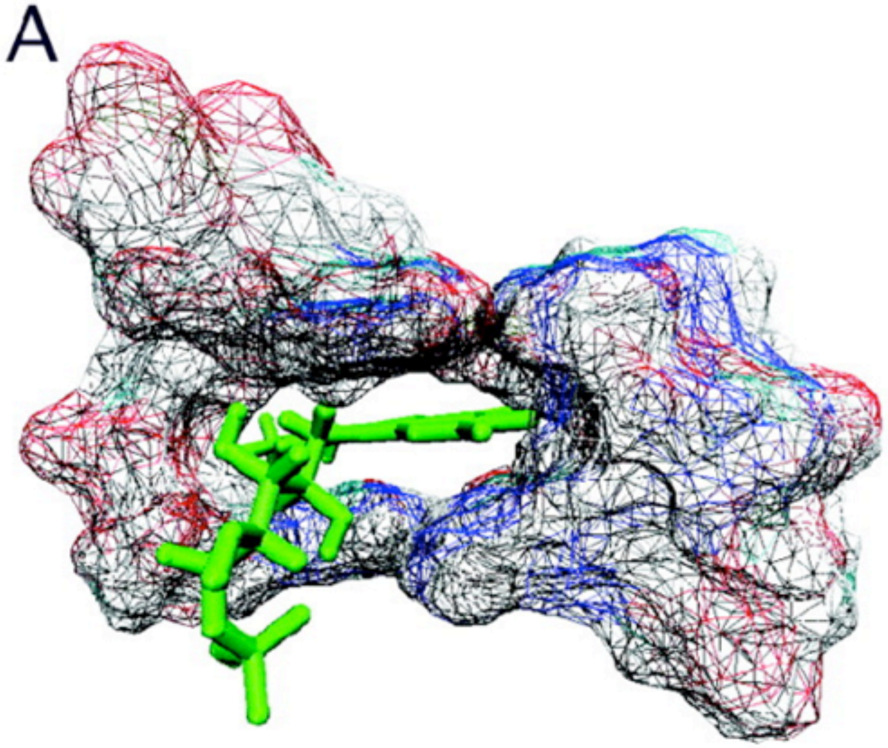


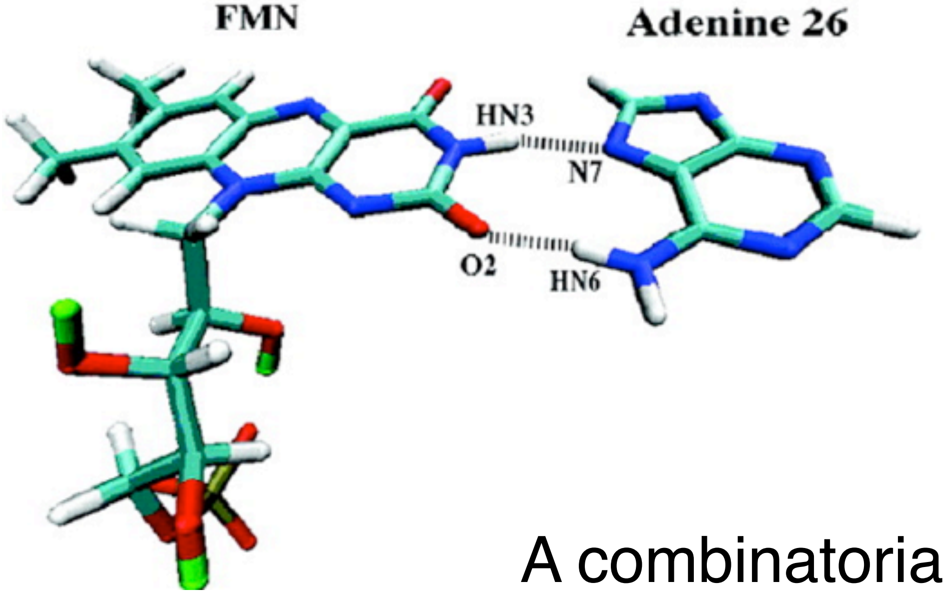
POSITIVE selection

For increased selectivity:
also include a *negative* selection step

Flavin aptamer



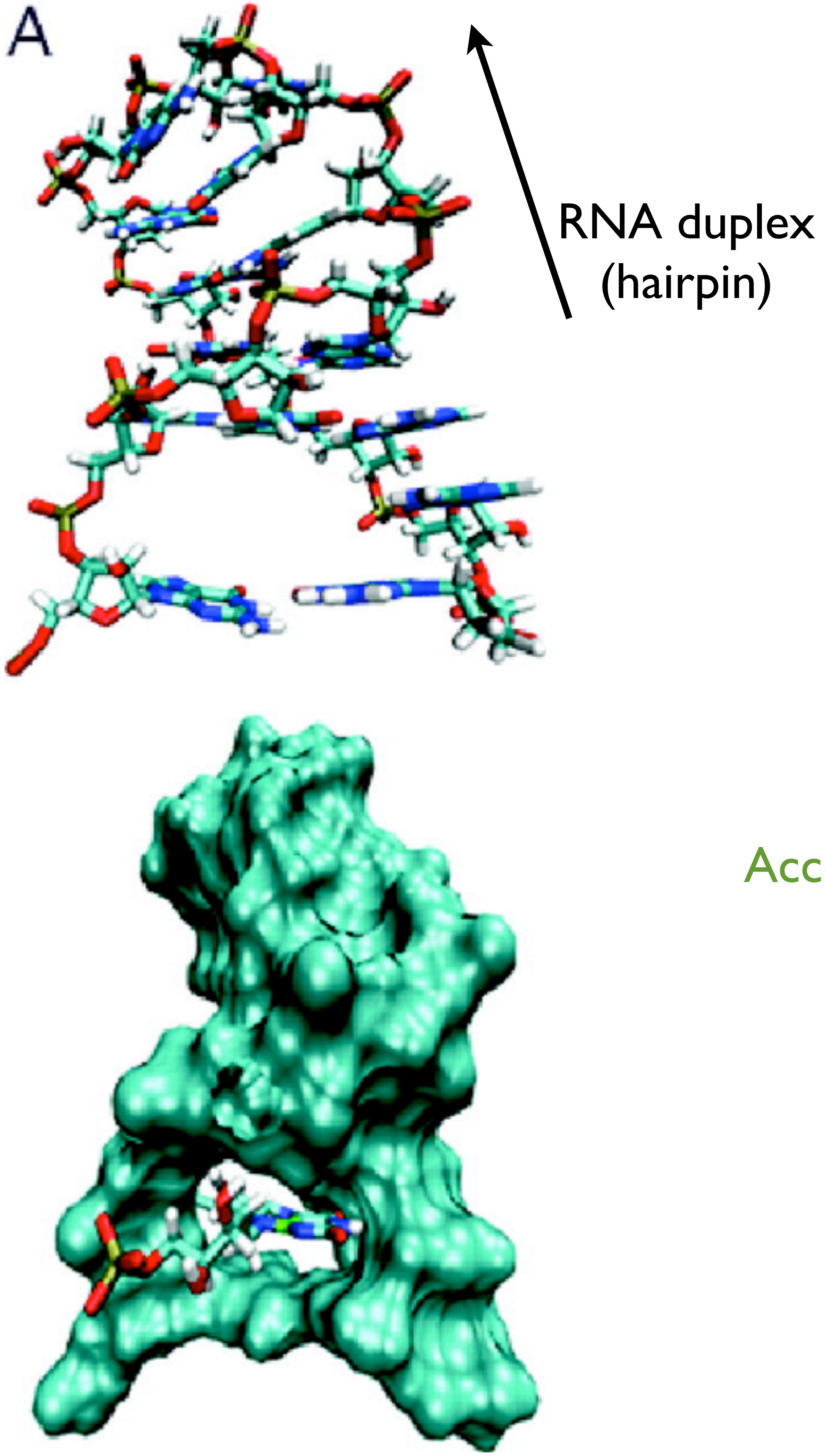


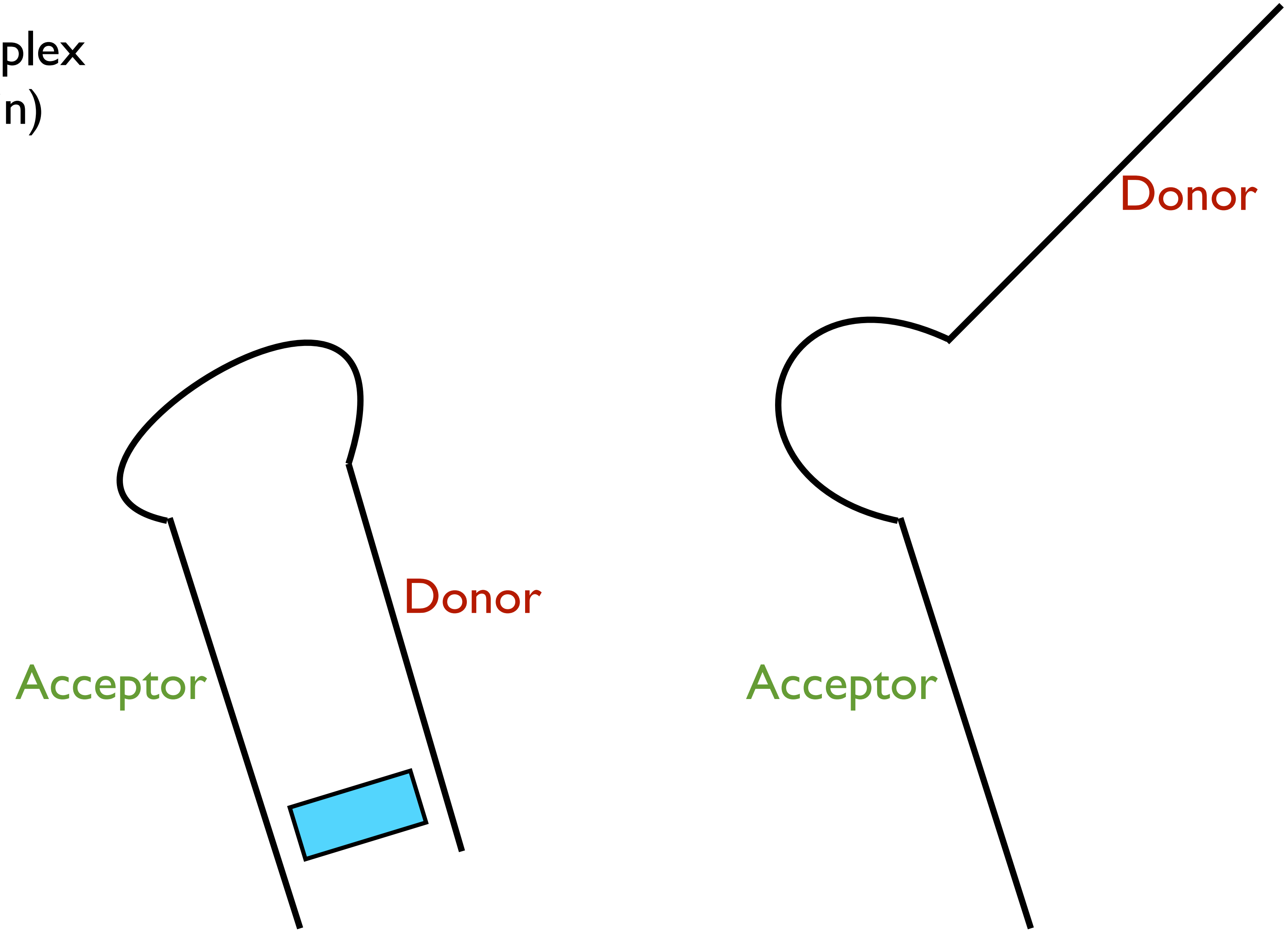


A combinatorial success!

Anderson, P. C. & Mecozzi, S. Identification of a 14mer RNA that recognizes and binds flavin mononucleotide with high affinity. Nucleic Acids Res 33, 6992-6999 (2005)

Flavin biosensor





FRET “on”

FRET “off”

Aptamers

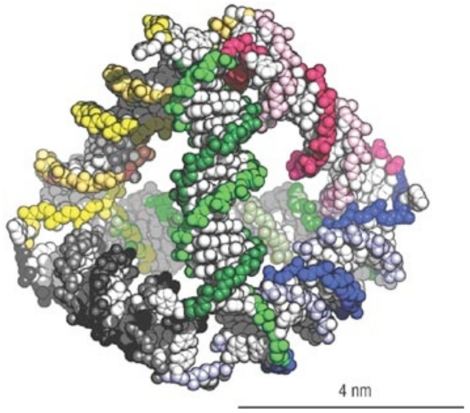
Selectivity in binding
  naturally stereospecific

Easy to generate
  it's all in the selection

Can be combined modularly

Can incorporate chemical modifications
  add unnatural chemical functionality
    (eg, fluorphores)

Nucleic Acids

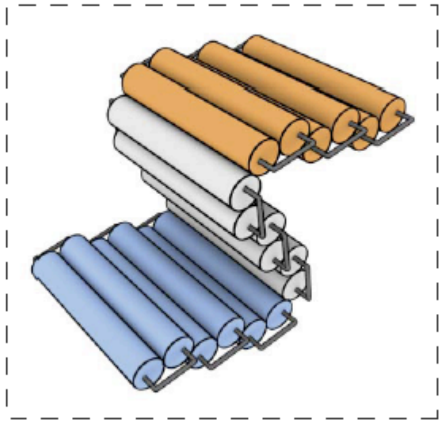


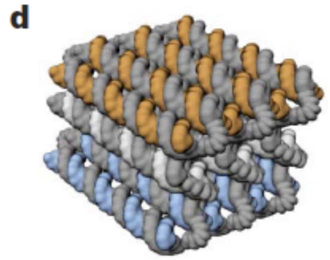
Excellent for rational design
    exquisitely addressable

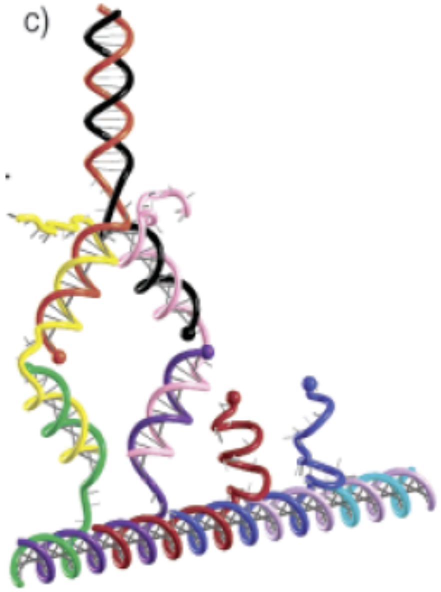
Synthesize chemically or biosynthetically

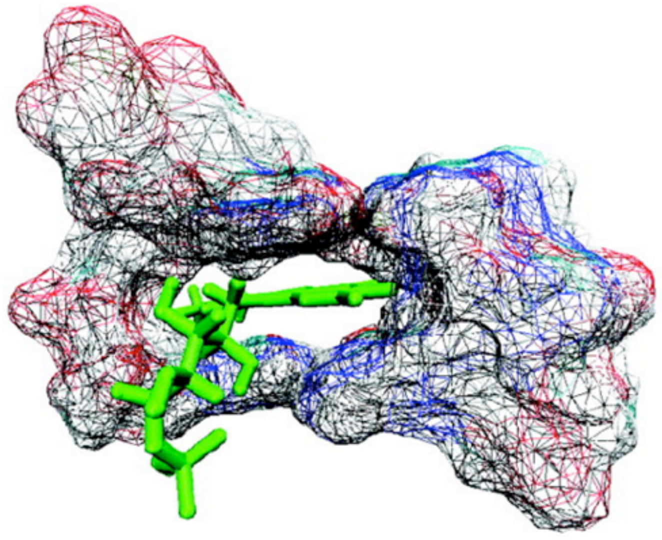
Combinatorial selection
    extremely powerful

Readily re-designed after selection

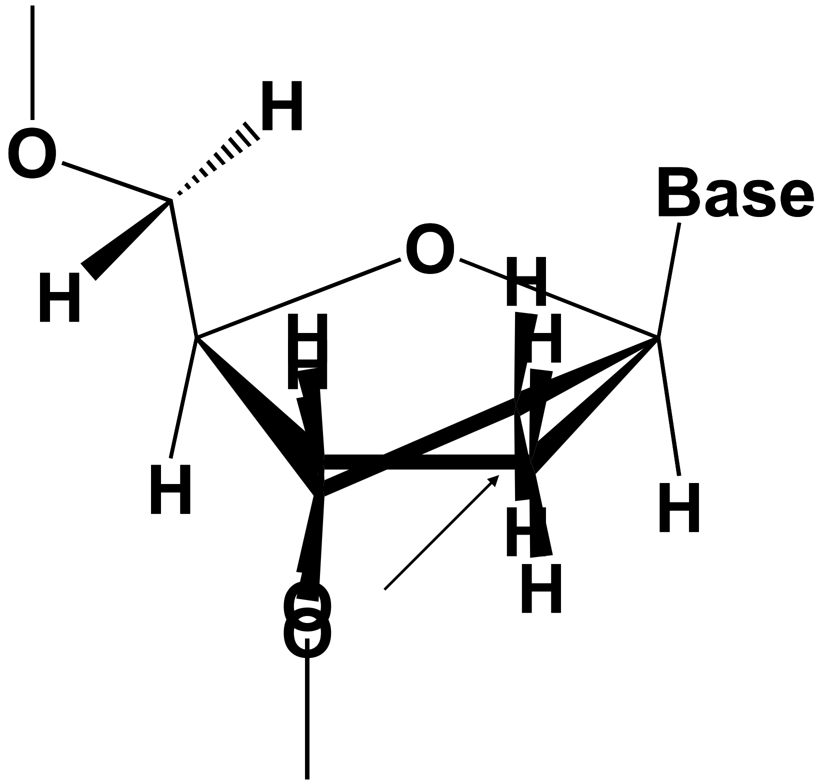


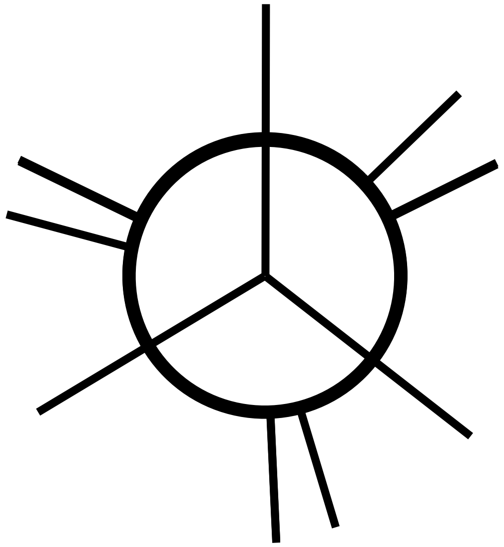


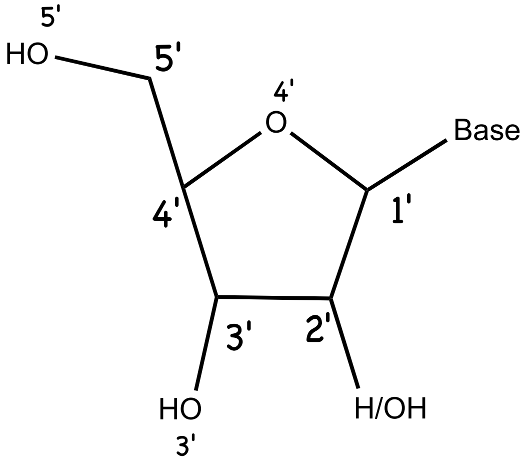


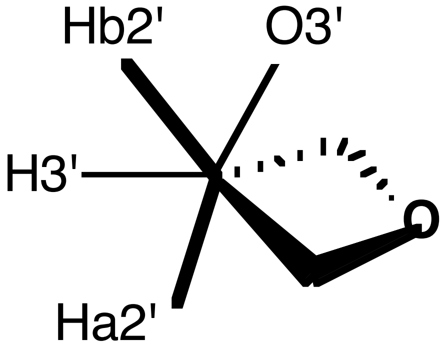


Furanose Sugar Ring

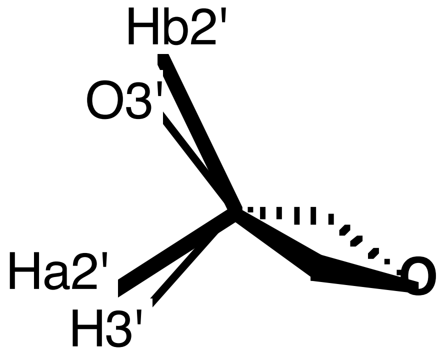






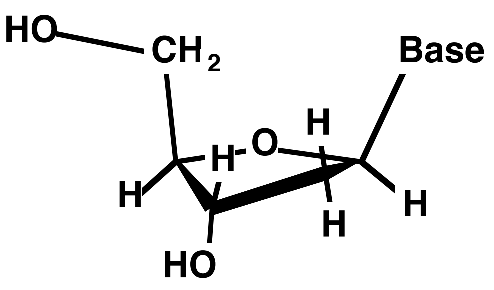


Staggered
Puckered

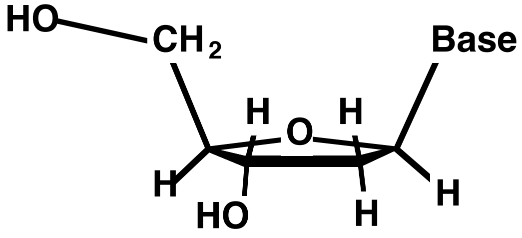


Eclipsed
Planar

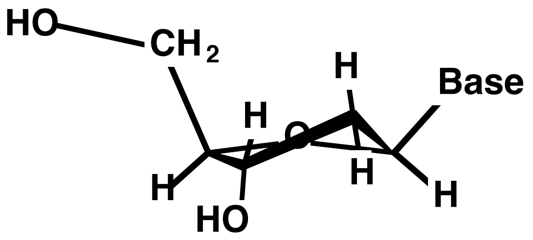
Furanose Sugar Ring



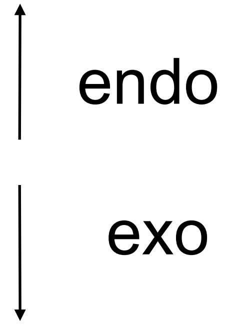
C3' exo

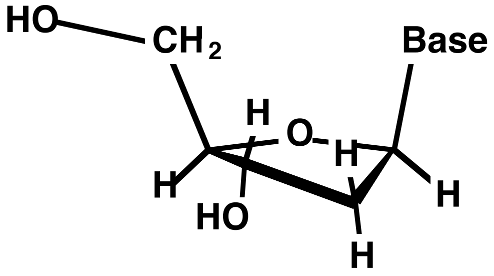


Unpuckered

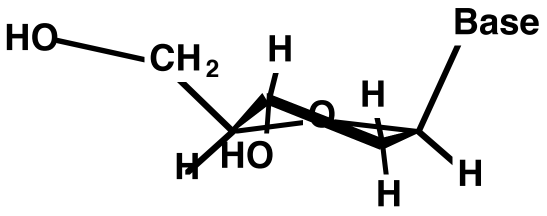


C2' endo

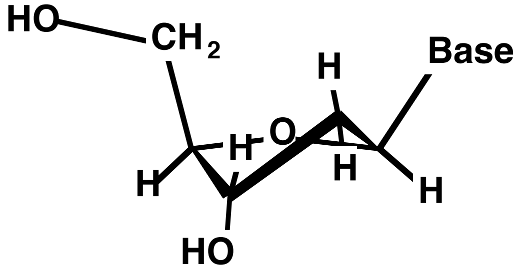




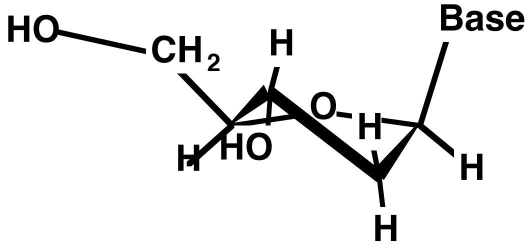
C2' exo



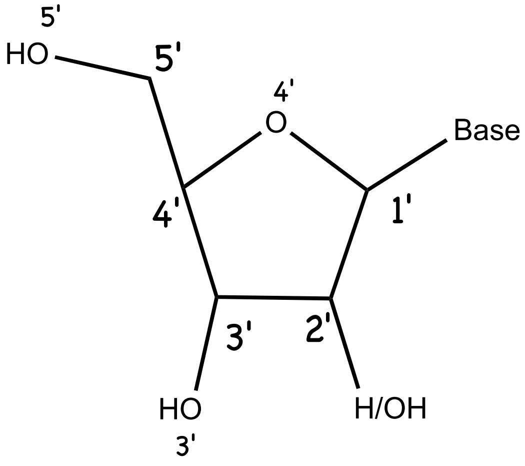
C3' endo

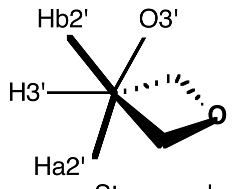


C2' endo - C3' exo

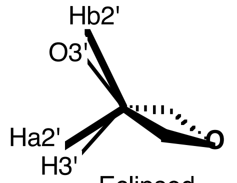


C2' exo - C3' endo

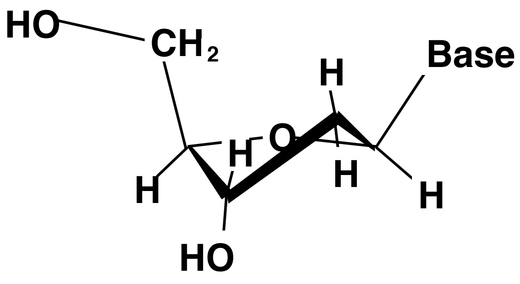




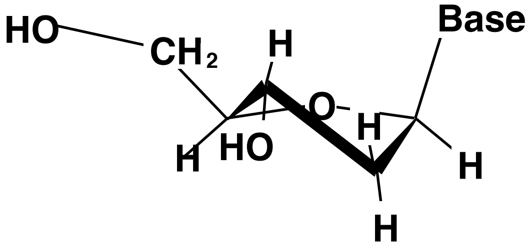
Staggered
Puckered



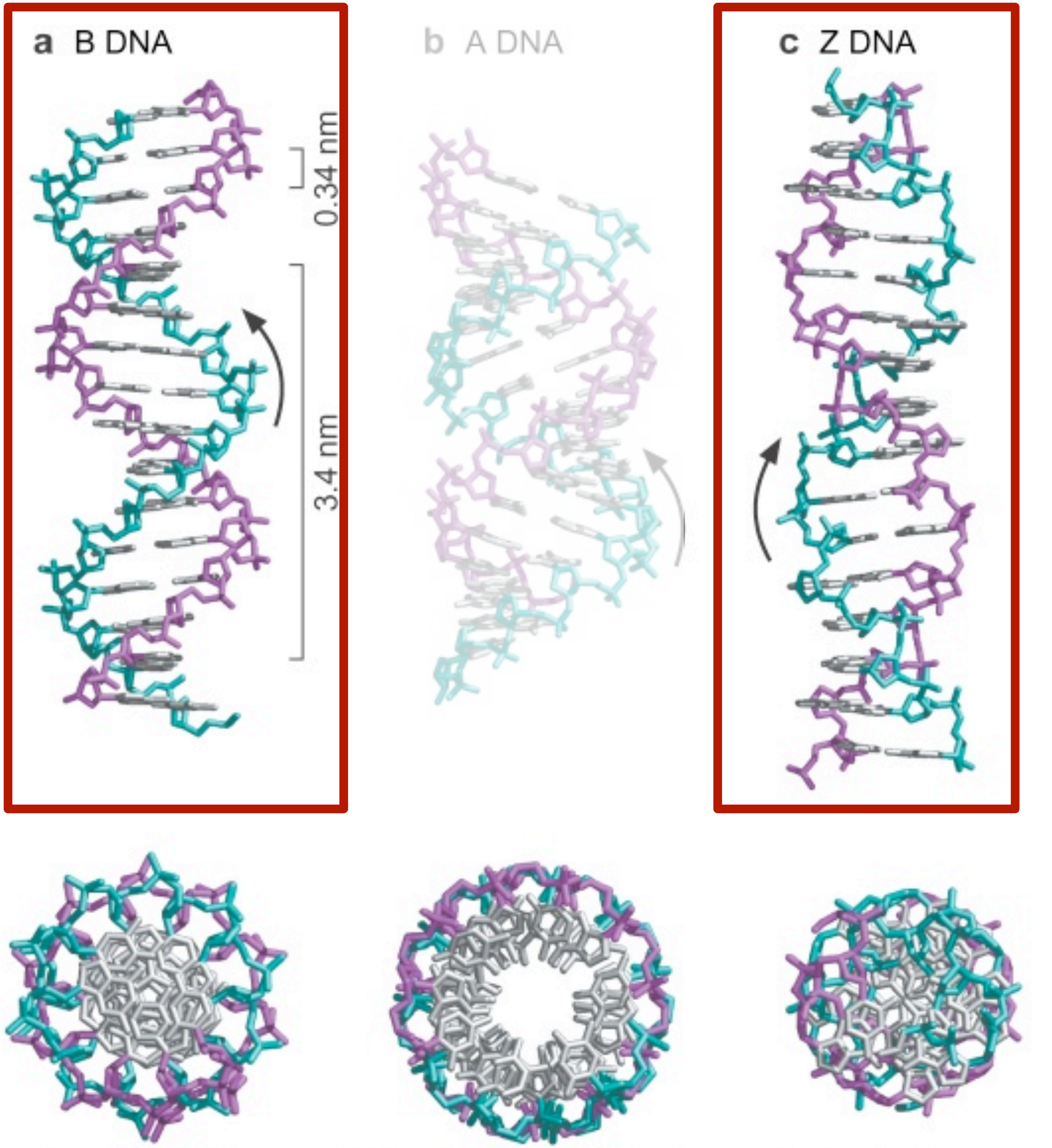
Eclipsed
Planar



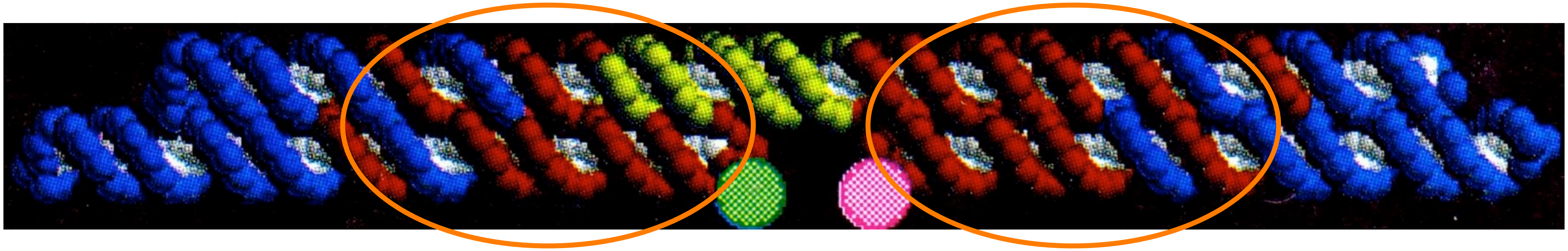
C2' endo - C3' exo

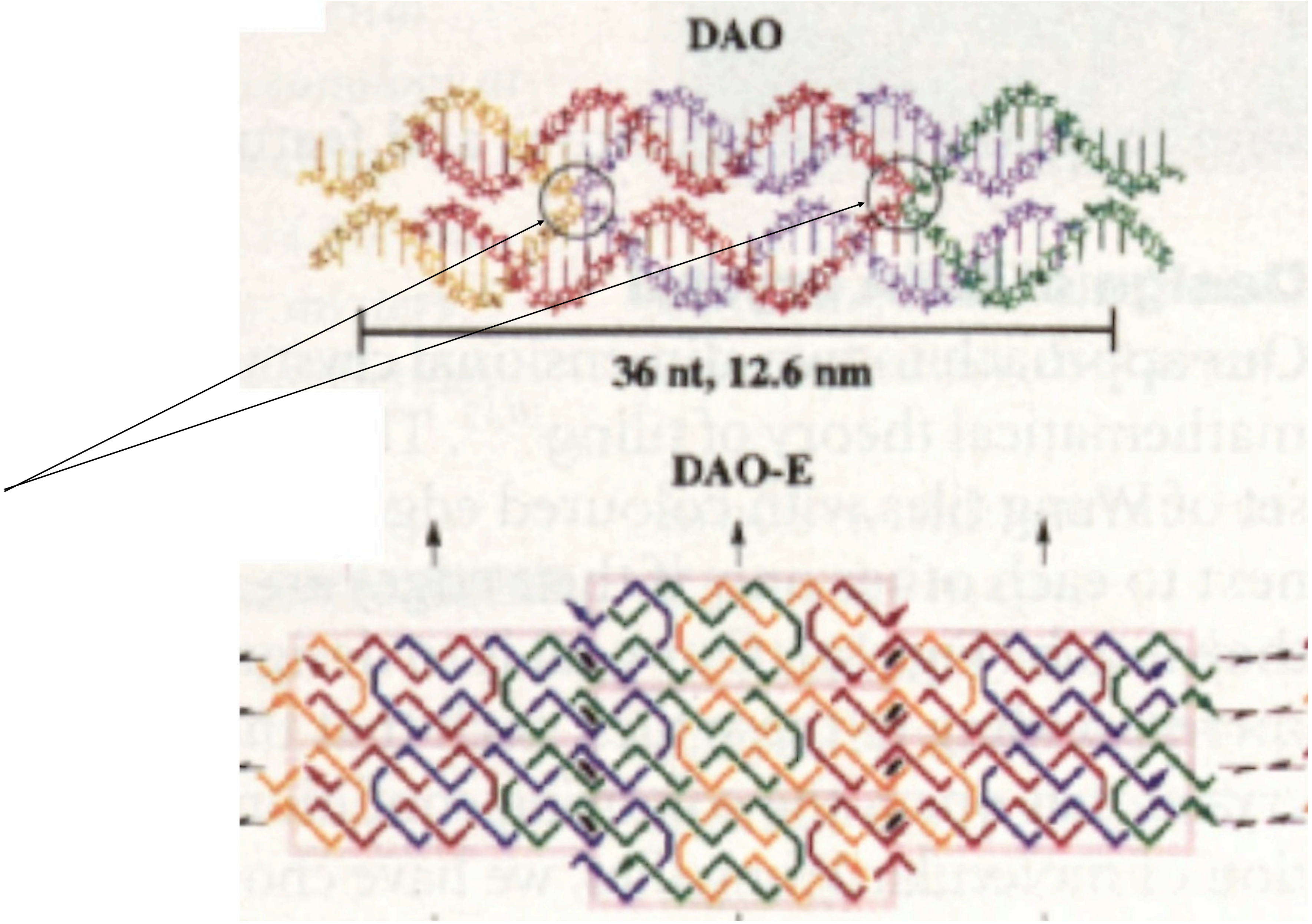


C2' exo - C3' endo



Setup - Double crossover "**wings**"



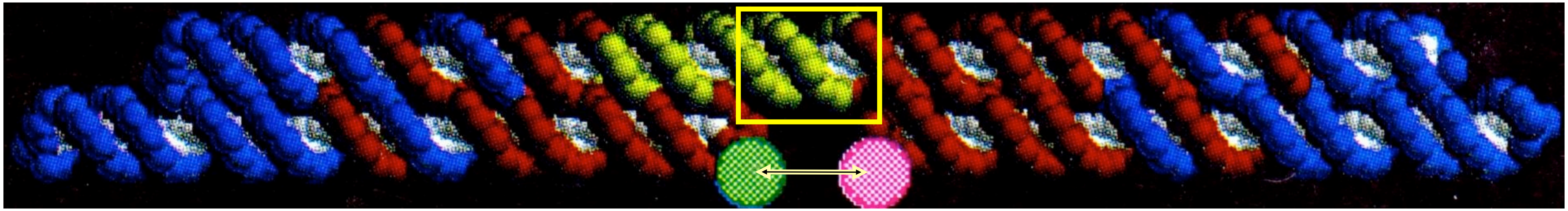


Color schemes **not** the same

Mao, C., Sun, W., Shen, Z. & Seeman, N. C. A nanomechanical device based on the B-Z transition of DNA. Nature 397, 144-146 (1999)

B->Z leads to Rotation, leads to Δd

B-DNA

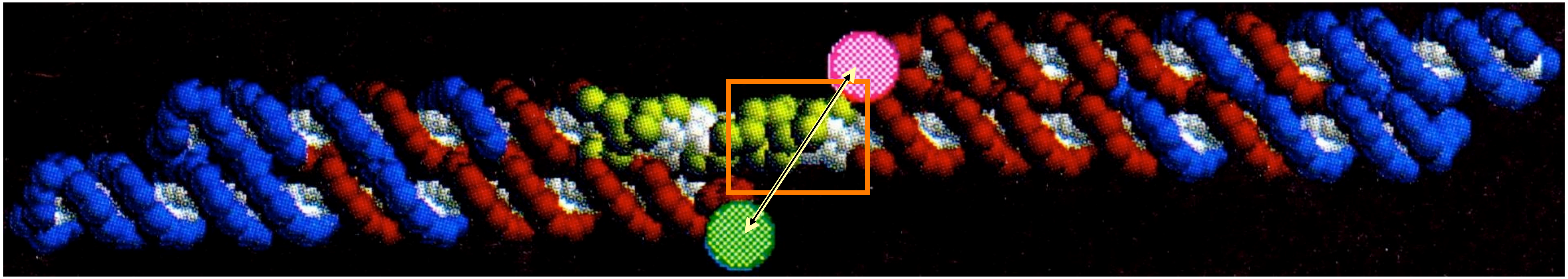


Rise per nucleotide = 3.38 Å

Rotation per nucleotide = 36°

Difference = 6° per nucleotide

Z-DNA

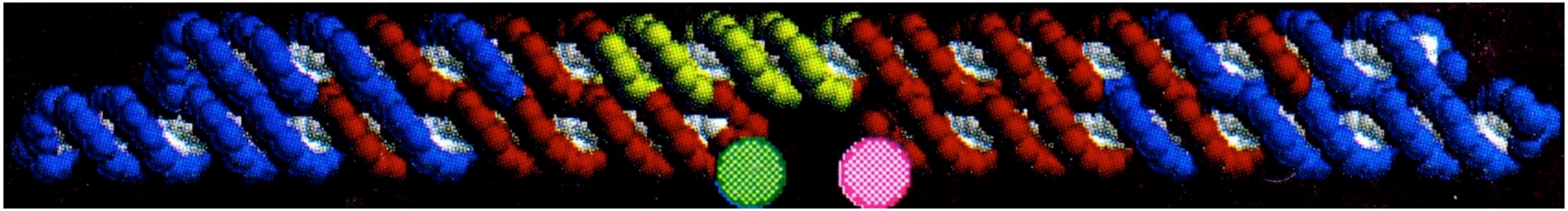


Ave. rise per nucleotide = 3.71 Å

Ave. rotation per nucleotide = 30°

Mao, C., Sun, W., Shen, Z. & Seeman, N. C. A nanomechanical device based on the B-Z transition of DNA. Nature 397, 144-146 (1999)

Förster Energy Transfer - Setup

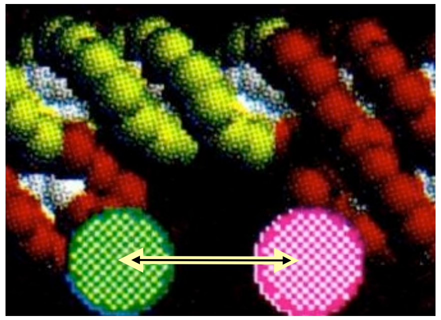


Donor:
Fluorescein

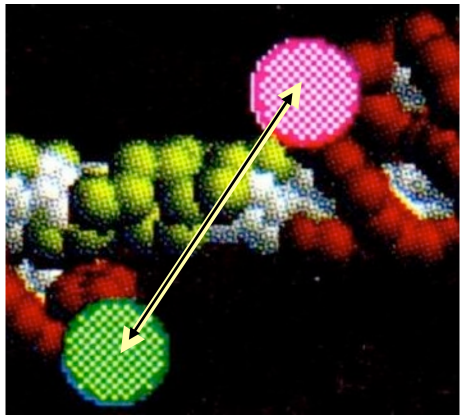
Acceptor:
Cy3 (a sulfoindocarbocyanine)

Efficiency of transfer:

$$E = \frac{R_o^6}{R_o^6 + R^6}$$



51 Å
B-form

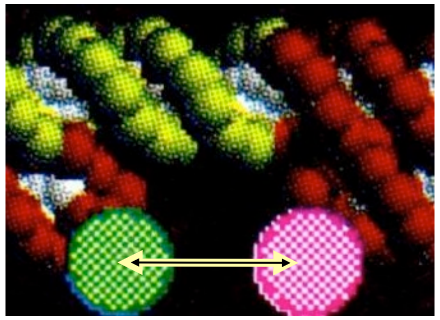


70 Å
Z-form

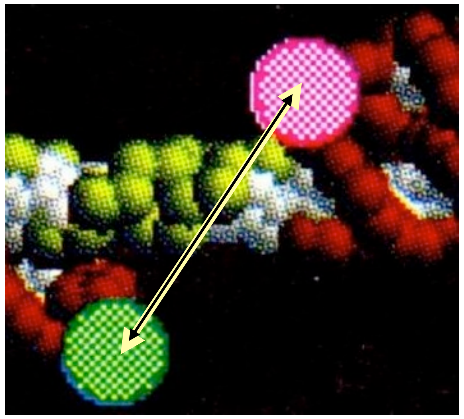
Distances from model building

Mao, C., Sun, W., Shen, Z. & Seeman, N. C. A nanomechanical device based on the B-Z transition of DNA. Nature 397, 144-146 (1999)

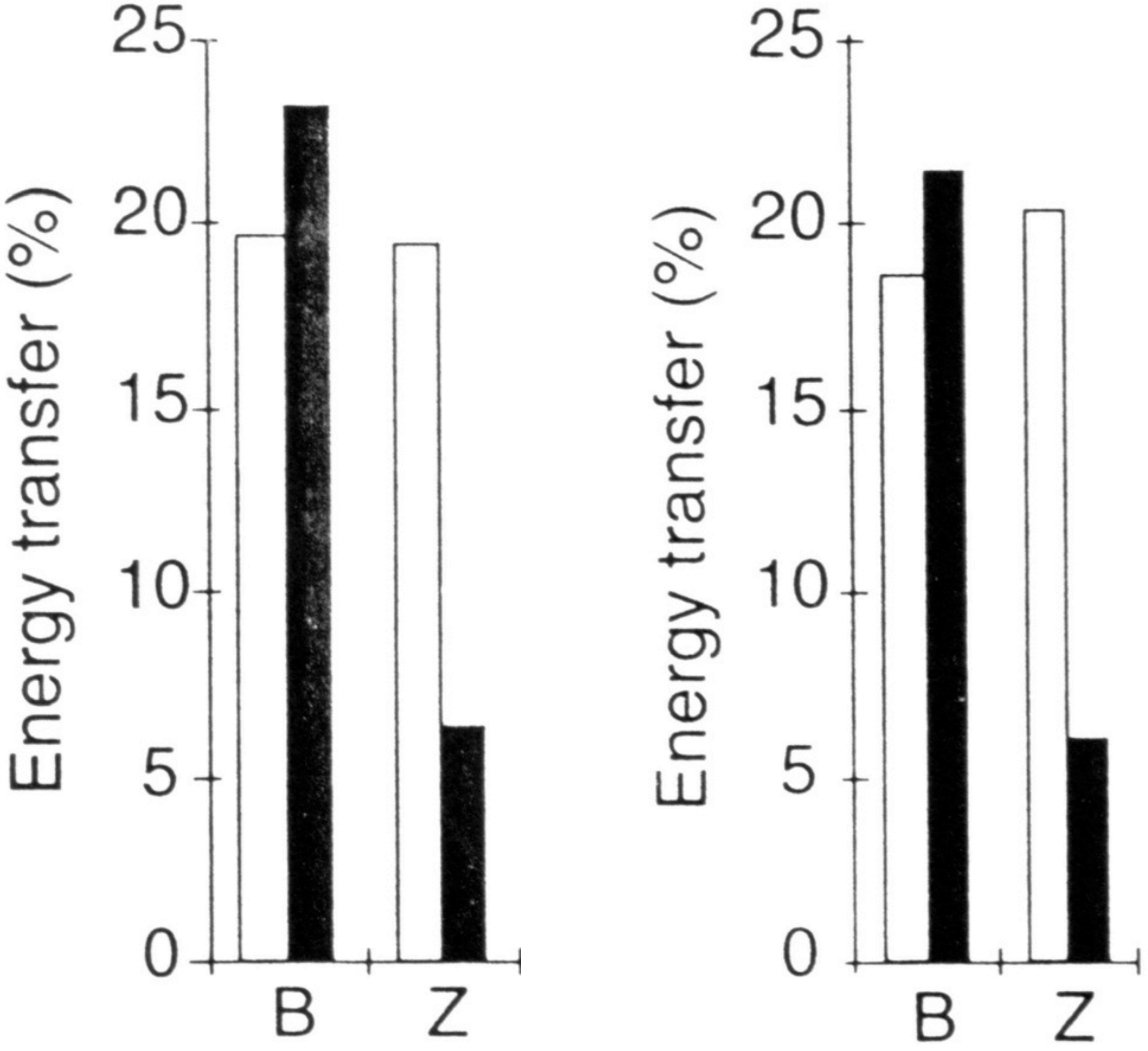
Förster Energy Transfer - Results



B-form



Z-form



From decrease in donor fluorescence

From increase in acceptor fluorescence

☐ Control

■ Proto-Z $d(CG)_{10}$

Conditions:

10 mM cacodylate pH 7.5
100 mM $MgCl_2$
100 mM NaCl

plus

“**B**” 0.00 mM $Co(NH_3)_6Cl_3$

“**Z**” 0.25 mM $Co(NH_3)_6Cl_3$

R_o for the pair (calculated) = 55.7 Å (B-DNA conditions)
= 56.1 Å (Z-DNA conditions)

Mao, C., Sun, W., Shen, Z. & Seeman, N. C. A nanomechanical device based on the B-Z transition of DNA. Nature 397, 144-146 (1999)

For next time...

Which is most stable?

5′-ACCGCCGACGT-3′
3′-TGGCGGCTGCA-5′

5′-ACCGCCGACGT-3′ ←
3′-AGGCGGCTGCC-5′

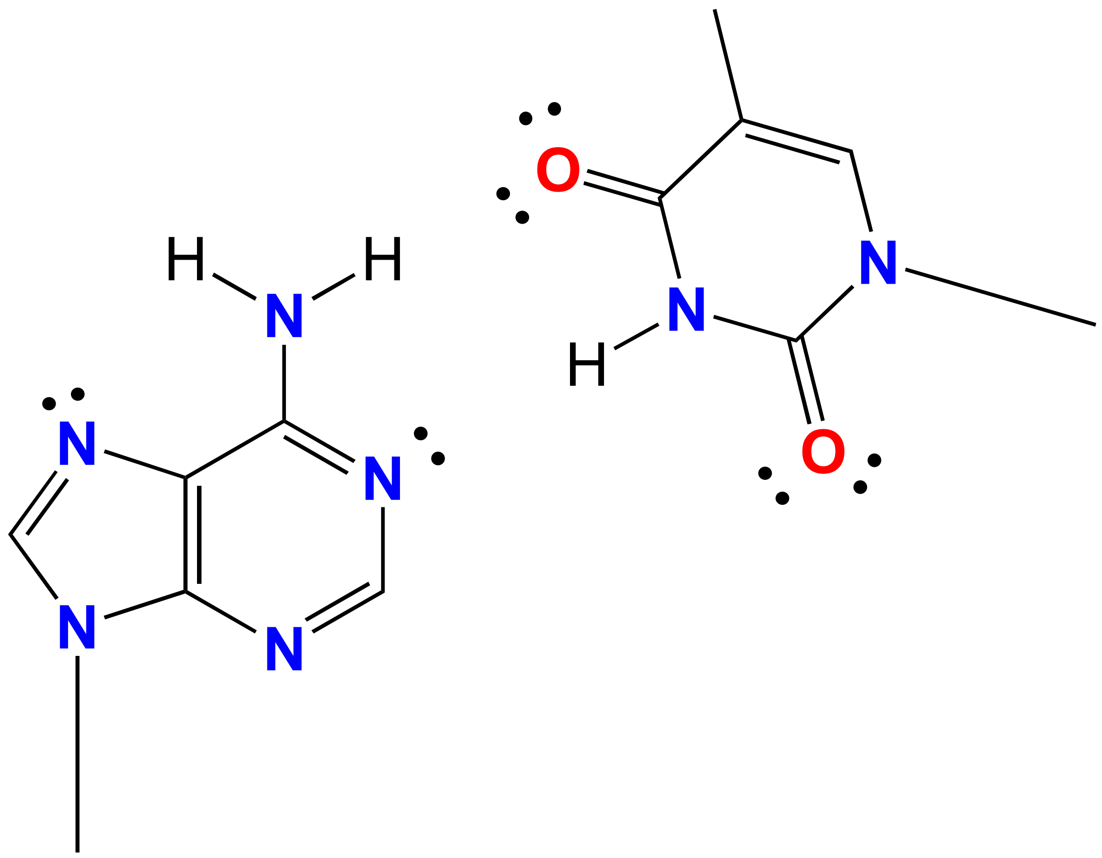
What stabilizes protein structures?

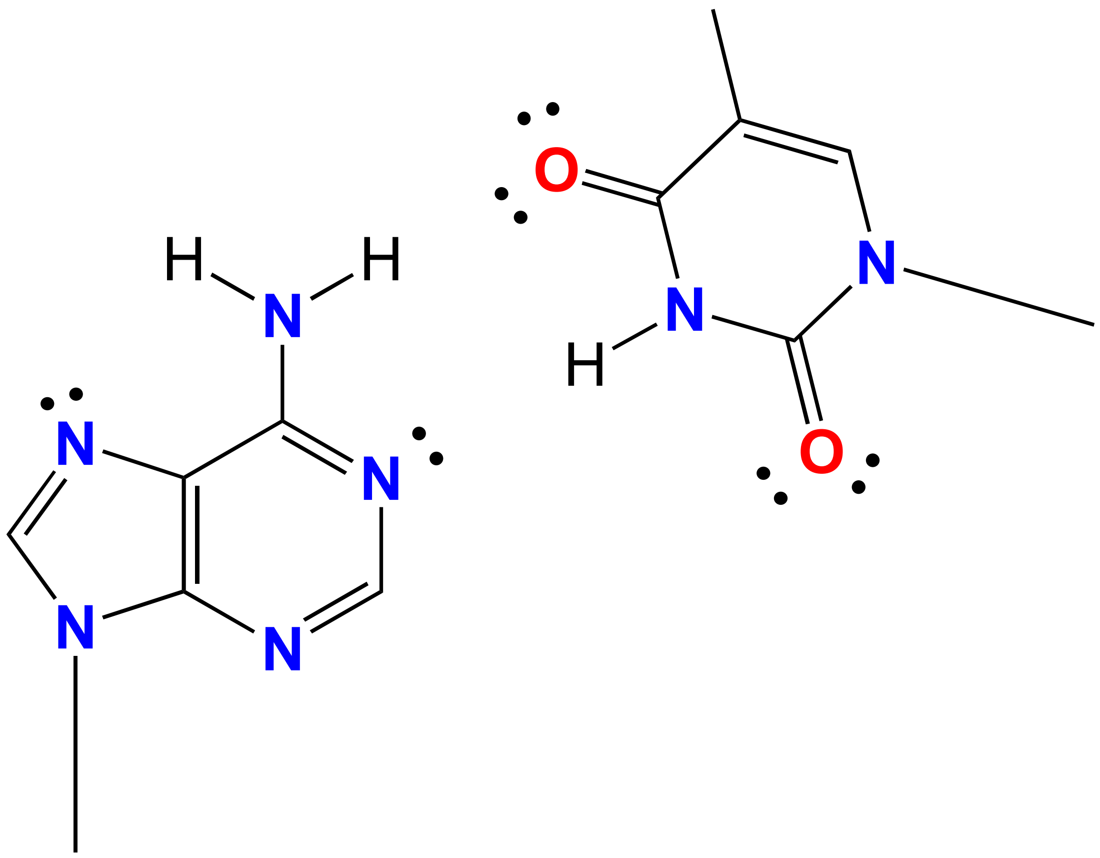
What *directs* protein structures?

The DNA Duplex

What stabilizes the duplex?

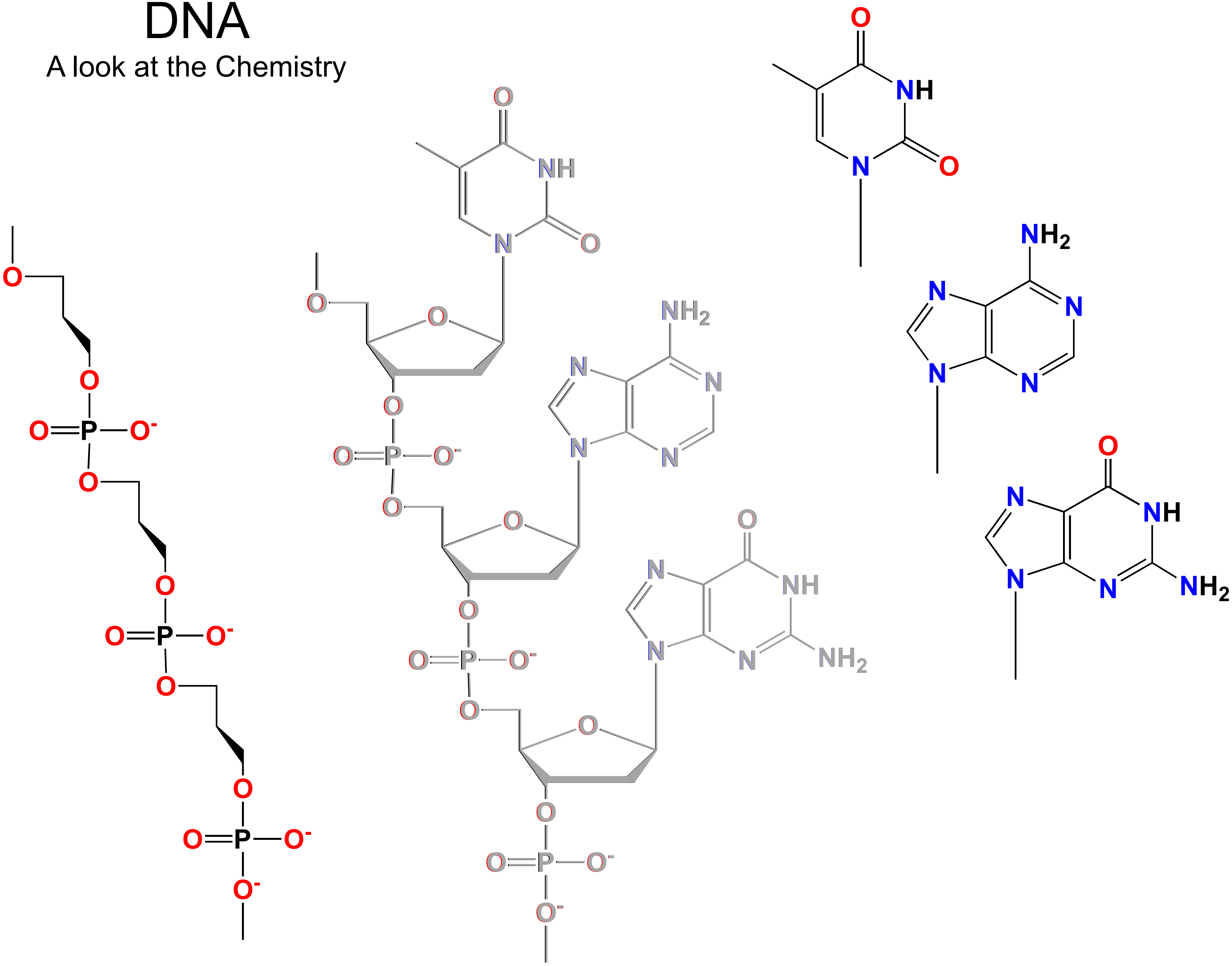
What *directs* duplex structure?





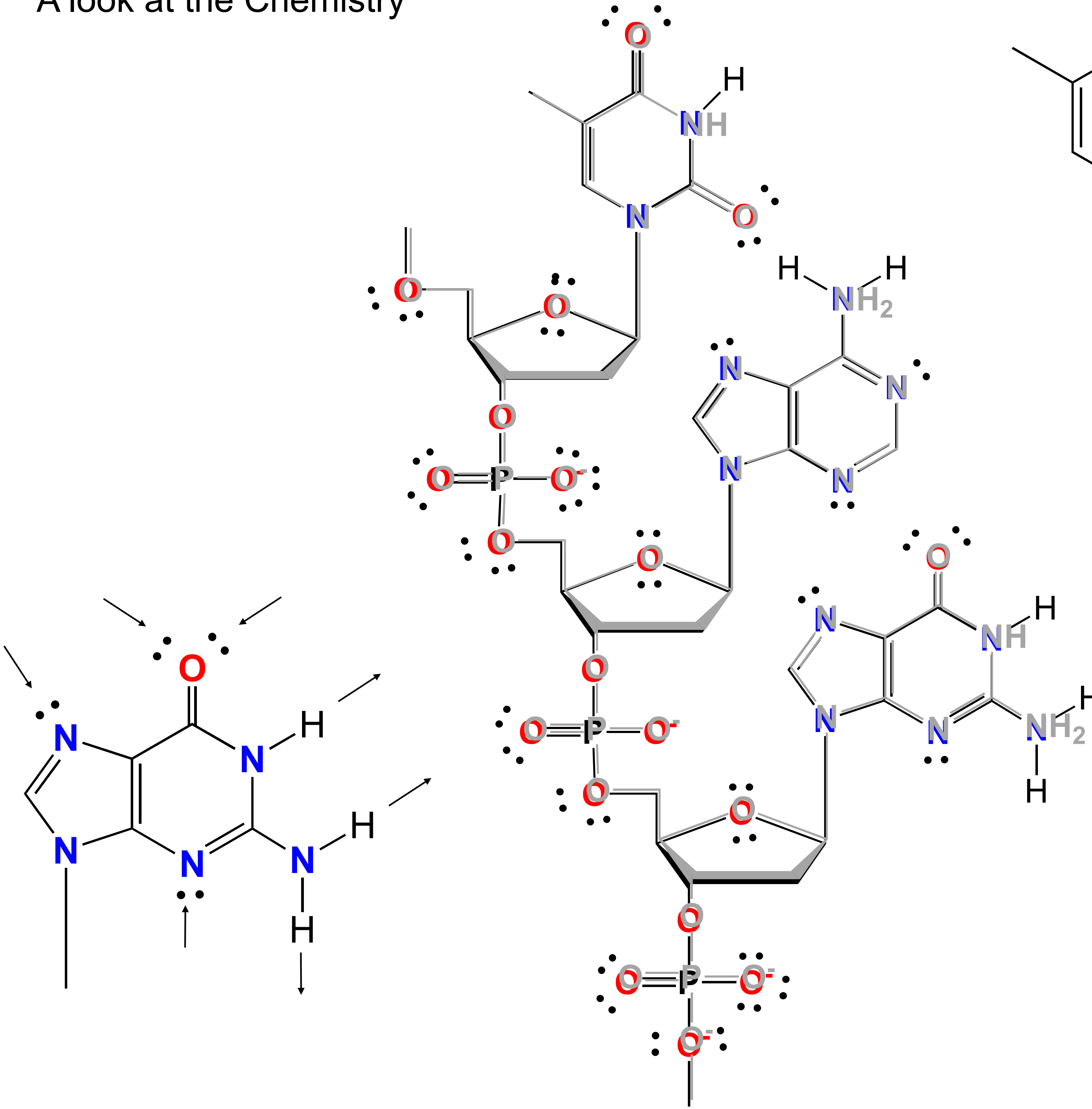
DNA

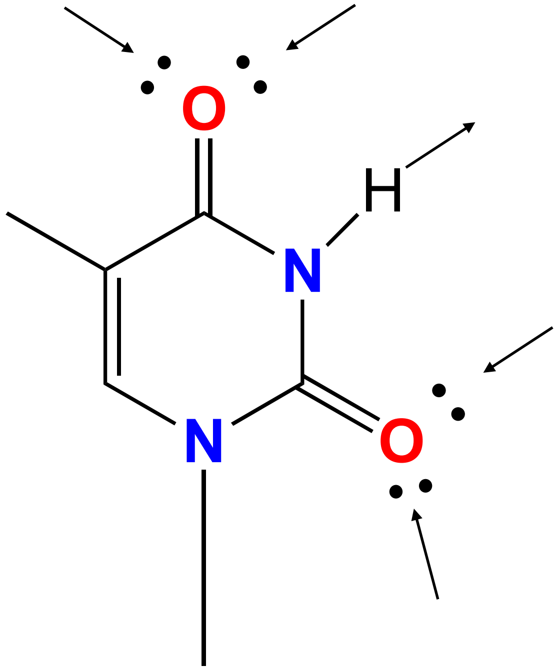
A look at the Chemistry

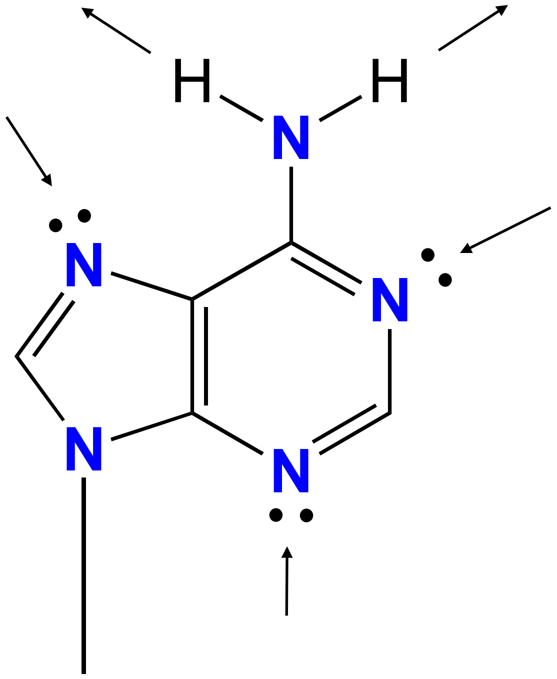


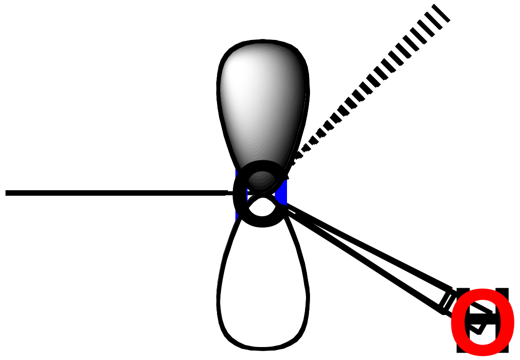
DNA

A look at the Chemistry

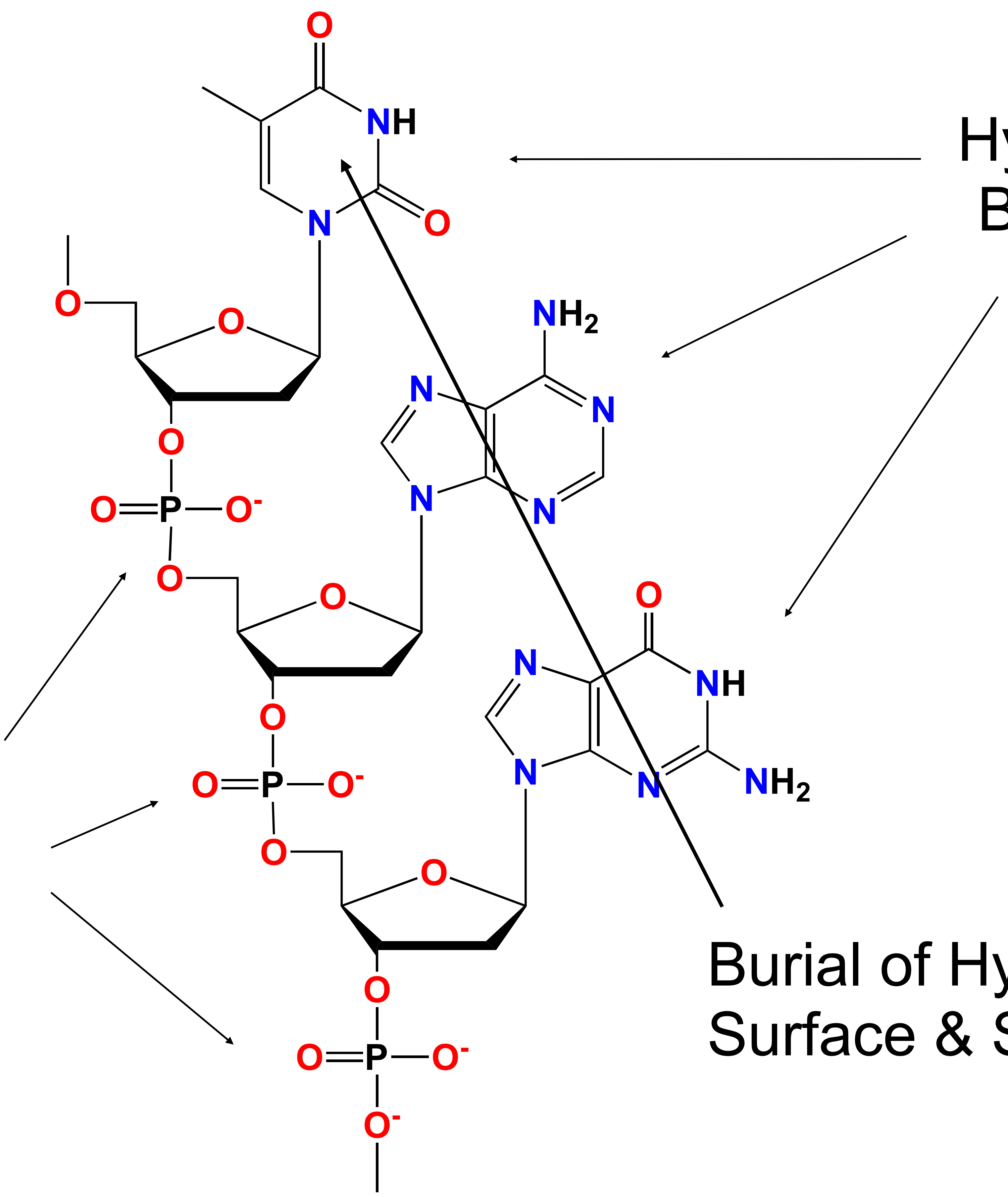








What forces are important?



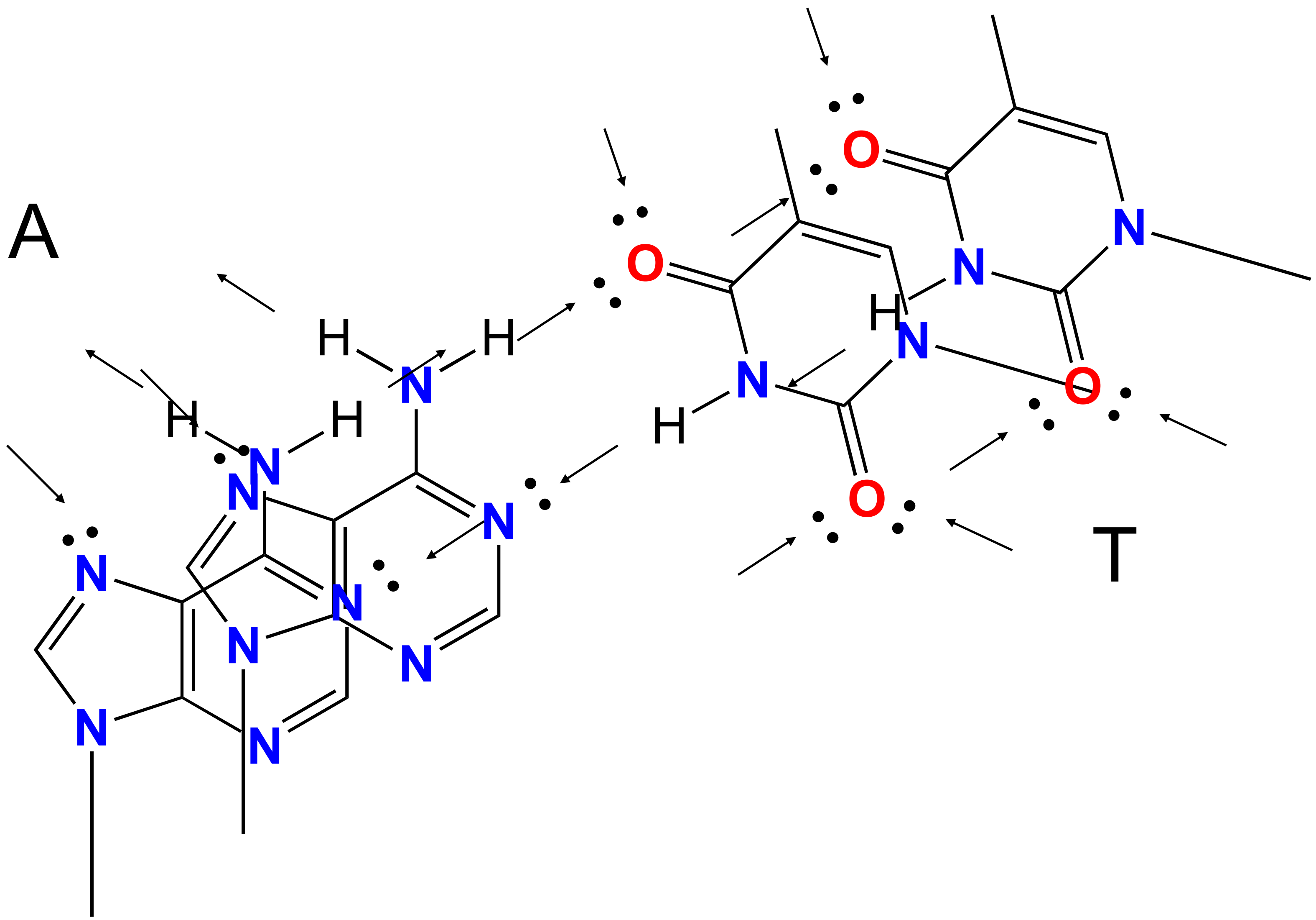
Hydrogen Bonding

Electrostatics

Burial of Hydrophobic Surface & Stacking

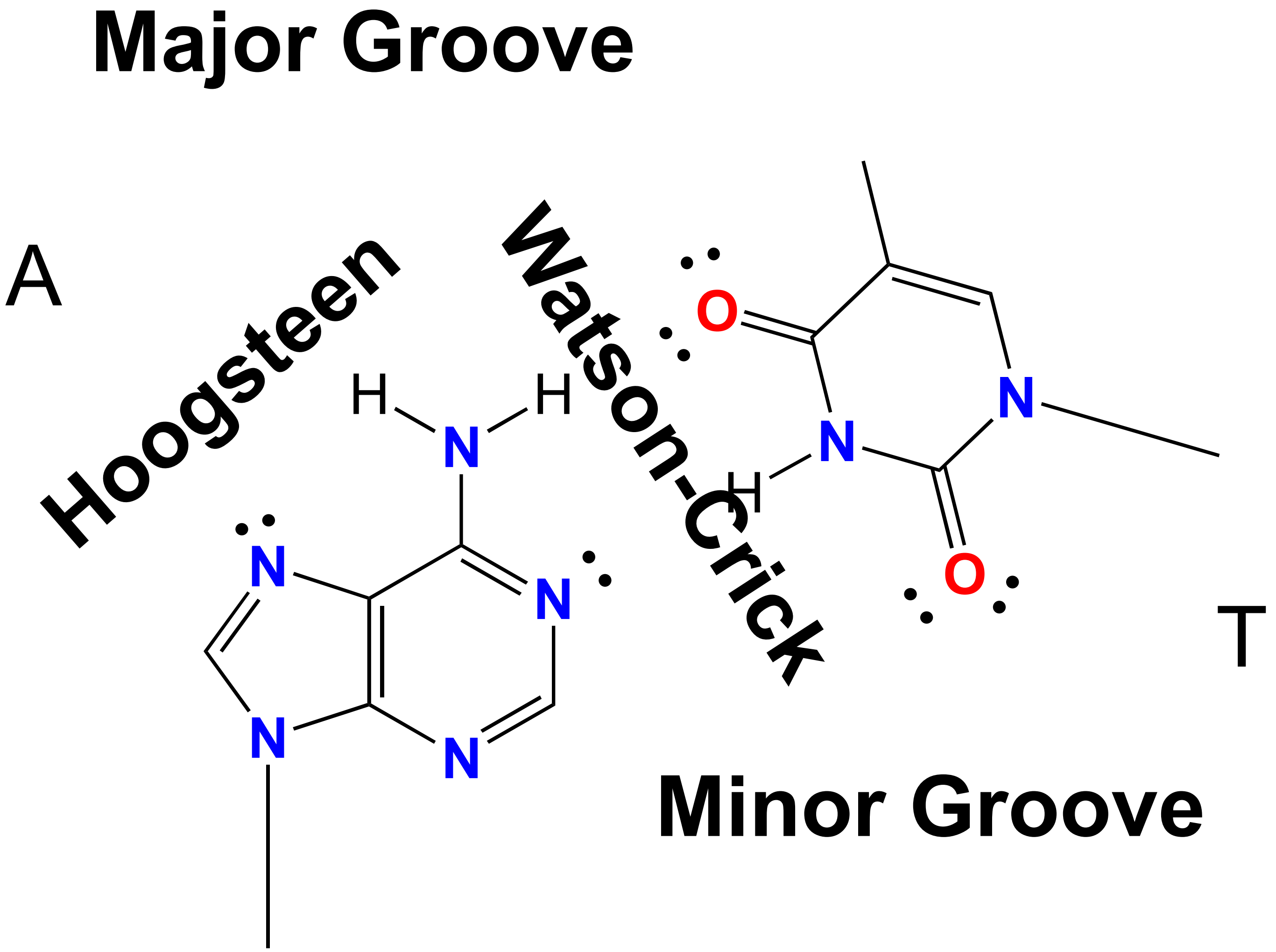
Base Pairing

(Donors matched to Acceptors)



Base Pairing

(Donors matched to Acceptors)



Major Groove

A

Hoogsteen

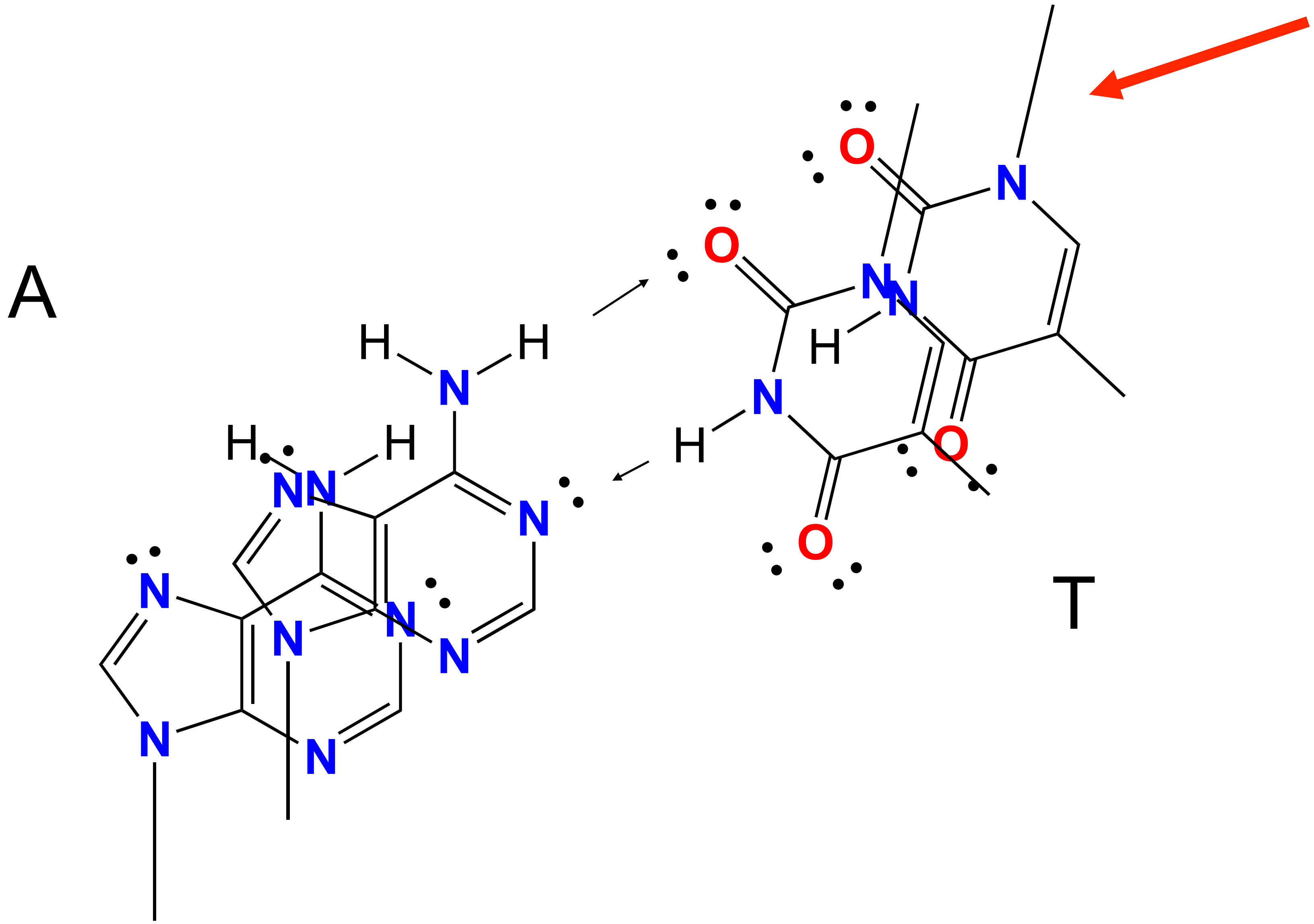
Watson-Crick

T

Minor Groove

Base Pairing

(Donors matched to Acceptors)



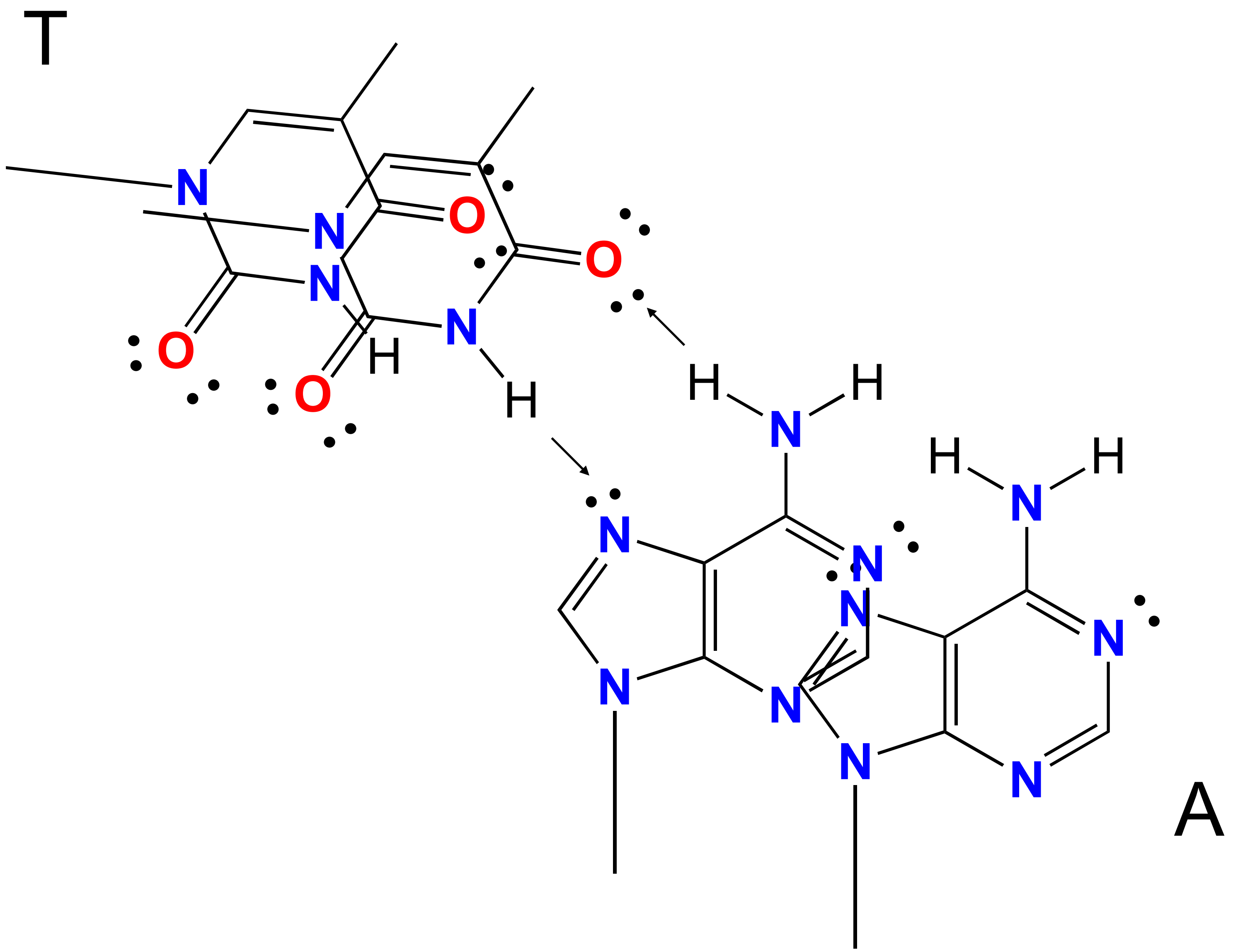
Good base pairing

Watson-Crick facing

but *Anti-Watson-Crick* orientation

Base Pairing

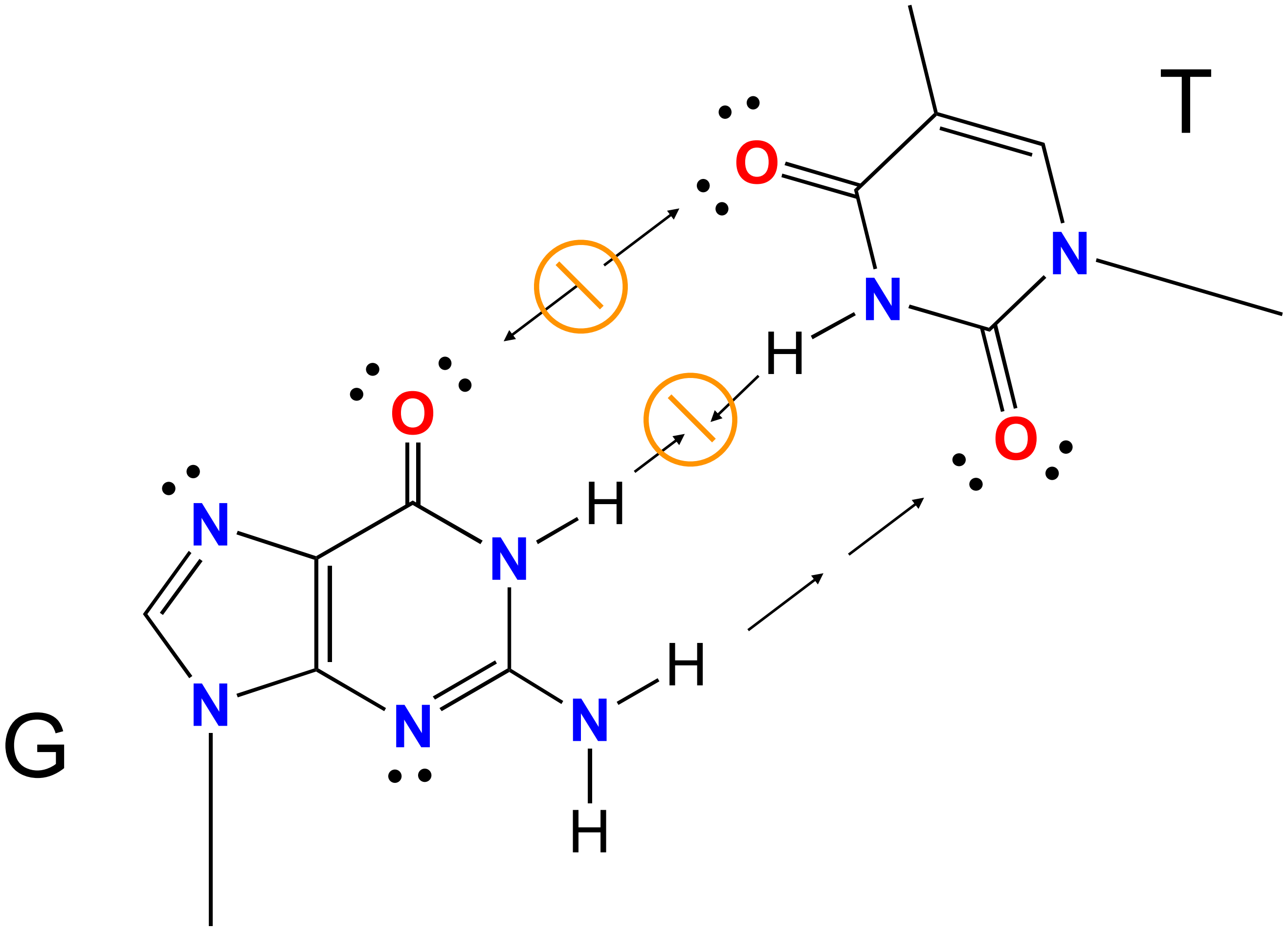
(Donors matched to Acceptors)



Good base pairing
WC-Hoogsteen facing

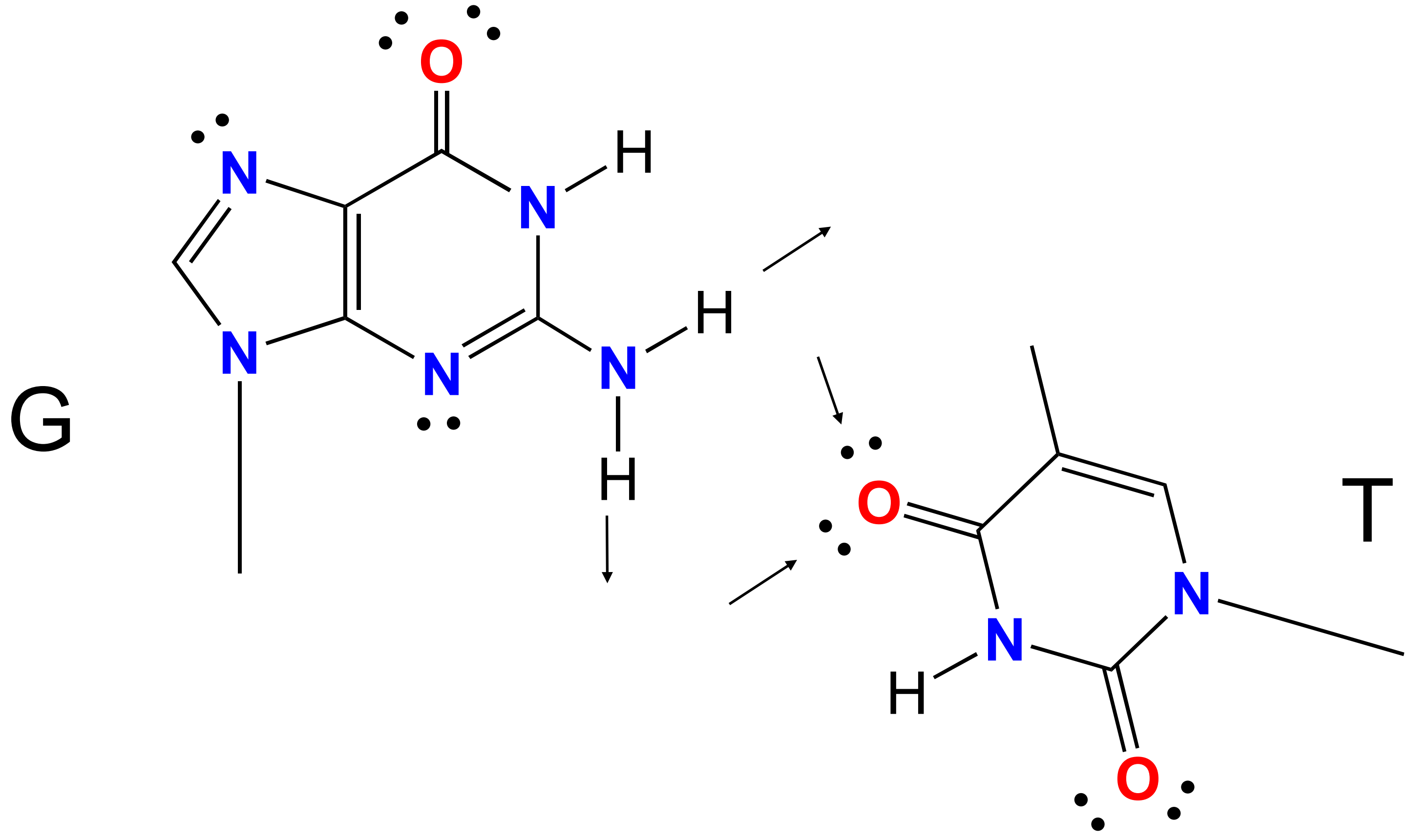
Bad Base Pairing

(Donors ***not*** matched to Acceptors)

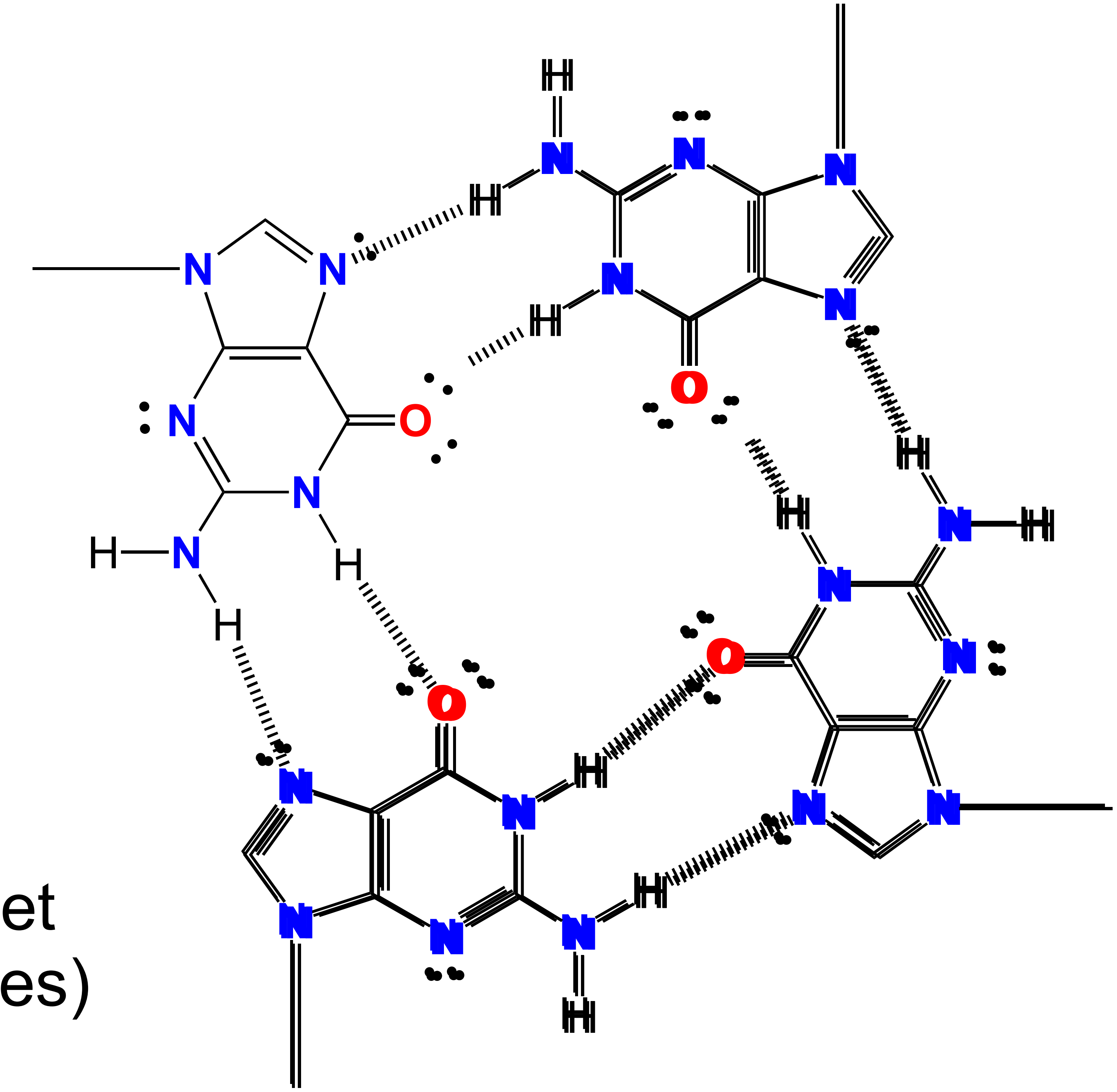


Bad Base Pairing

(Donors to Acceptors with ***terrible angles***)

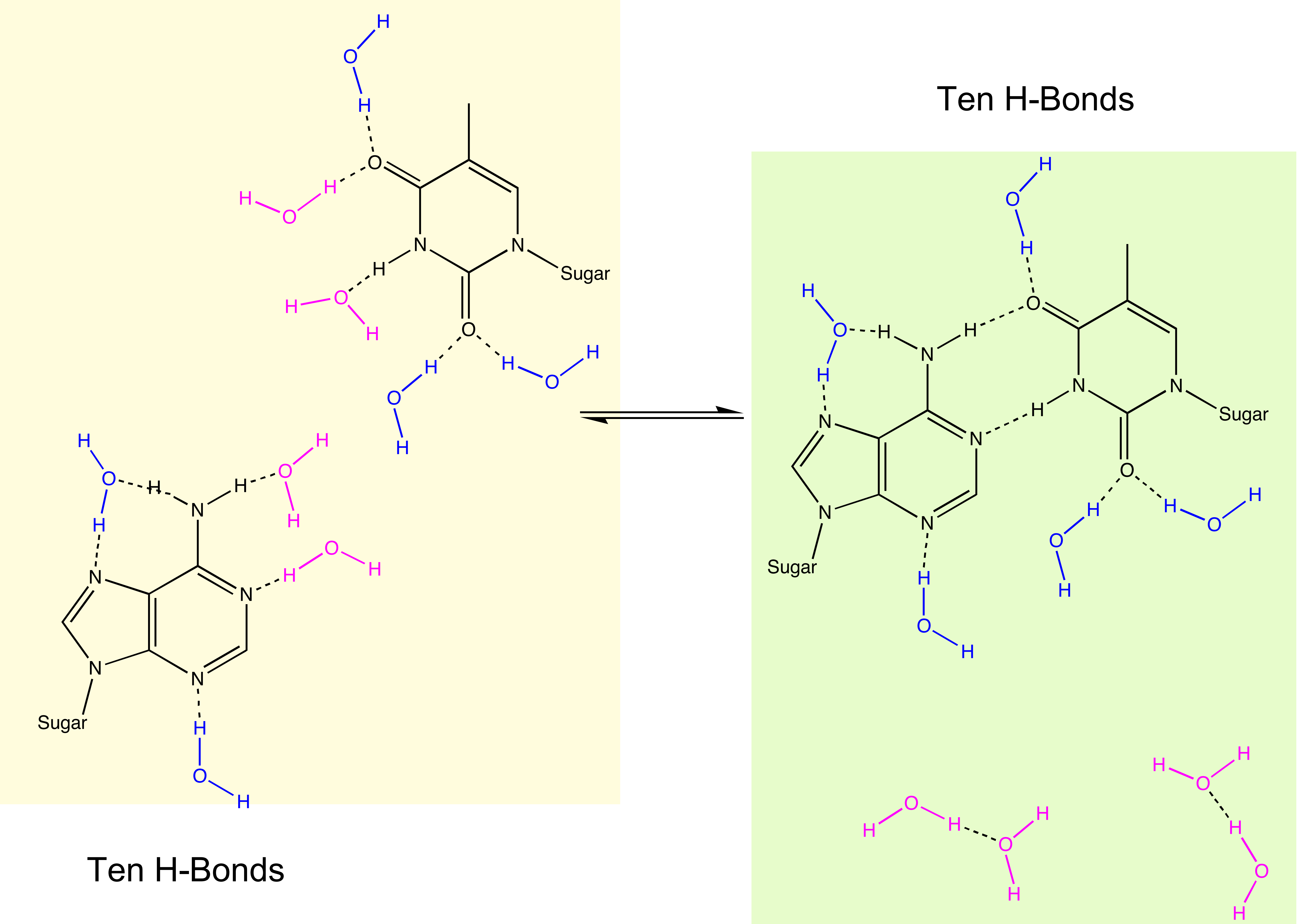


Wild (but good) Base Pairing



G-quartet
(Telomeres)

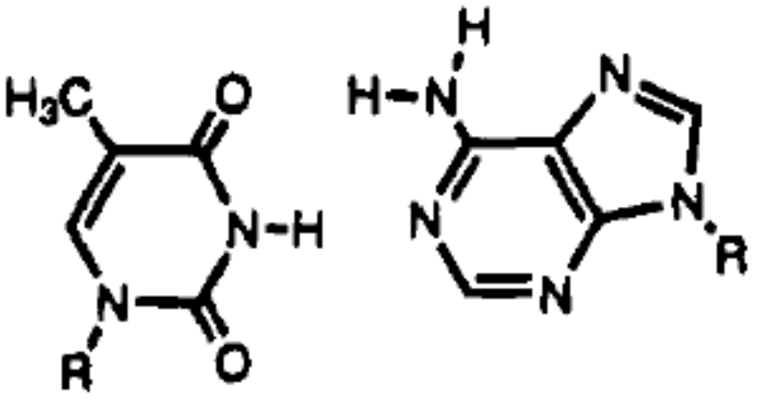
AT Base Pair



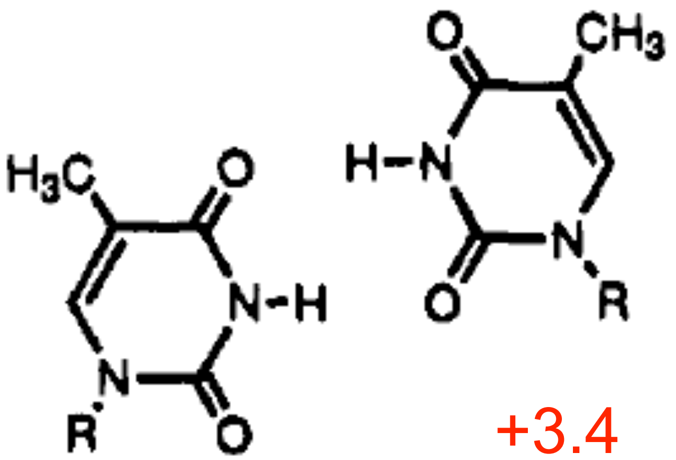
Ten H-Bonds

Ten H-Bonds

How important are H-bonds in DNA?

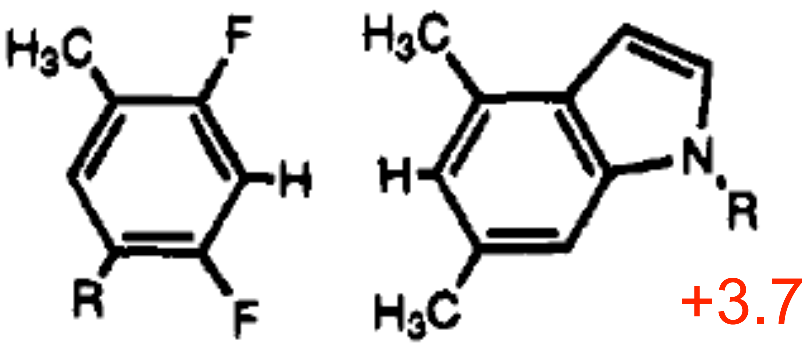


T - A



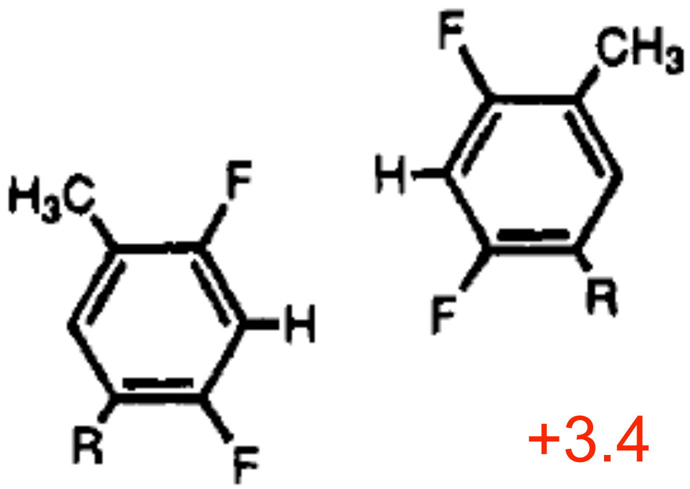
+3.4

T - T



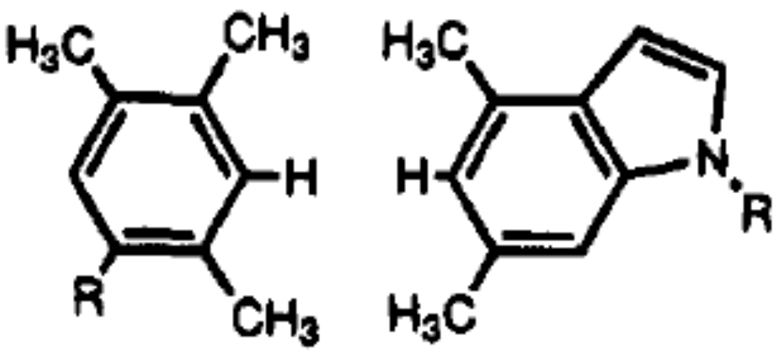
+3.7

F - D



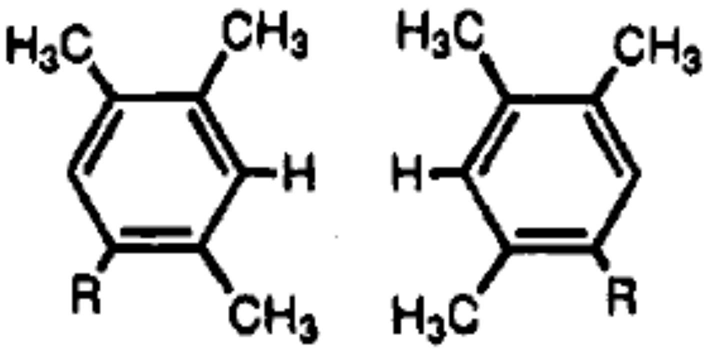
+3.4

F - F



+3.5

B - D



+3.0

B - B

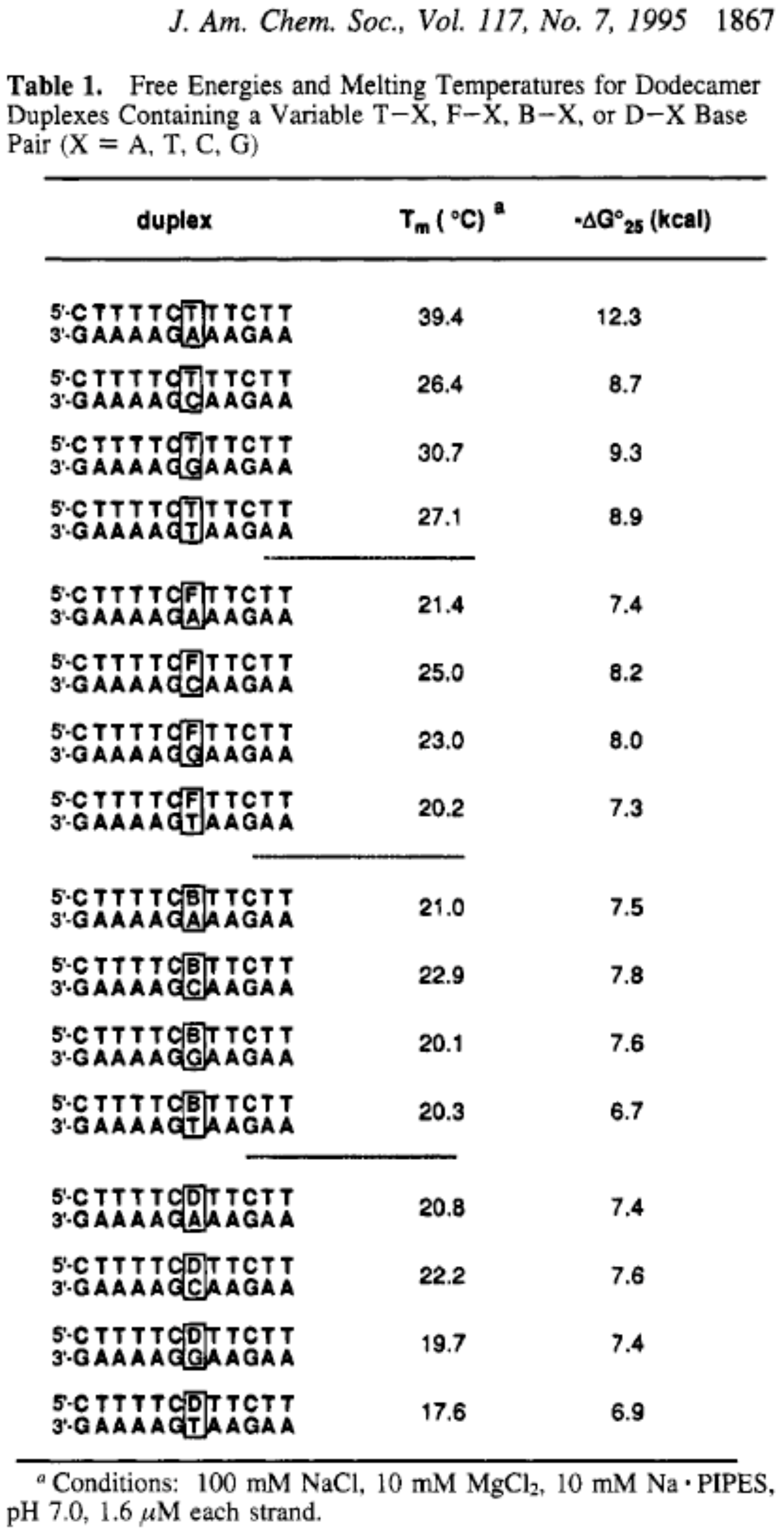
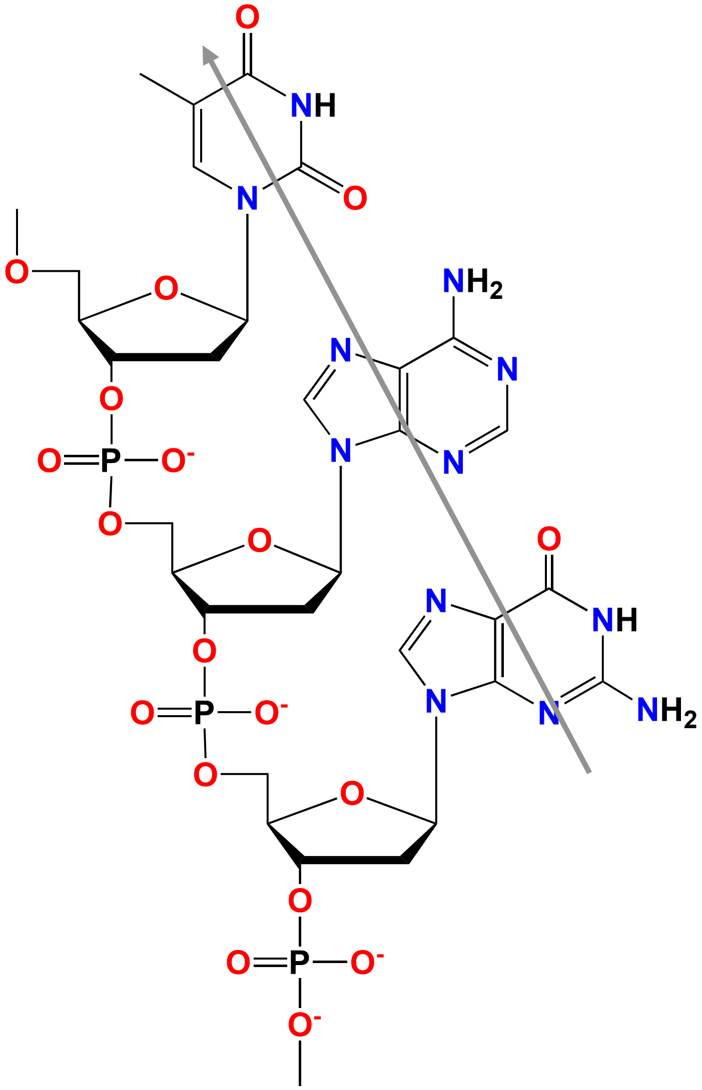

J. Am. Chem. Soc., Vol. 117, No. 7, 1995 1867

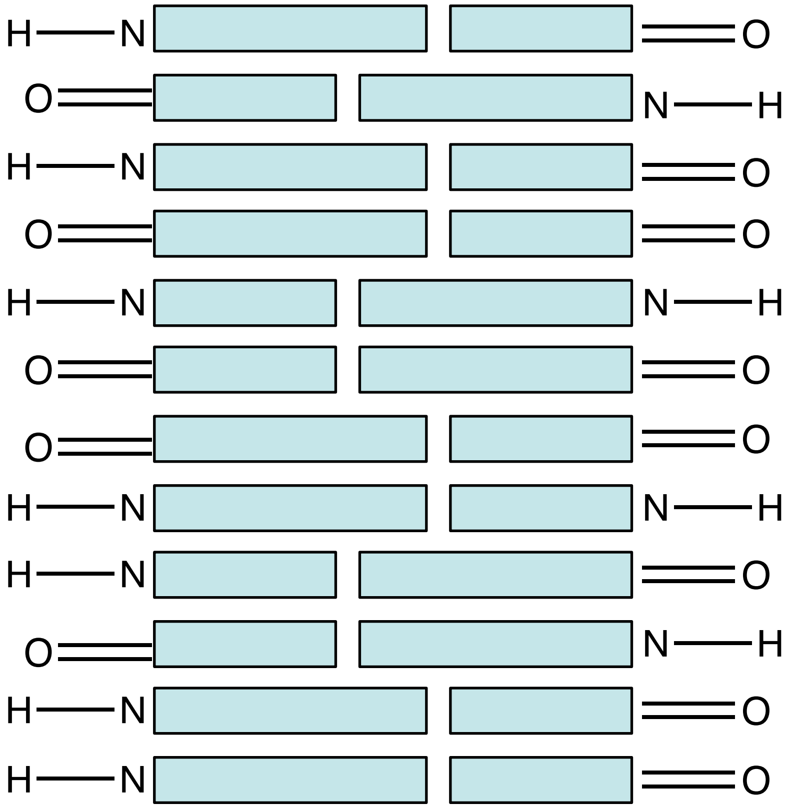
Table 1. Free Energies and Melting Temperatures for Dodecamer Duplexes Containing a Variable T−X, F−X, B−X, or D−X Base Pair (X = A, T, C, G)

duplex	T_m (°C)[a]	-$\Delta G°_{25}$ (kcal)
5'-CTTTTCTTTCTT 3'-GAAAAGAAAGAA	39.4	12.3
5'-CTTTTCTTTCTT 3'-GAAAAGCAAGAA	26.4	8.7
5'-CTTTTCTTTCTT 3'-GAAAAGGAAGAA	30.7	9.3
5'-CTTTTCTTTCTT 3'-GAAAAGTAAGAA	27.1	8.9
5'-CTTTTCFTTCTT 3'-GAAAAGAAAGAA	21.4	7.4
5'-CTTTTCFTTCTT 3'-GAAAAGCAAGAA	25.0	8.2
5'-CTTTTCFTTCTT 3'-GAAAAGGAAGAA	23.0	8.0
5'-CTTTTCFTTCTT 3'-GAAAAGTAAGAA	20.2	7.3
5'-CTTTTCBTTCTT 3'-GAAAAGAAAGAA	21.0	7.5
5'-CTTTTCBTTCTT 3'-GAAAAGCAAGAA	22.9	7.8
5'-CTTTTCBTTCTT 3'-GAAAAGGAAGAA	20.1	7.6
5'-CTTTTCBTTCTT 3'-GAAAAGTAAGAA	20.3	6.7
5'-CTTTTCDTTCTT 3'-GAAAAGAAAGAA	20.8	7.4
5'-CTTTTCDTTCTT 3'-GAAAAGCAAGAA	22.2	7.6
5'-CTTTTCDTTCTT 3'-GAAAAGGAAGAA	19.7	7.4
5'-CTTTTCDTTCTT 3'-GAAAAGTAAGAA	17.6	6.9

[a] Conditions: 100 mM NaCl, 10 mM $MgCl_2$, 10 mM Na·PIPES, pH 7.0, 1.6 μM each strand.

Burial of hydrophobic surface drives helix formation (hydrophobic core / stacking interactions)

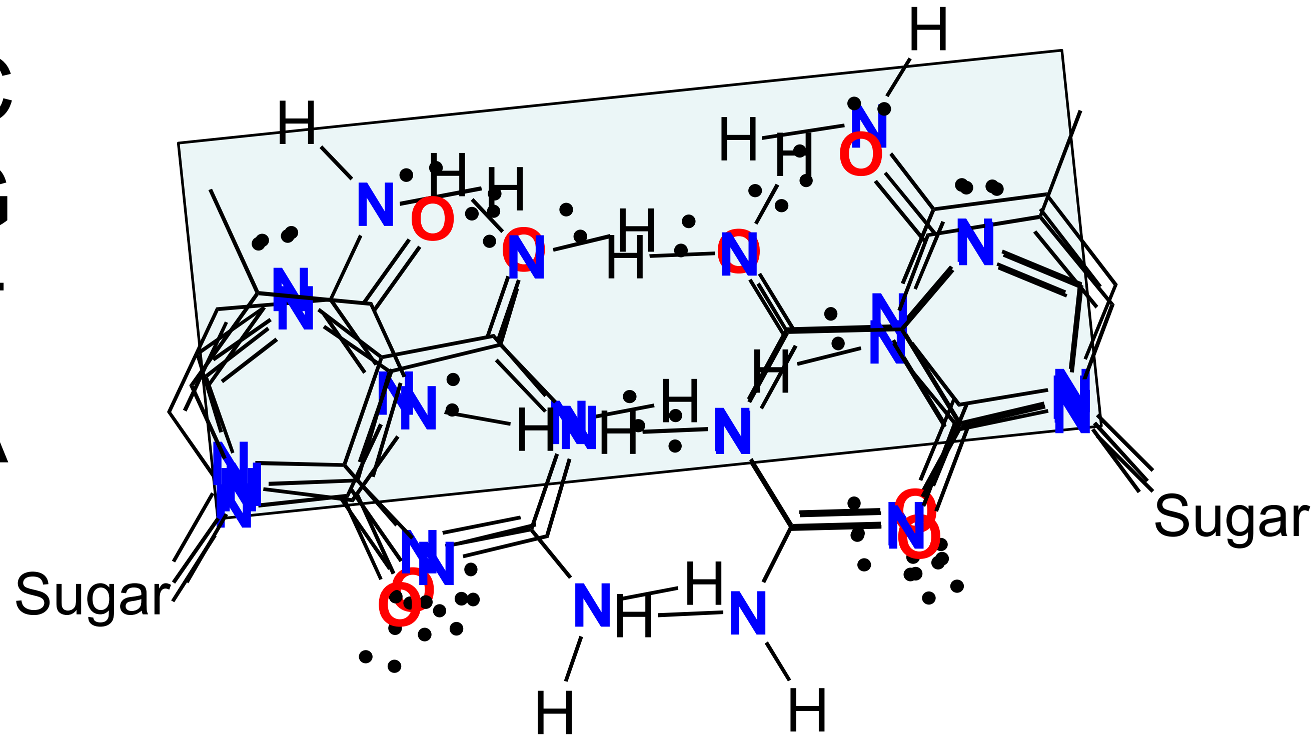




Flat faces are nonpolar
Edges are very polar (can H-bond)

Why is Watson-Crick so good?

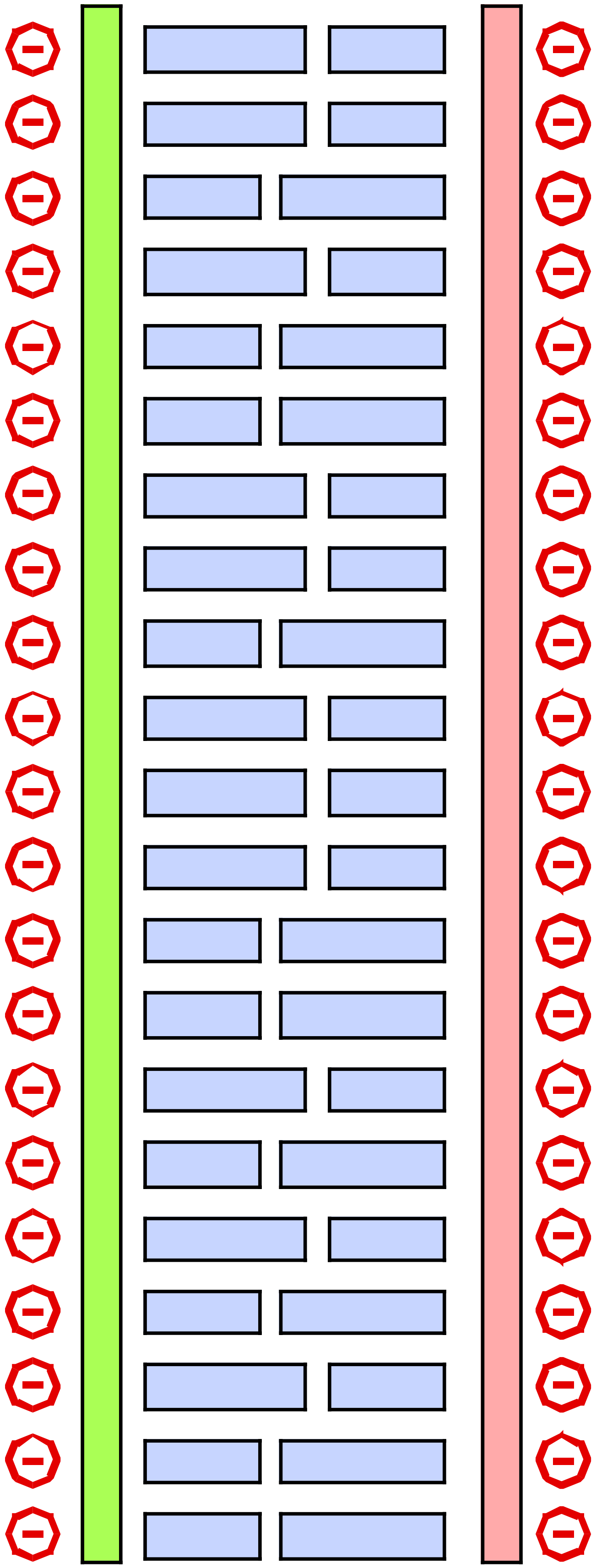
C
G
T
A

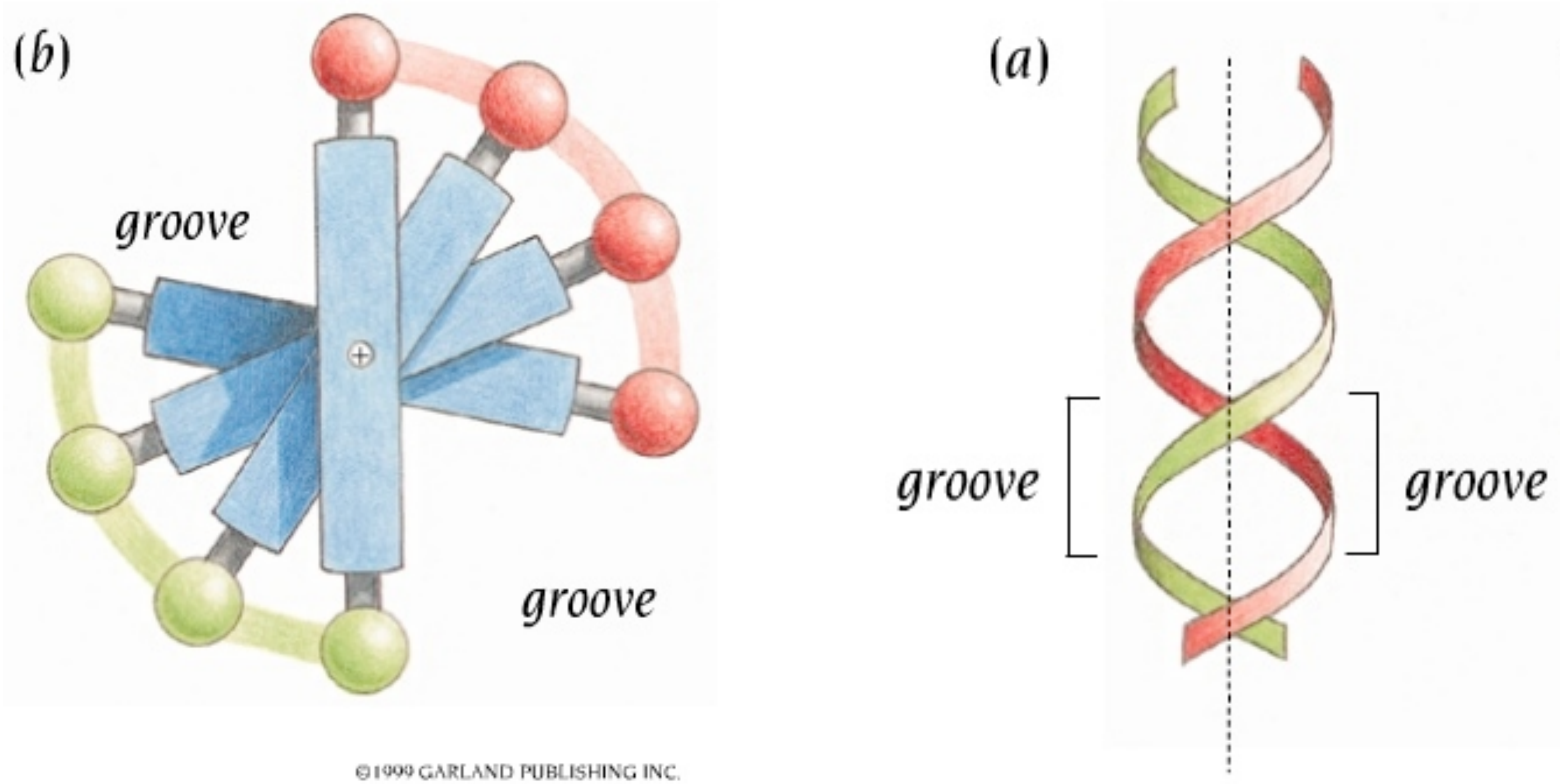


G
C
A
T

All four WC base pairs
are
isosteric

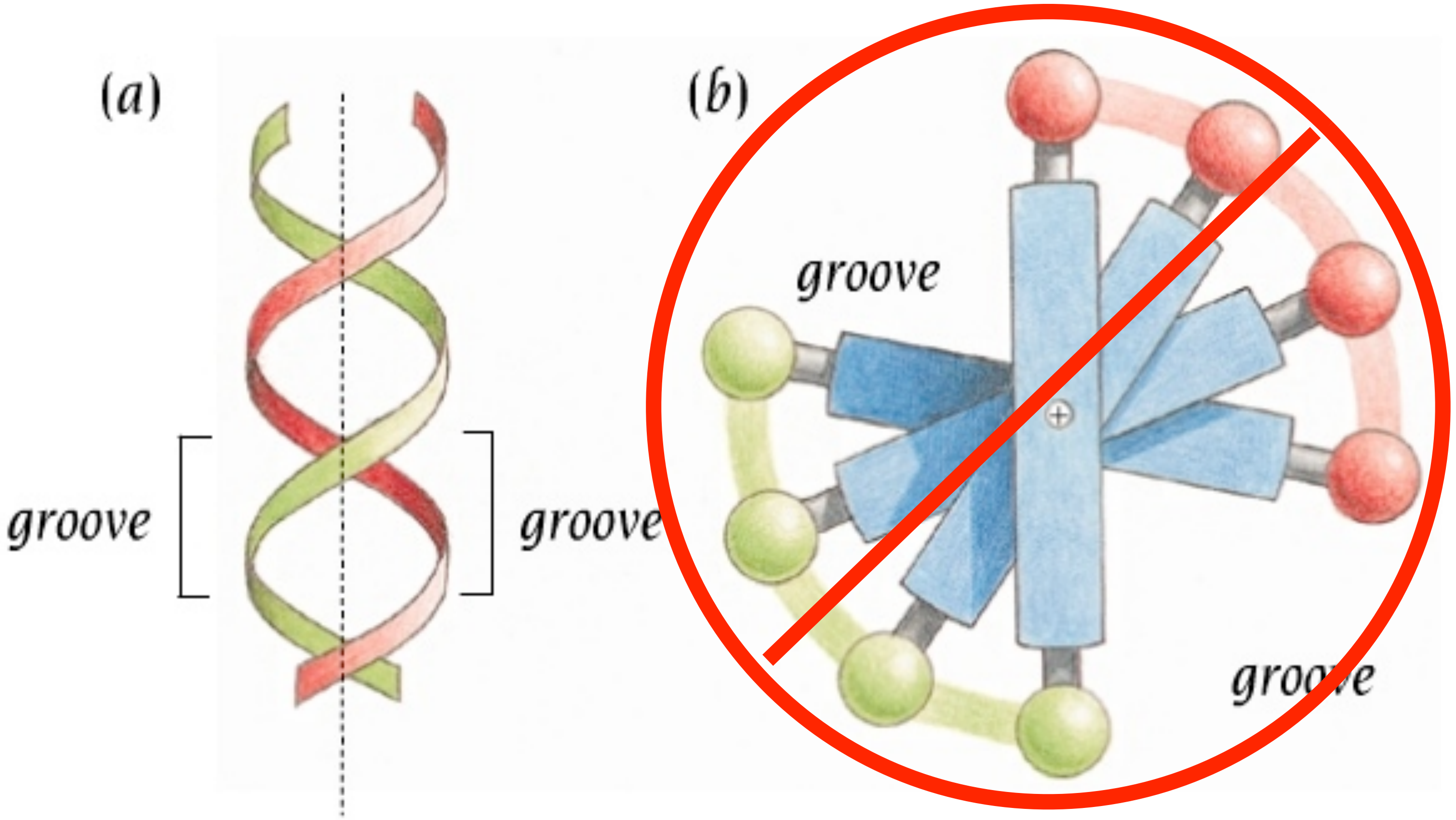
Why *a helix*?



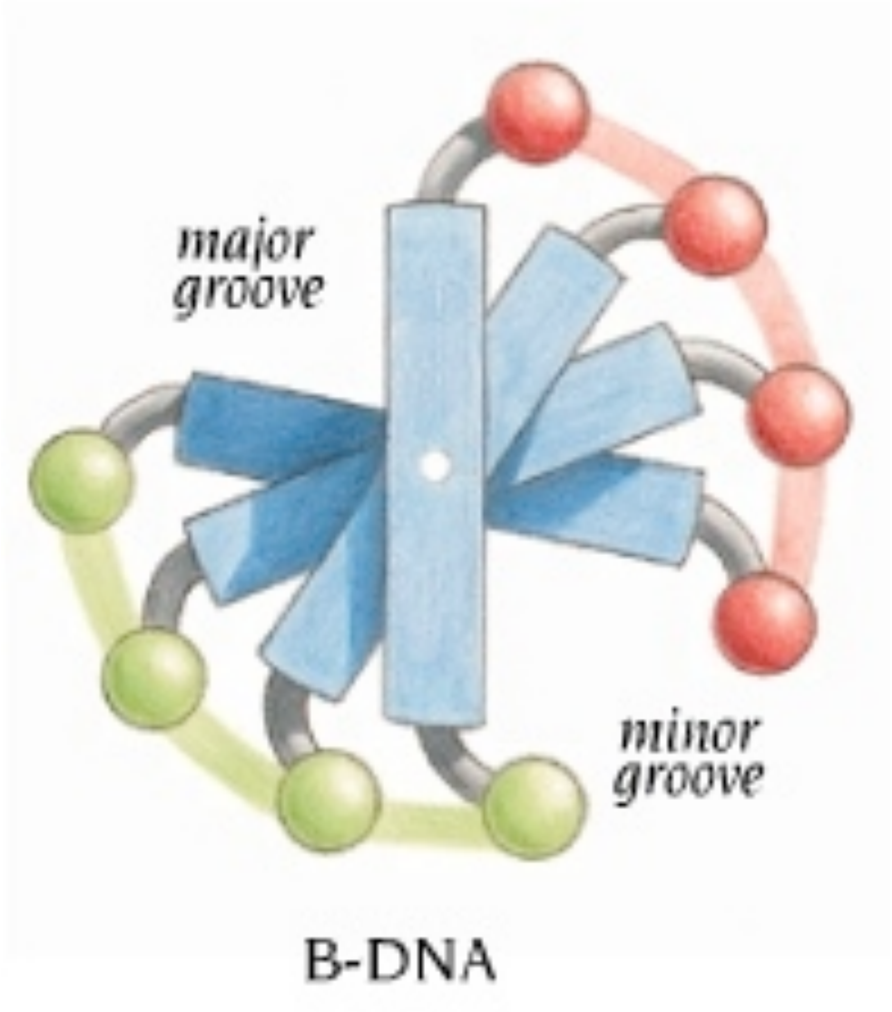


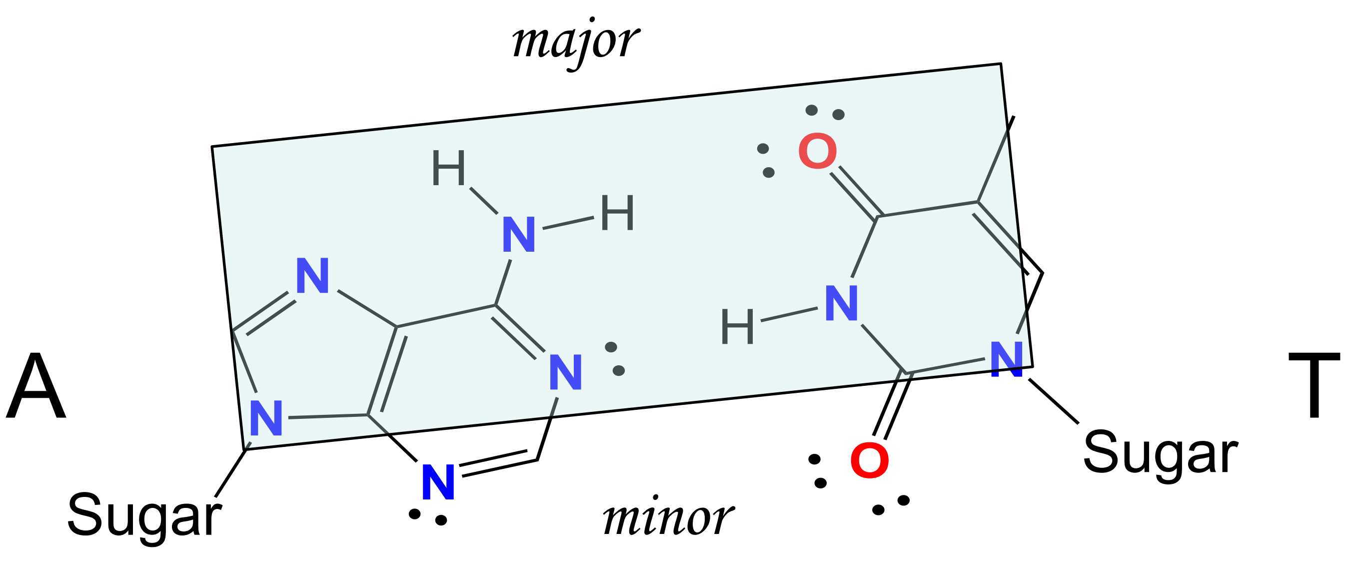
©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group

Why ***major*** and ***minor*** grooves?

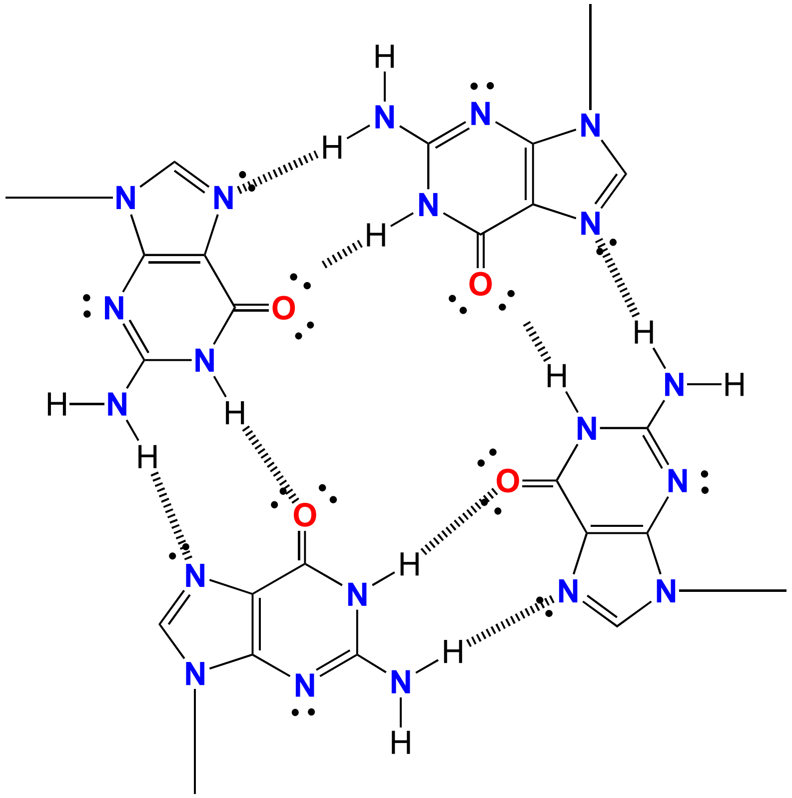


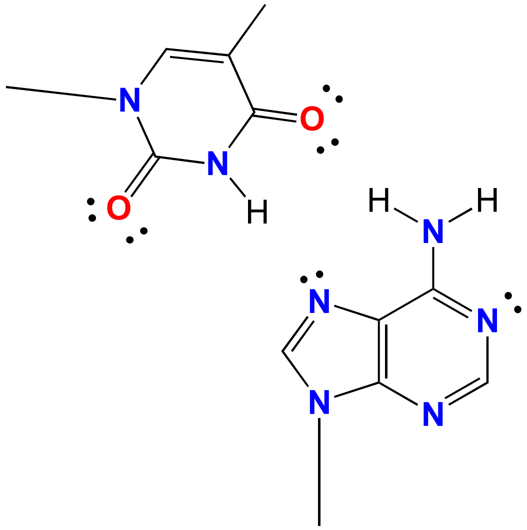
©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group

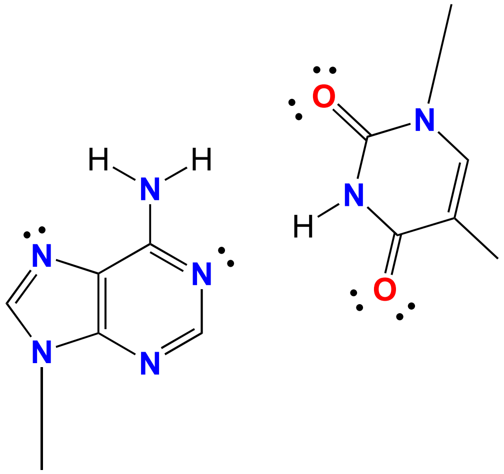


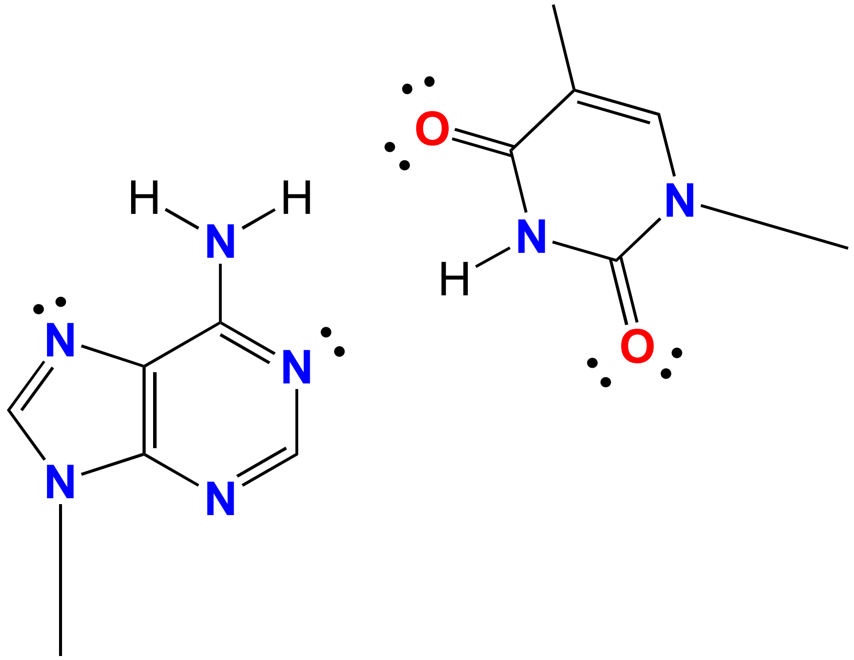


Nucleic Acid - Nucleic Acid Recognition





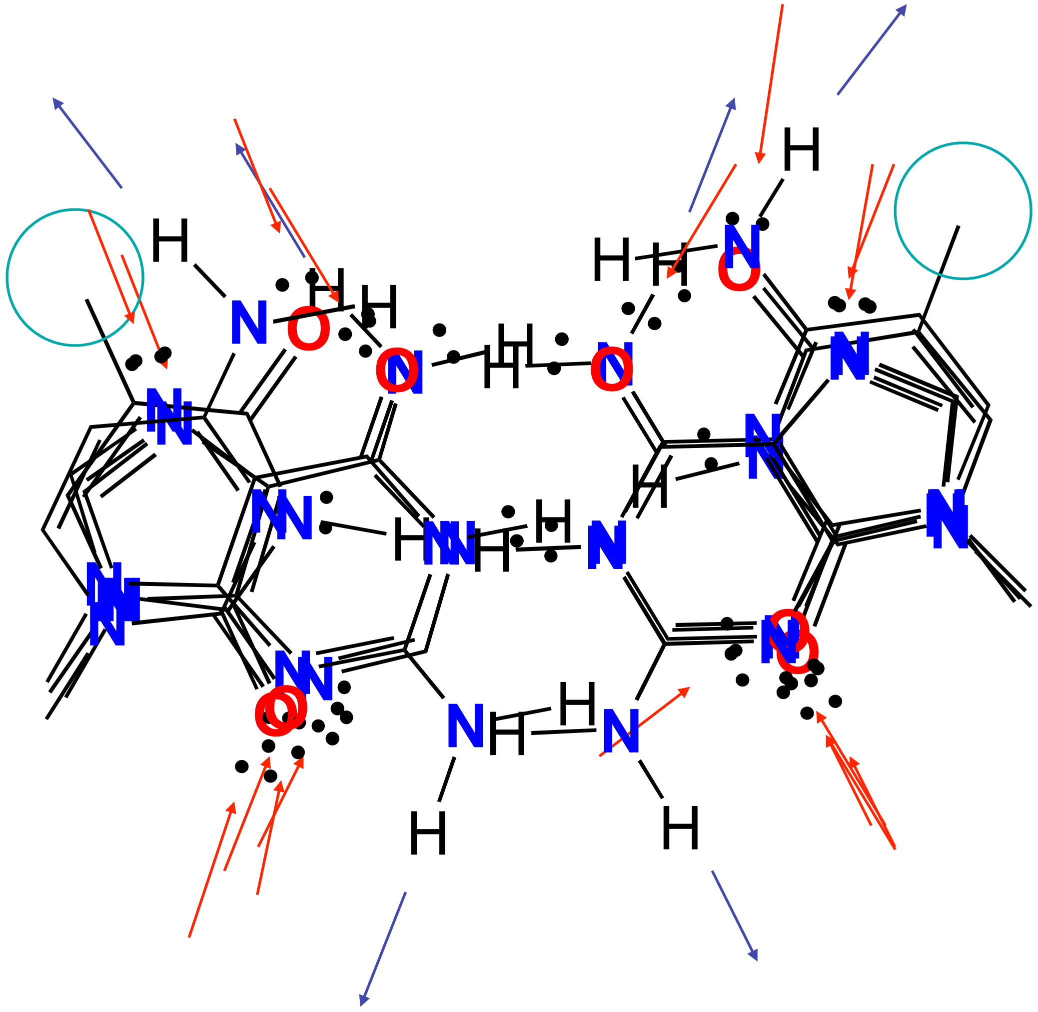




Why is the major groove so good?

Major Groove

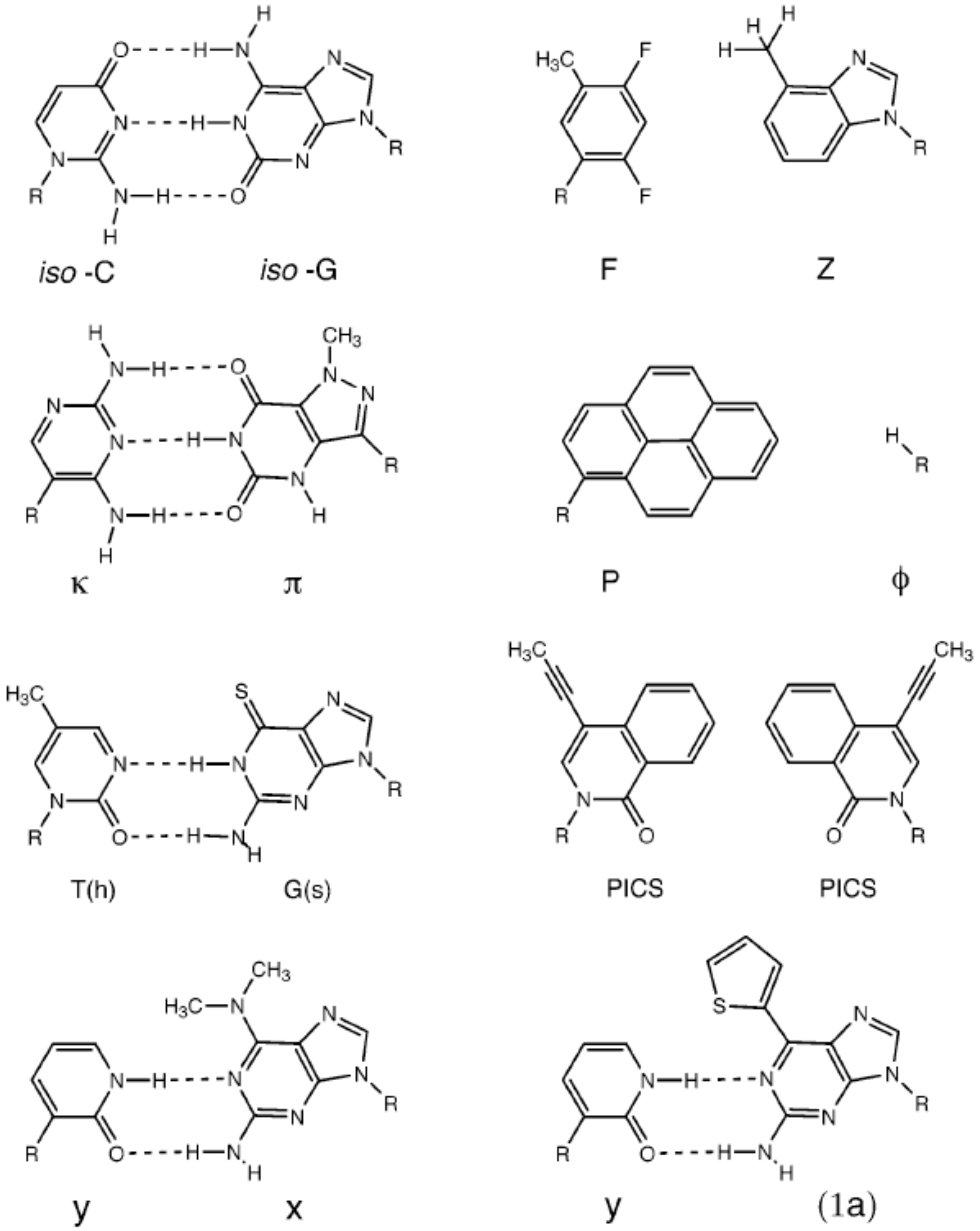
C
G
T
A



G
C
A
T

Minor Groove

Bases that polymerases like



iso -C *iso* -G F Z

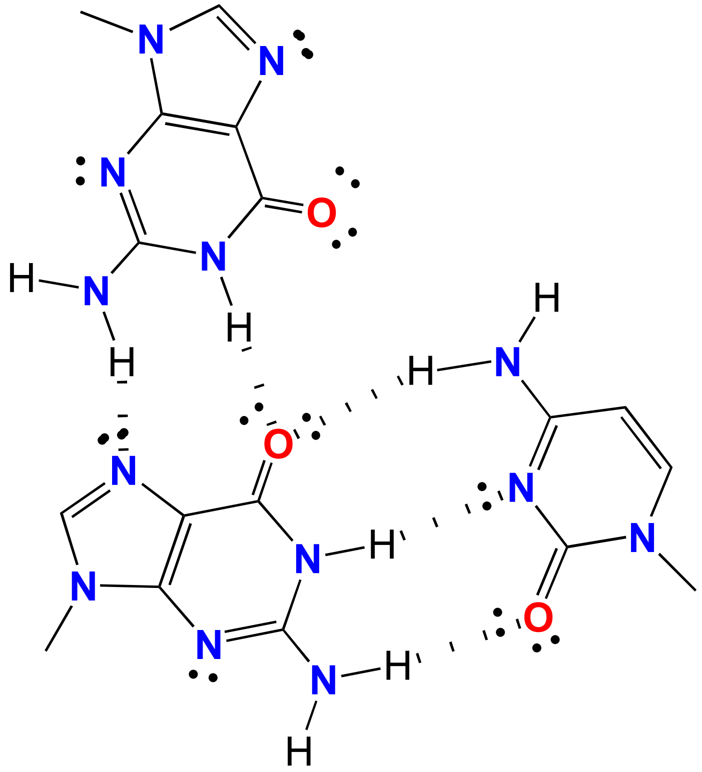
κ π P ϕ

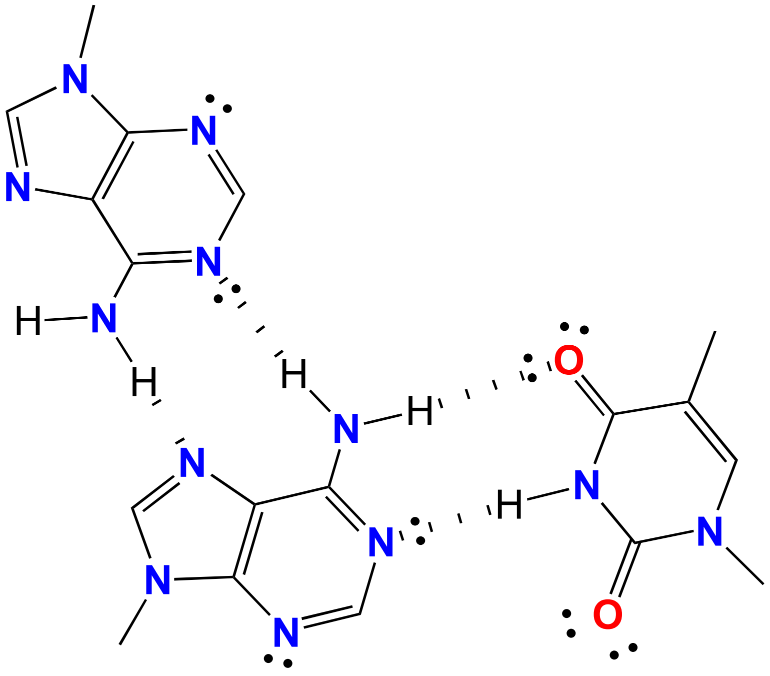
T(h) G(s) PICS PICS

y x y (1a)

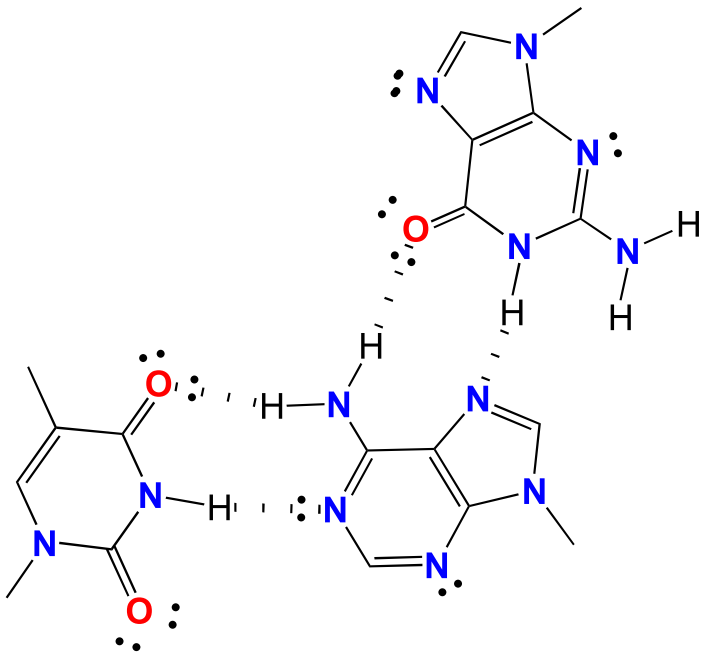
FIGURE 7. Examples of DNA base replacements designed to form stable pairs and/or to be replicated selectively by DNA polymerase enzymes.[17,23,24,49,54,55]

Nucleic Acid “Triples / Platforms”





Major Groove
Interactions

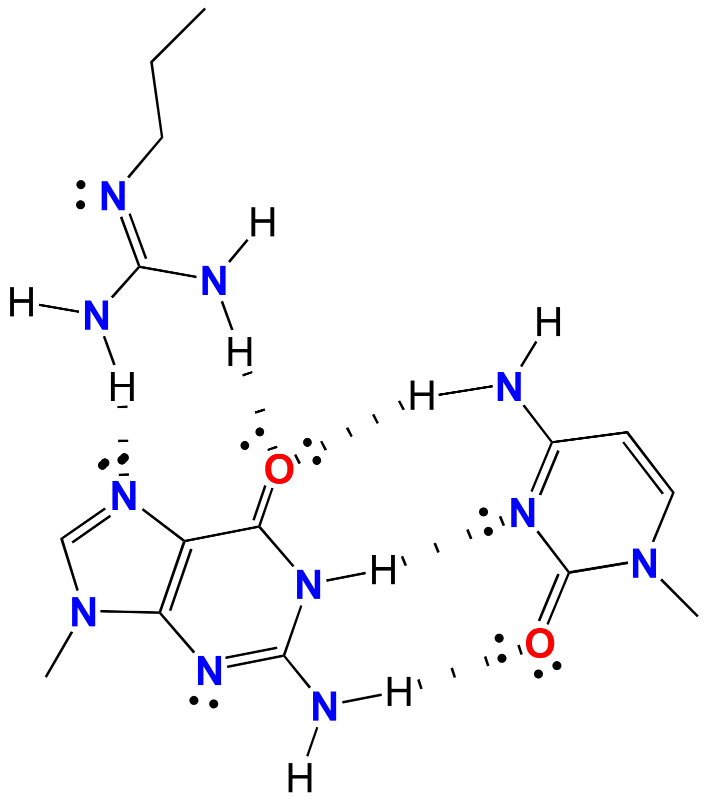


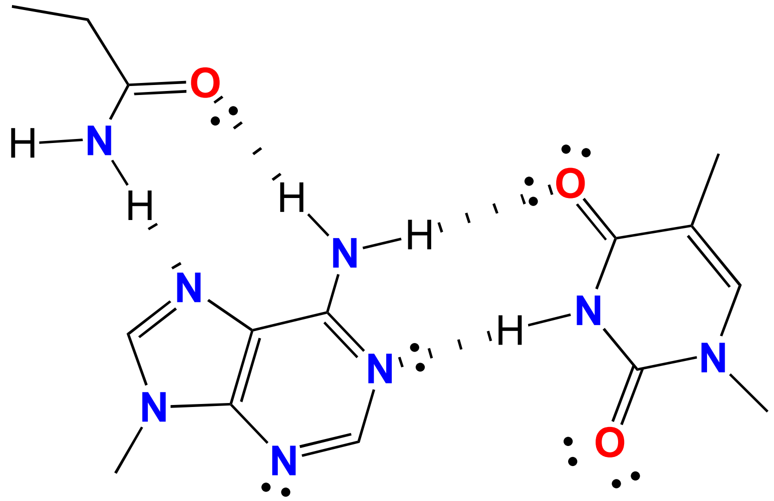
Protein - Nucleic Acid Interactions

Gln

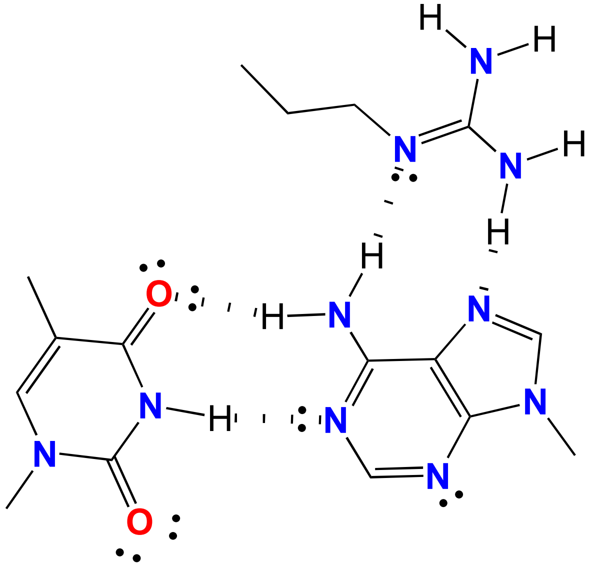
Asn

Arg





Major Groove Interactions



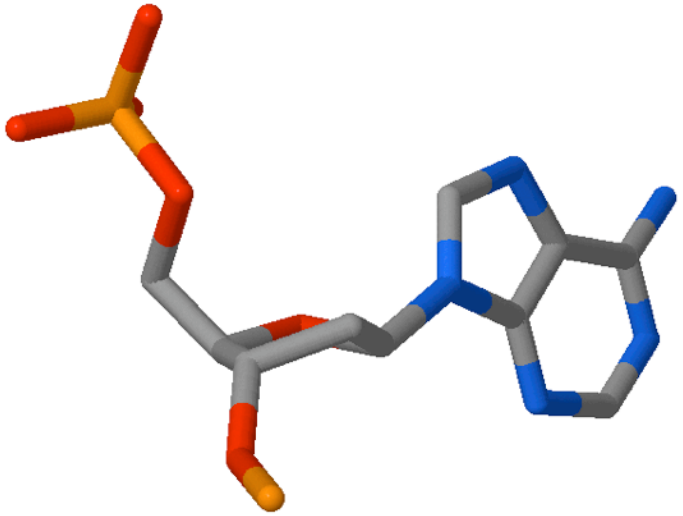
B-form DNA

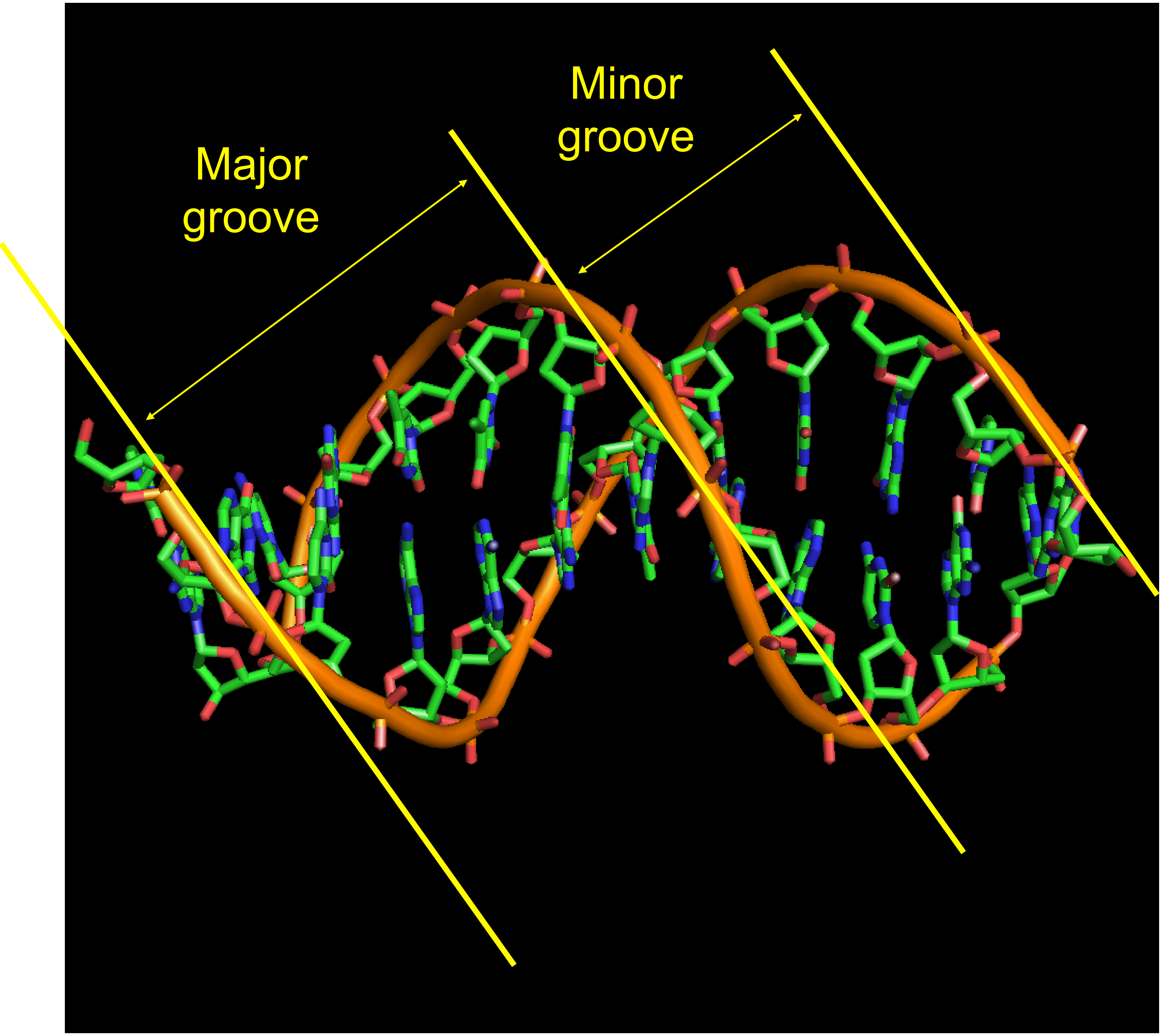
B-form

Residues per turn =10
Twist per base pair = 36°
Rise per pair = 3.4Å
c2’-endo

Minor groove width = 5.7Å
Major groove width = 11.7Å

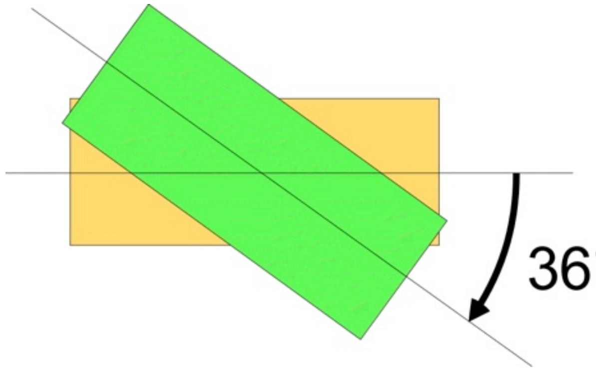
Minor groove depth = 7.5Å
Major groove depth = 8.8Å





B-form DNA

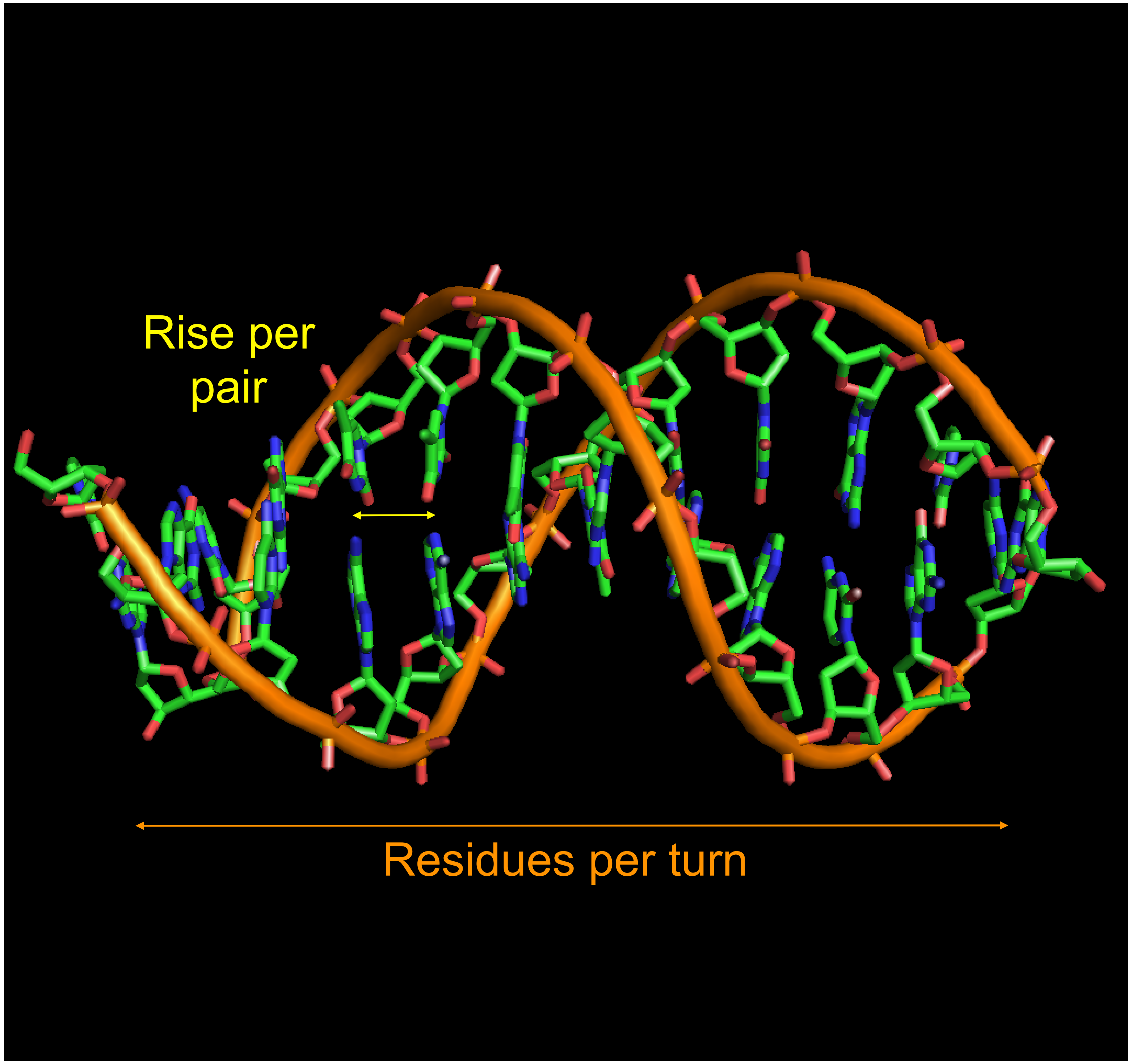
Residues per turn =10
Twist per base pair = 36°



Rise per pair = 3.4Å
c2’-endo

Minor groove width = 5.7Å
Major groove width = 11.7Å

Minor groove depth = 7.5Å
Major groove depth = 8.8Å

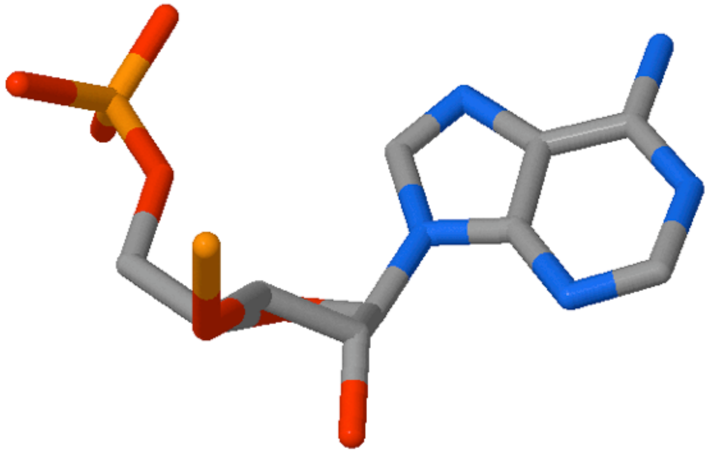


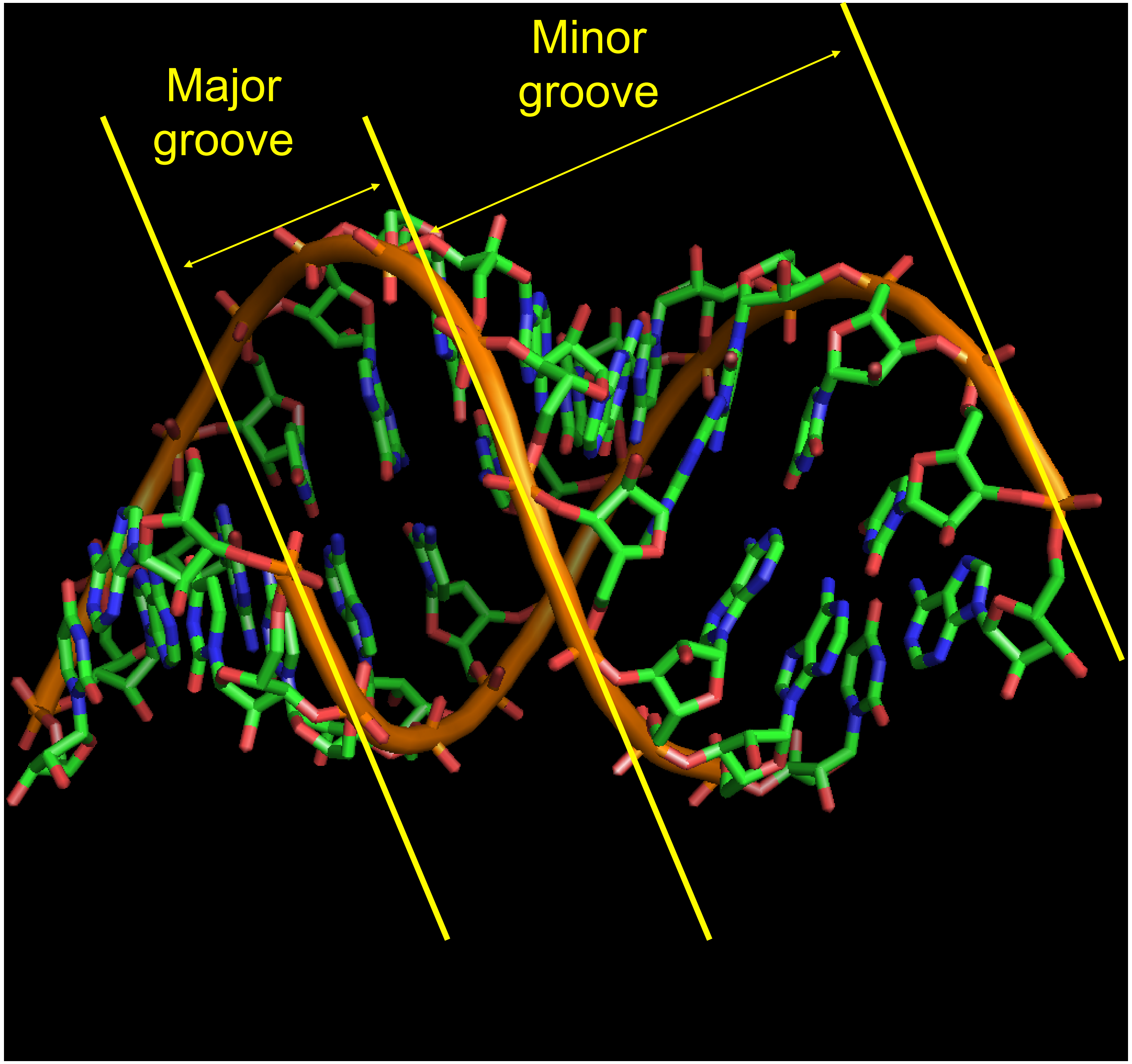
A-form RNA

Residues per turn =11
Twist per base pair = 33°
Rise per pair = 2.9Å
c3’-endo

Minor groove width = 11Å
Major groove width = 2.7Å

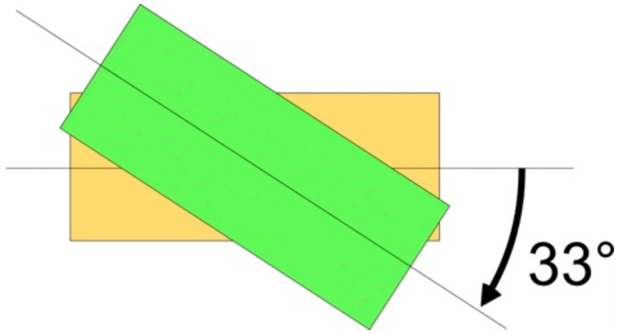
Minor groove depth = 2.8Å
Major groove depth = 13.5Å





A-form RNA

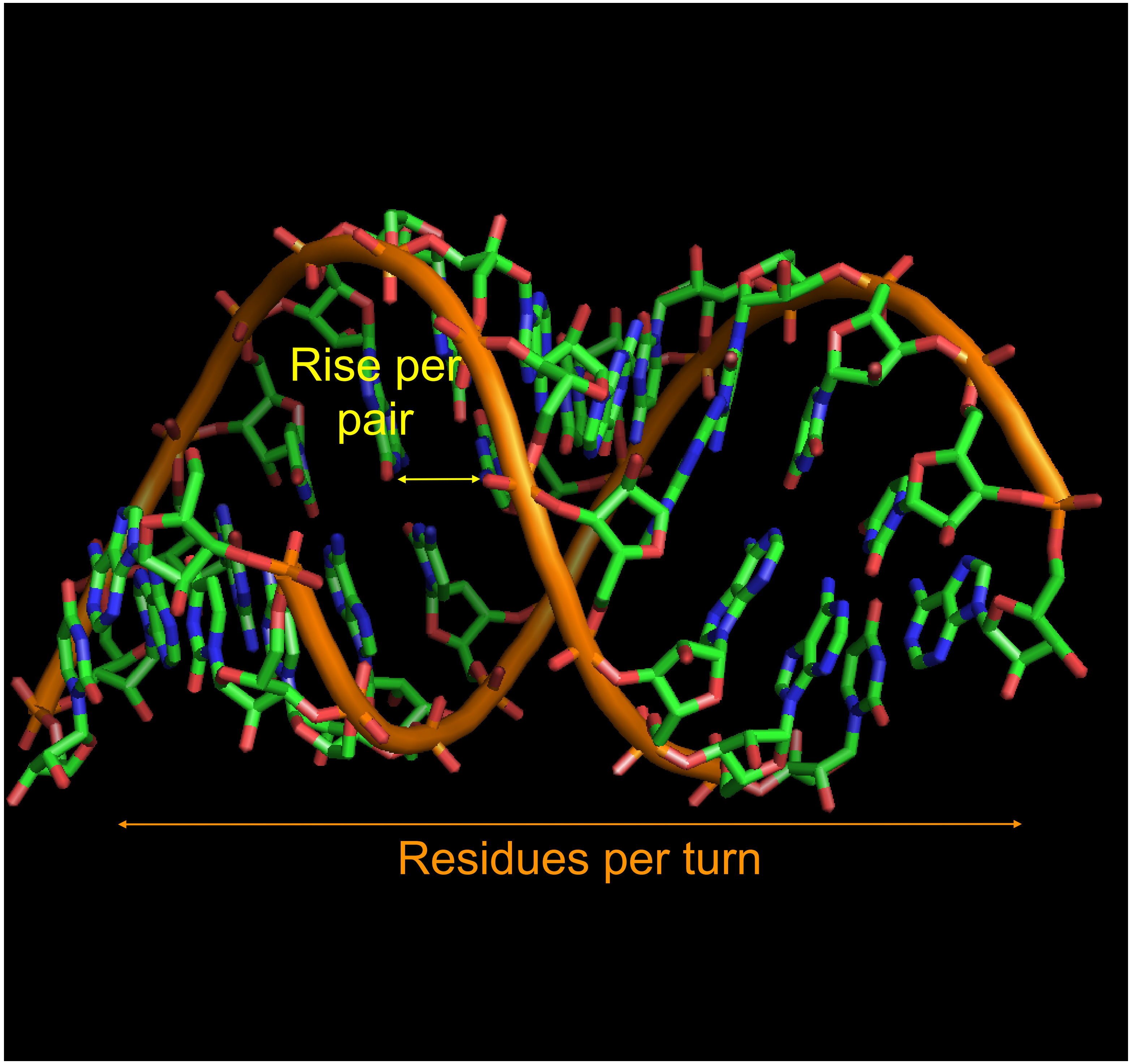
Residues per turn =11
Twist per base pair = 33°



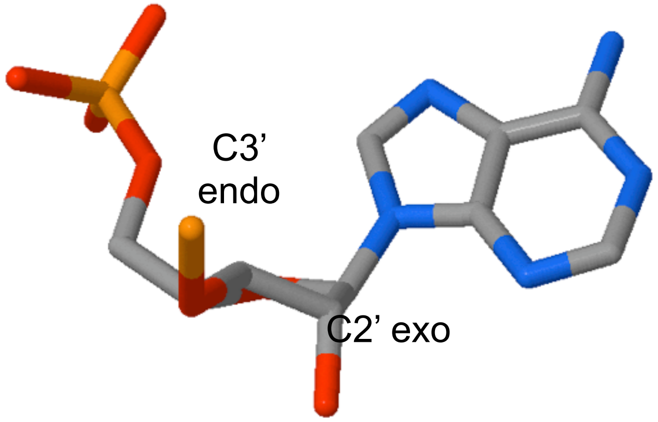
Rise per pair = 2.9Å
c3’-endo

Minor groove width = 11Å
Major groove width = 2.7Å

Minor groove depth = 2.8Å
Major groove depth = 13.5Å



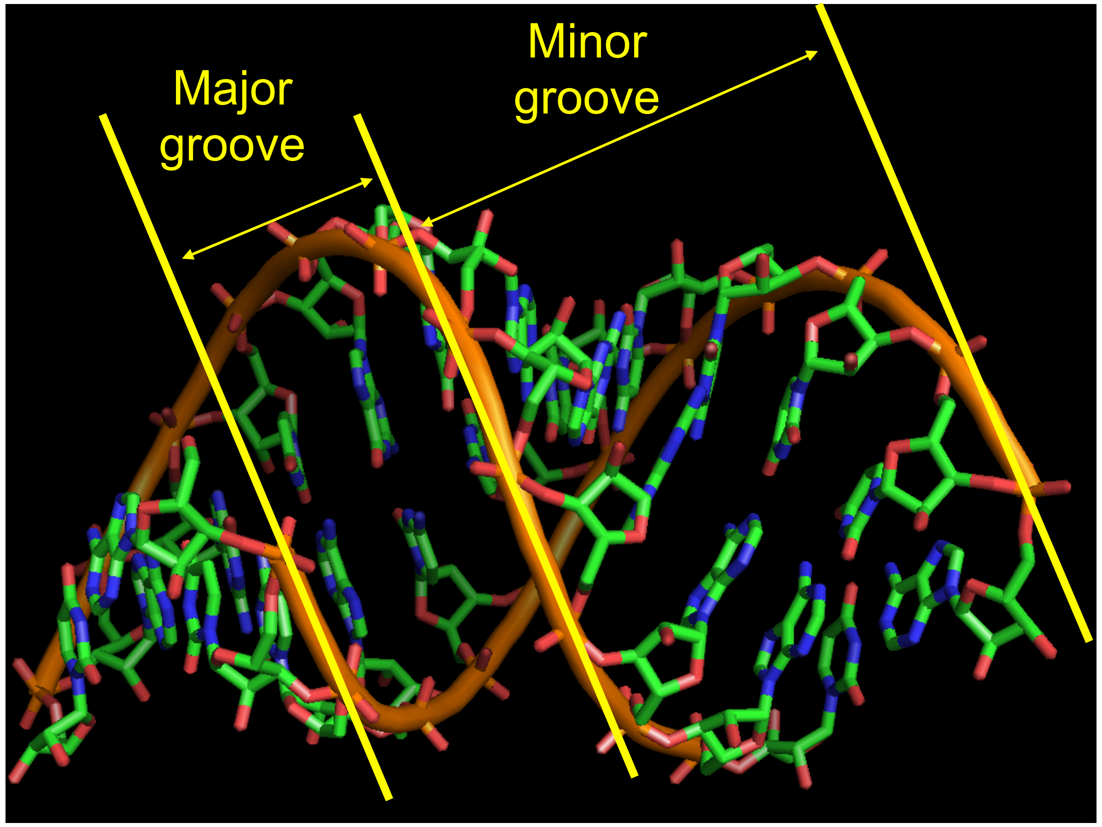
Compare

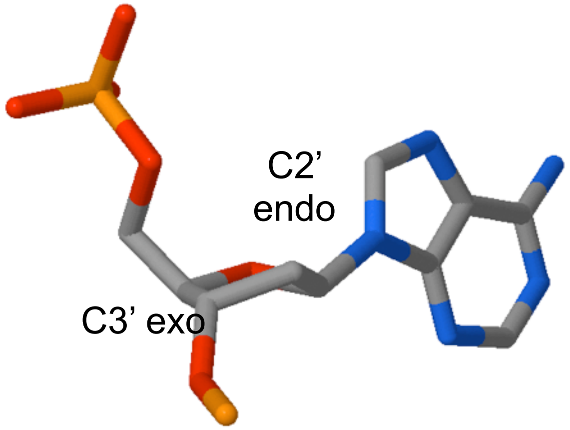


A-form (RNA)

Minor groove width = 11Å
Major groove width = 2.7Å

Minor groove depth = 2.8Å
Major groove depth = 13.5Å

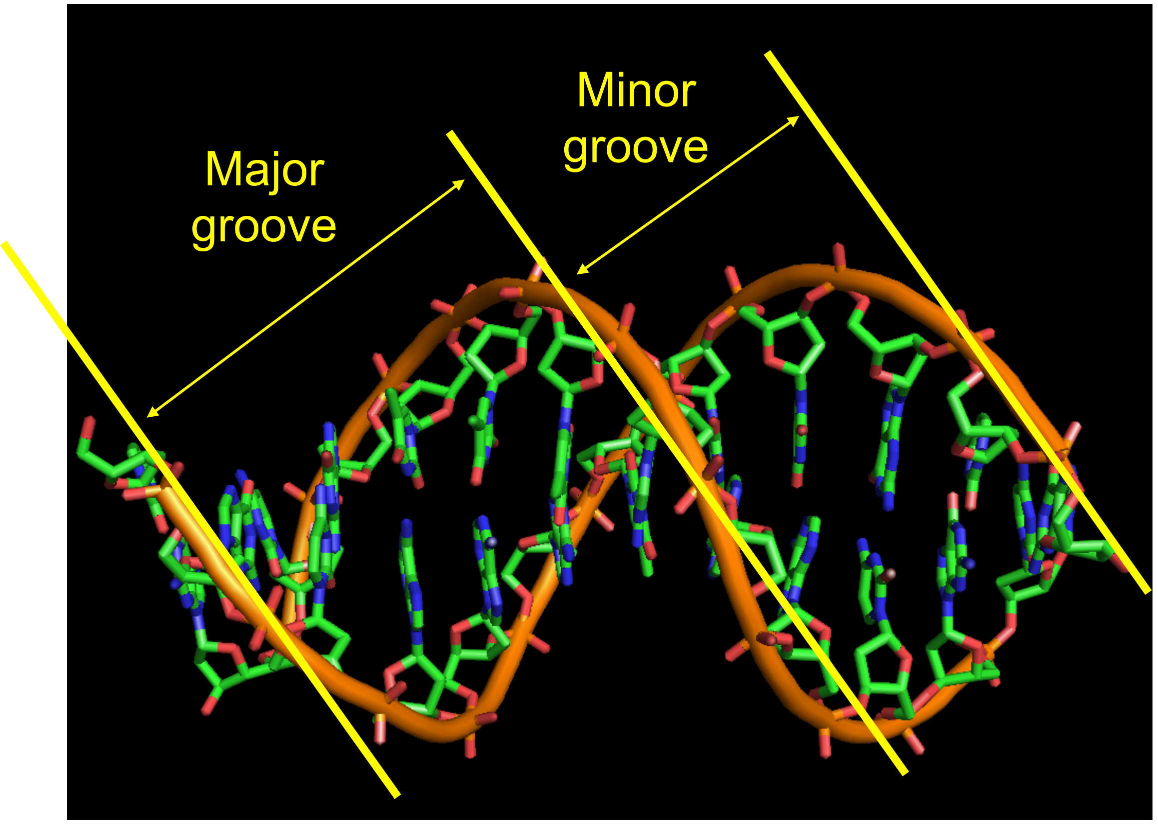




B-form (DNA)

Minor groove width = 5.7Å
Major groove width = 11.7Å

Minor groove depth = 7.5Å
Major groove depth = 8.8Å



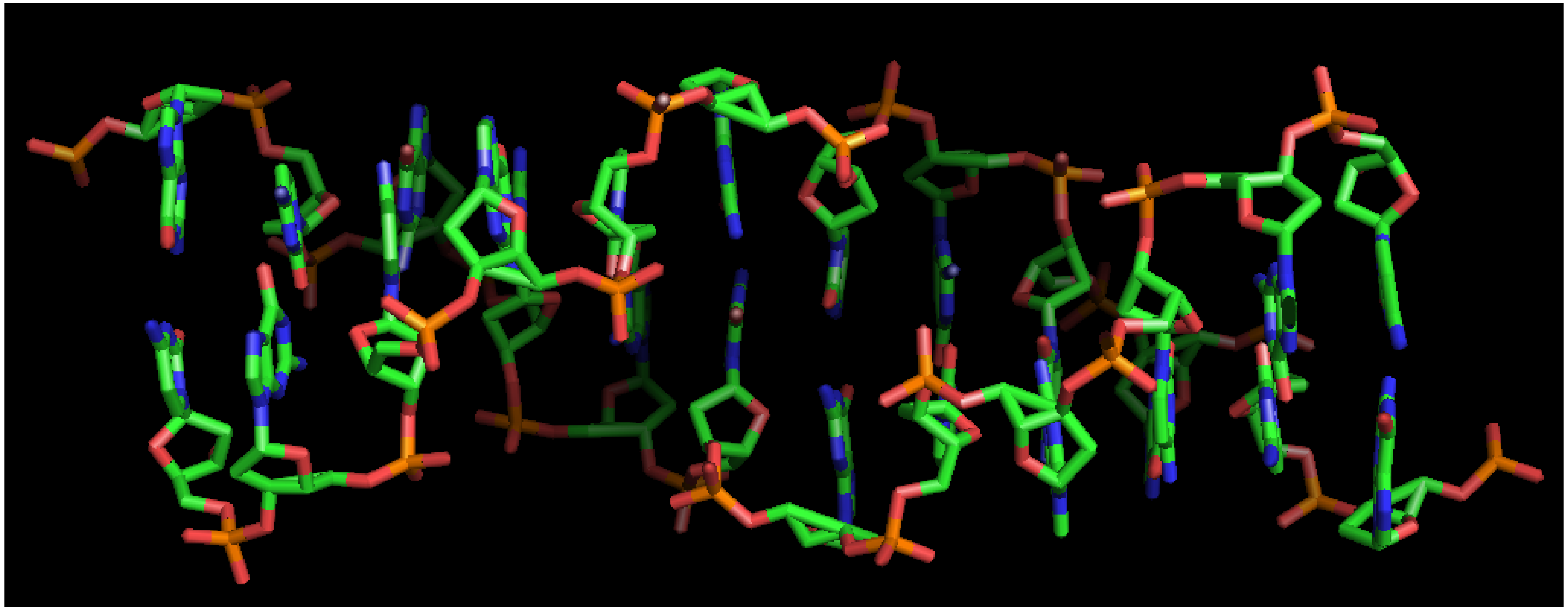
Z-DNA

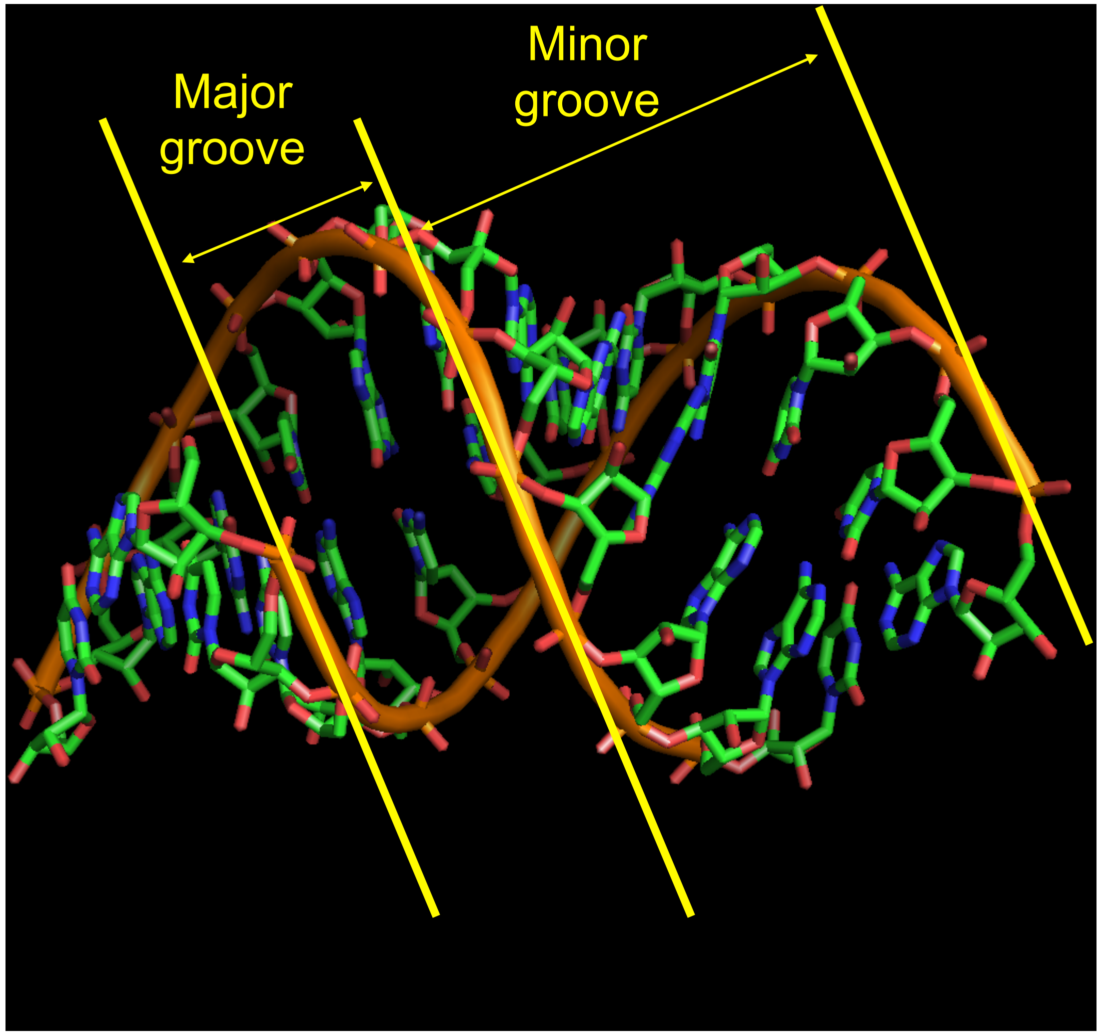
Residues per turn =12
Twist per base pair = -9 / -51°

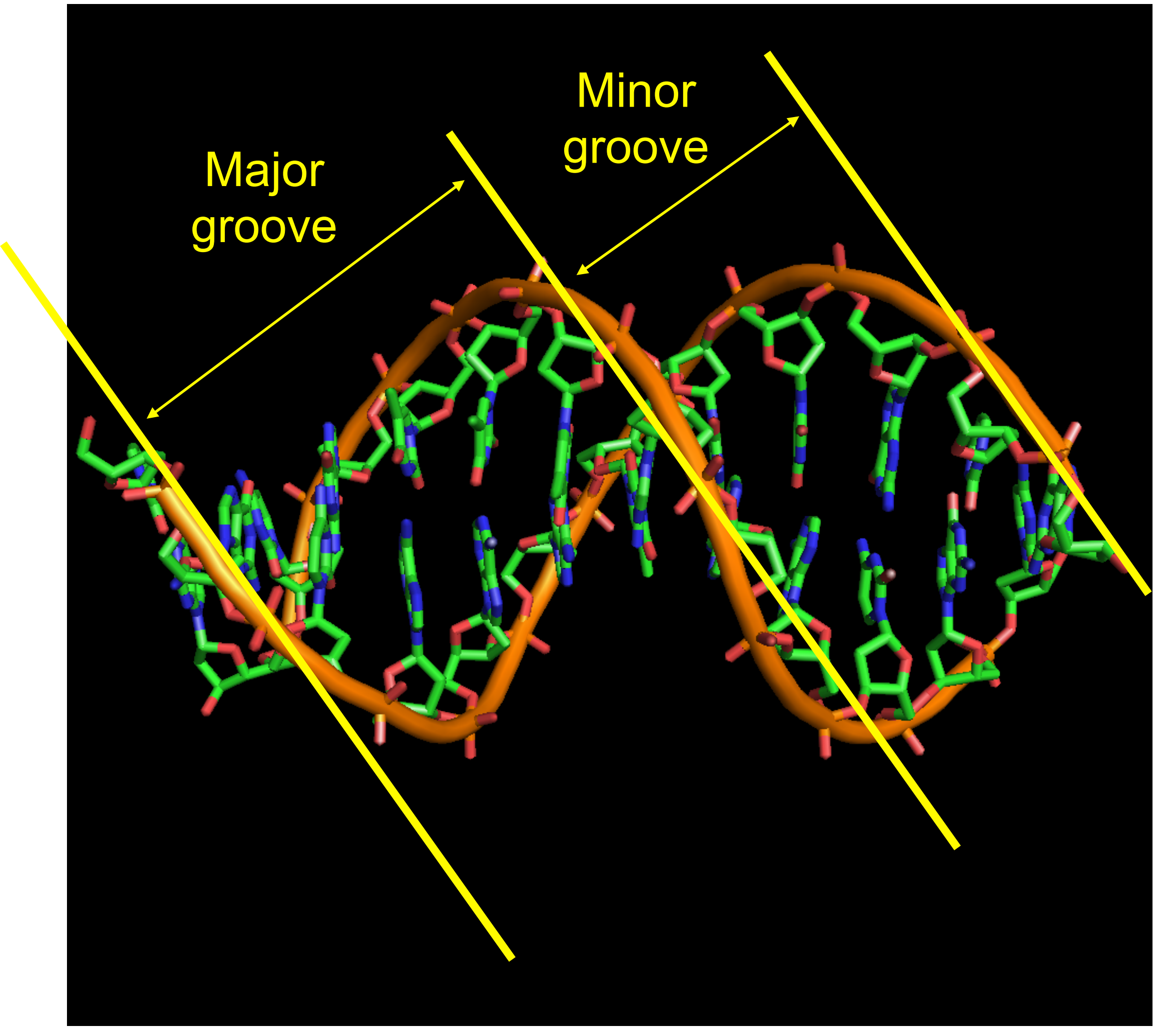
Rise per pair = 3.7Å
c3’-endo(syn) / c2’-endo

Minor groove width = 2.0Å
Major groove width = 8.8Å

Minor groove depth = 13.8Å
Major groove depth = 3.7Å

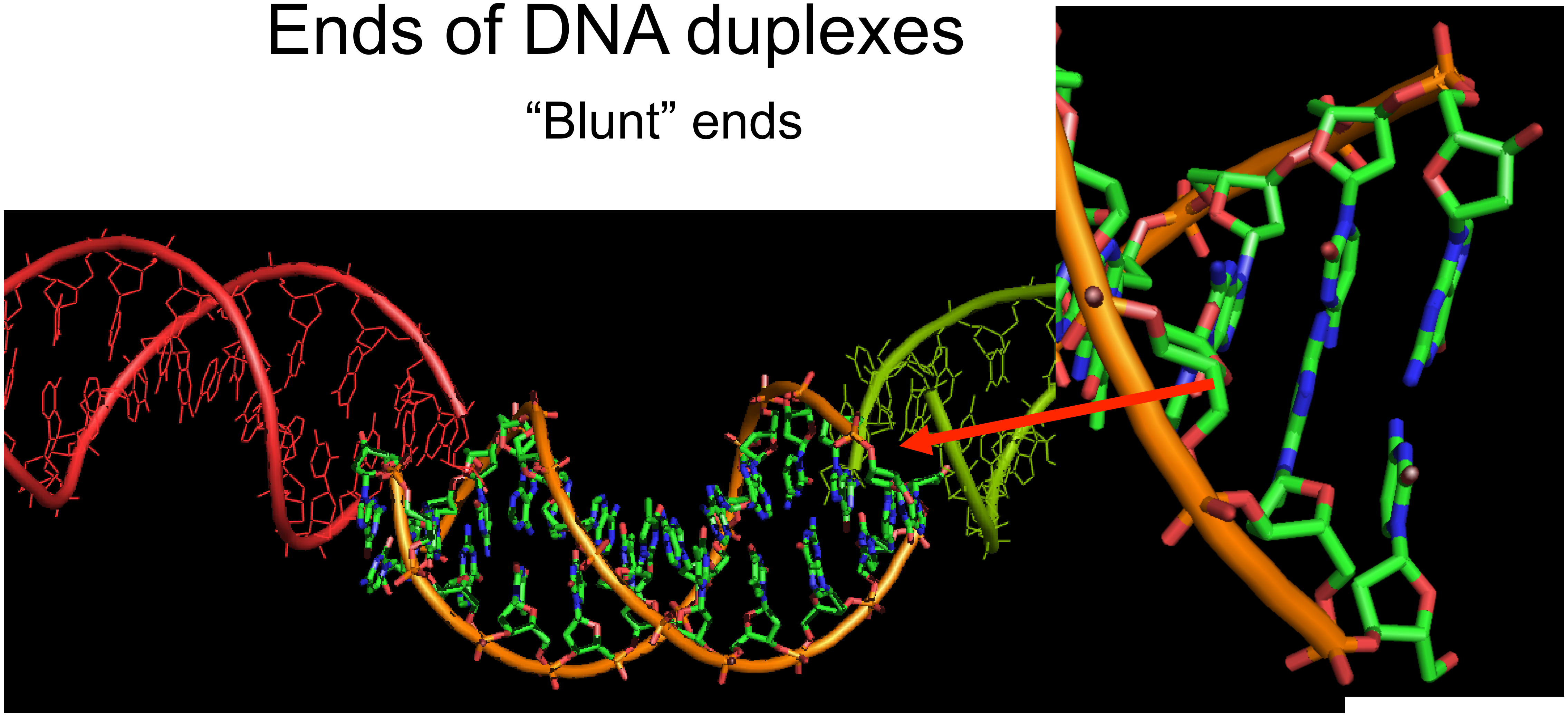


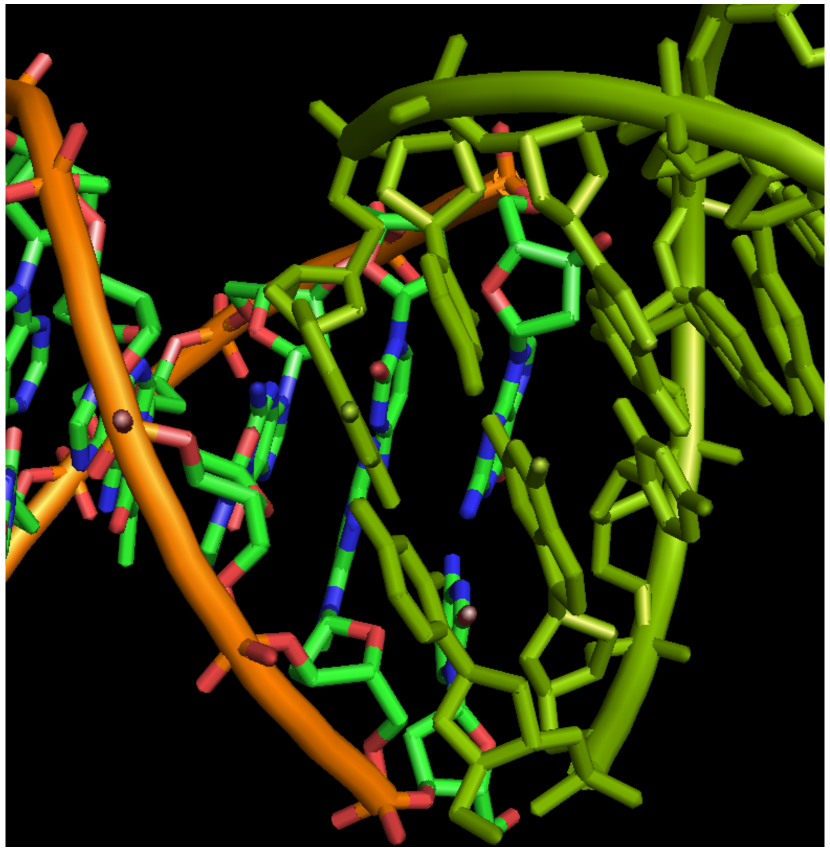




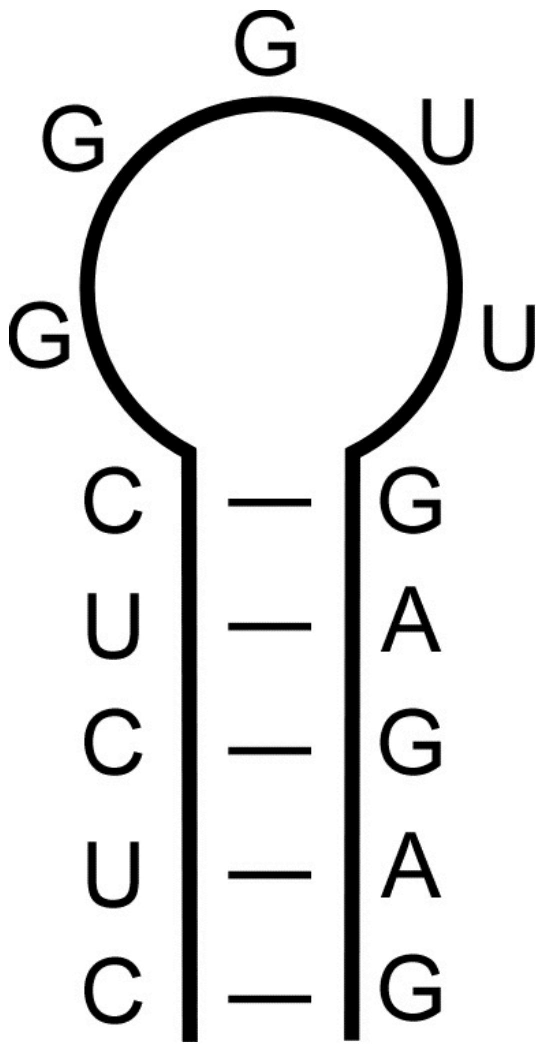
Ends of DNA duplexes

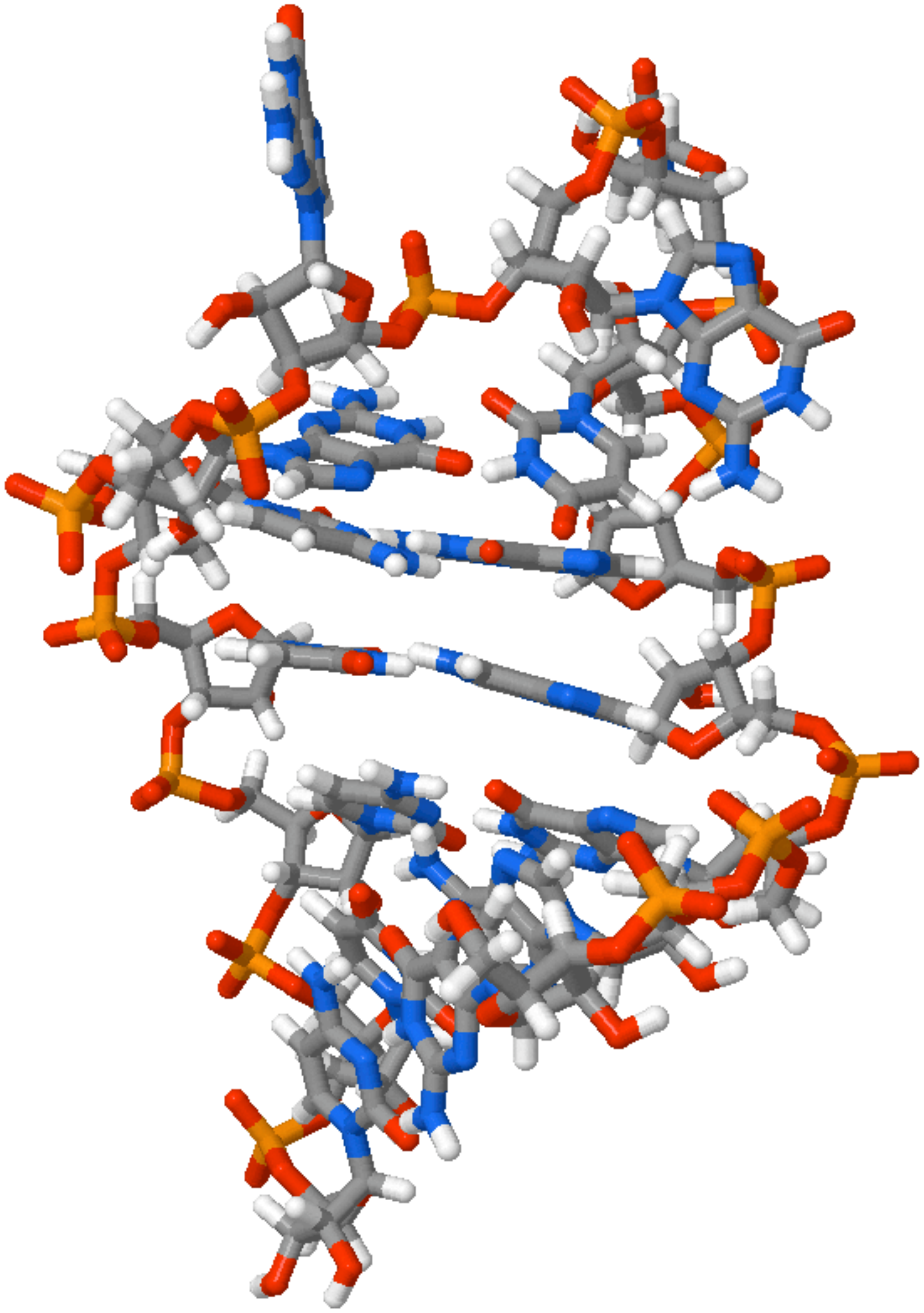
“Blunt” ends



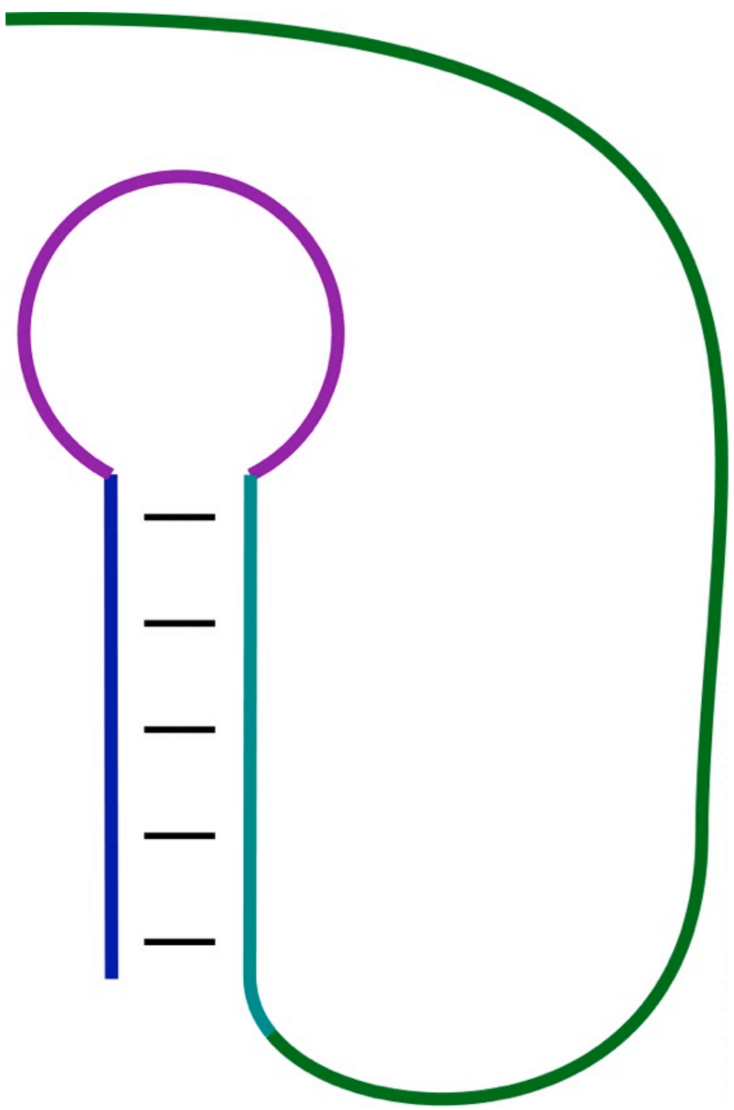


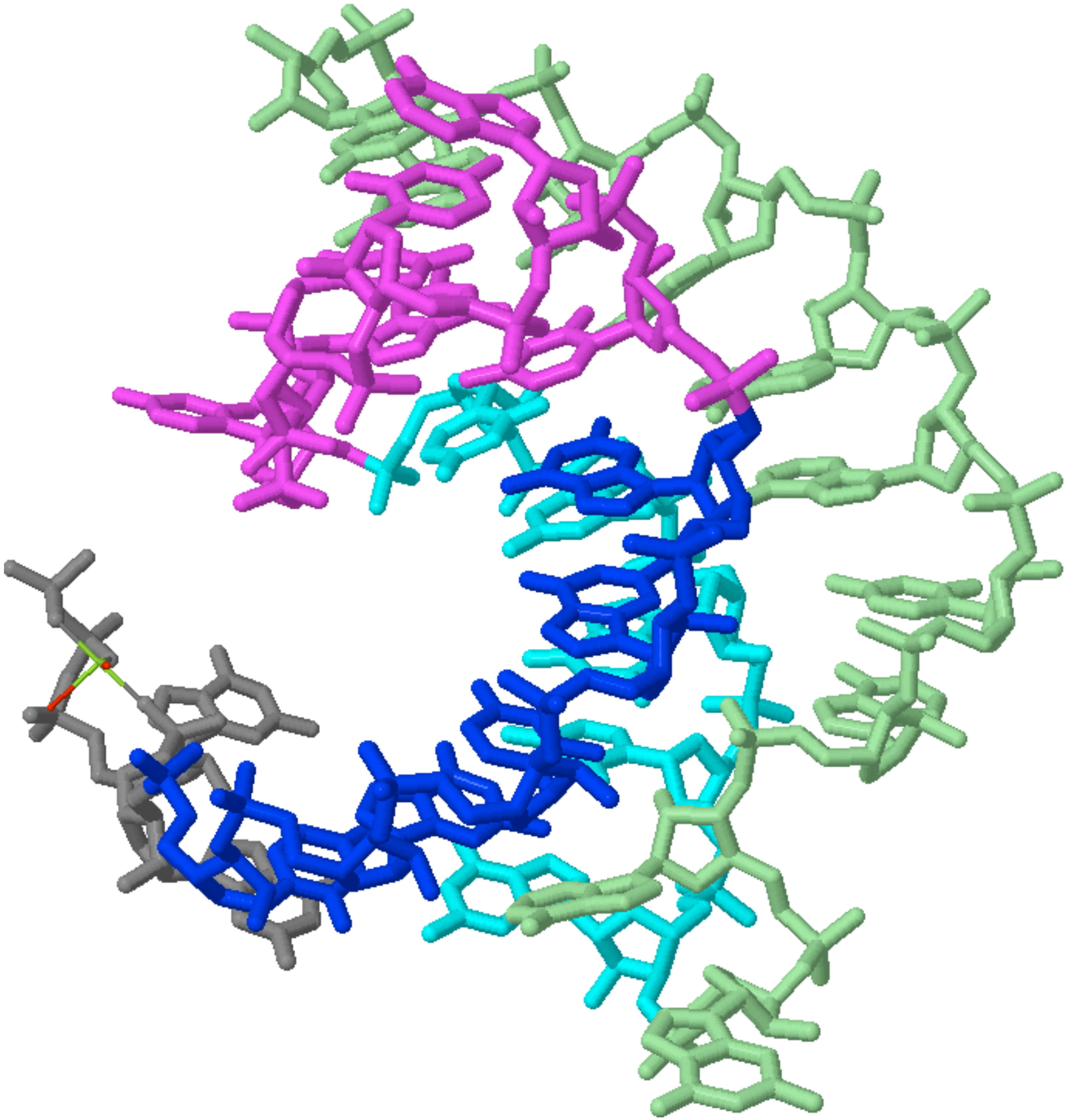
Simple Structure - Hairpin



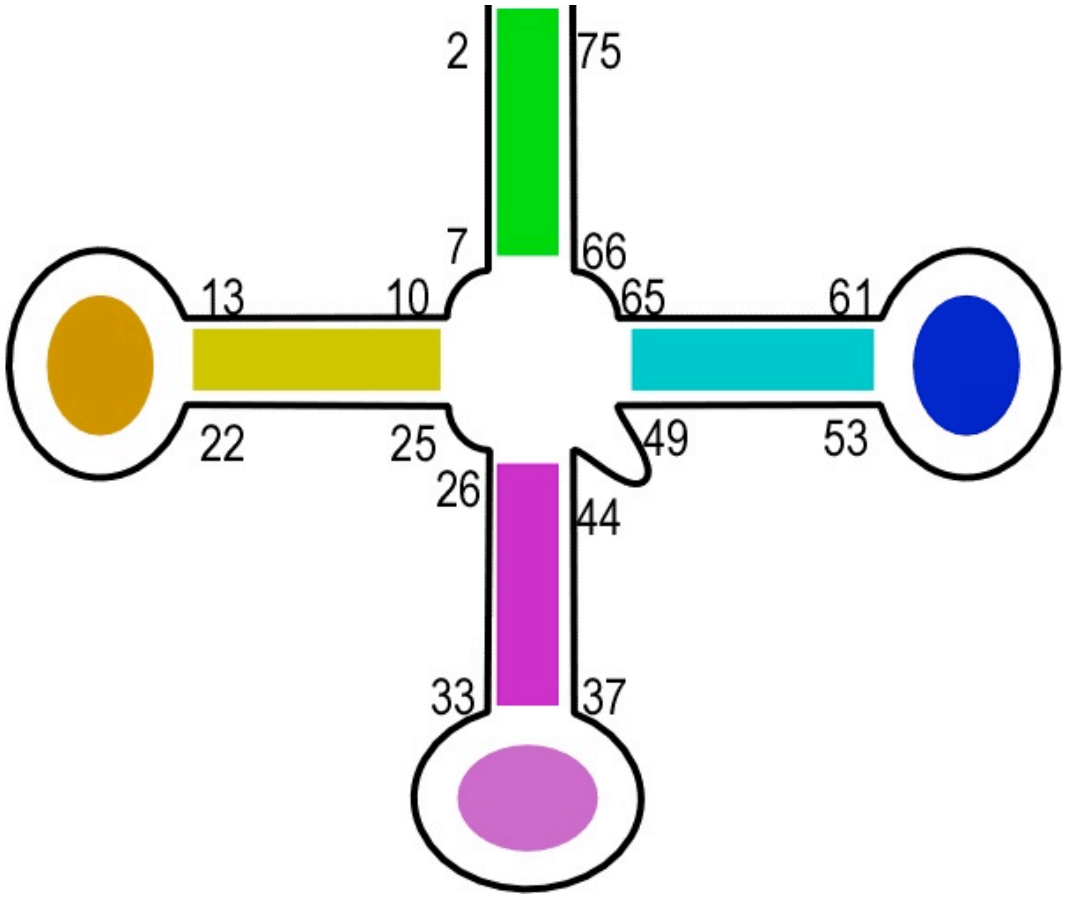


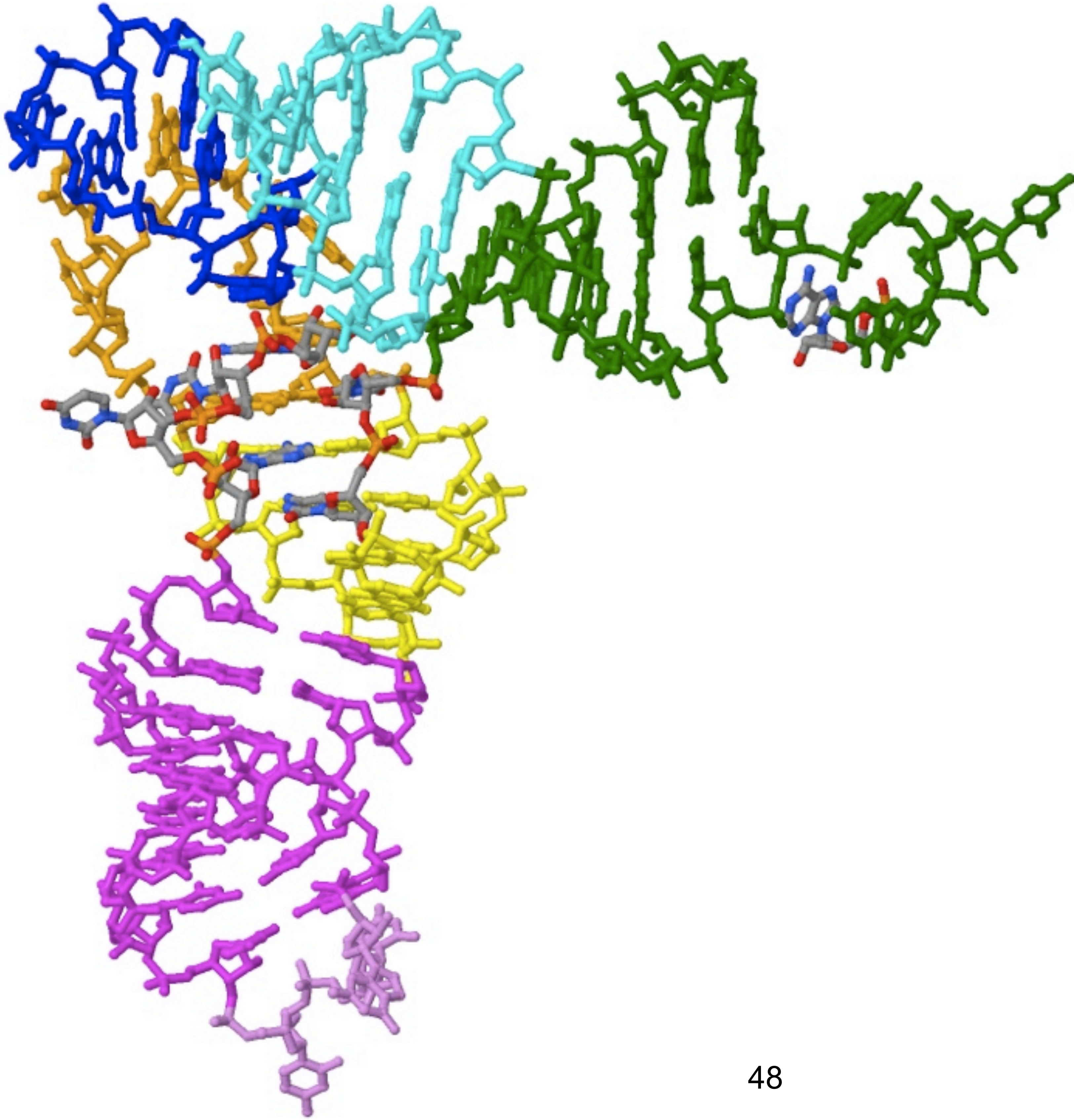
Classic Structure - Pseudoknot



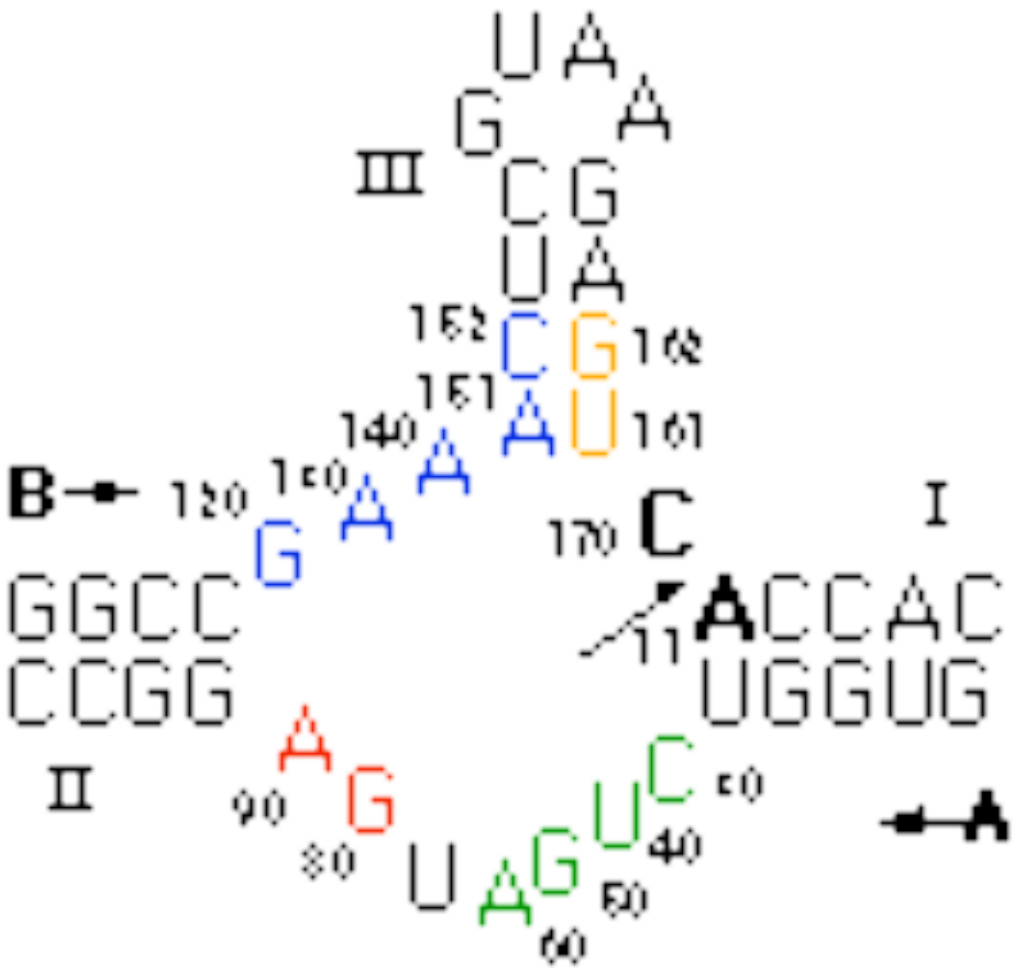


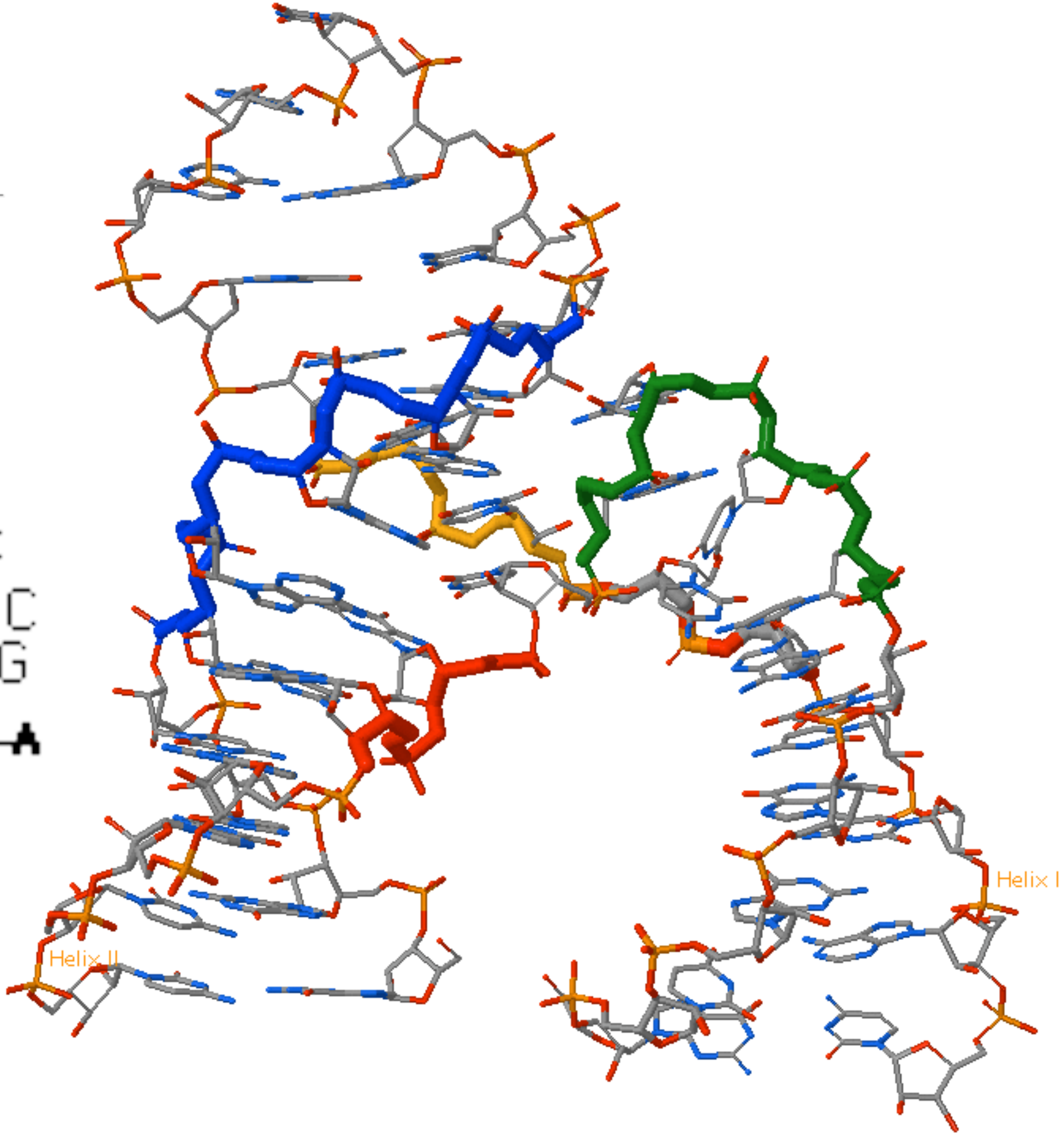
tRNA



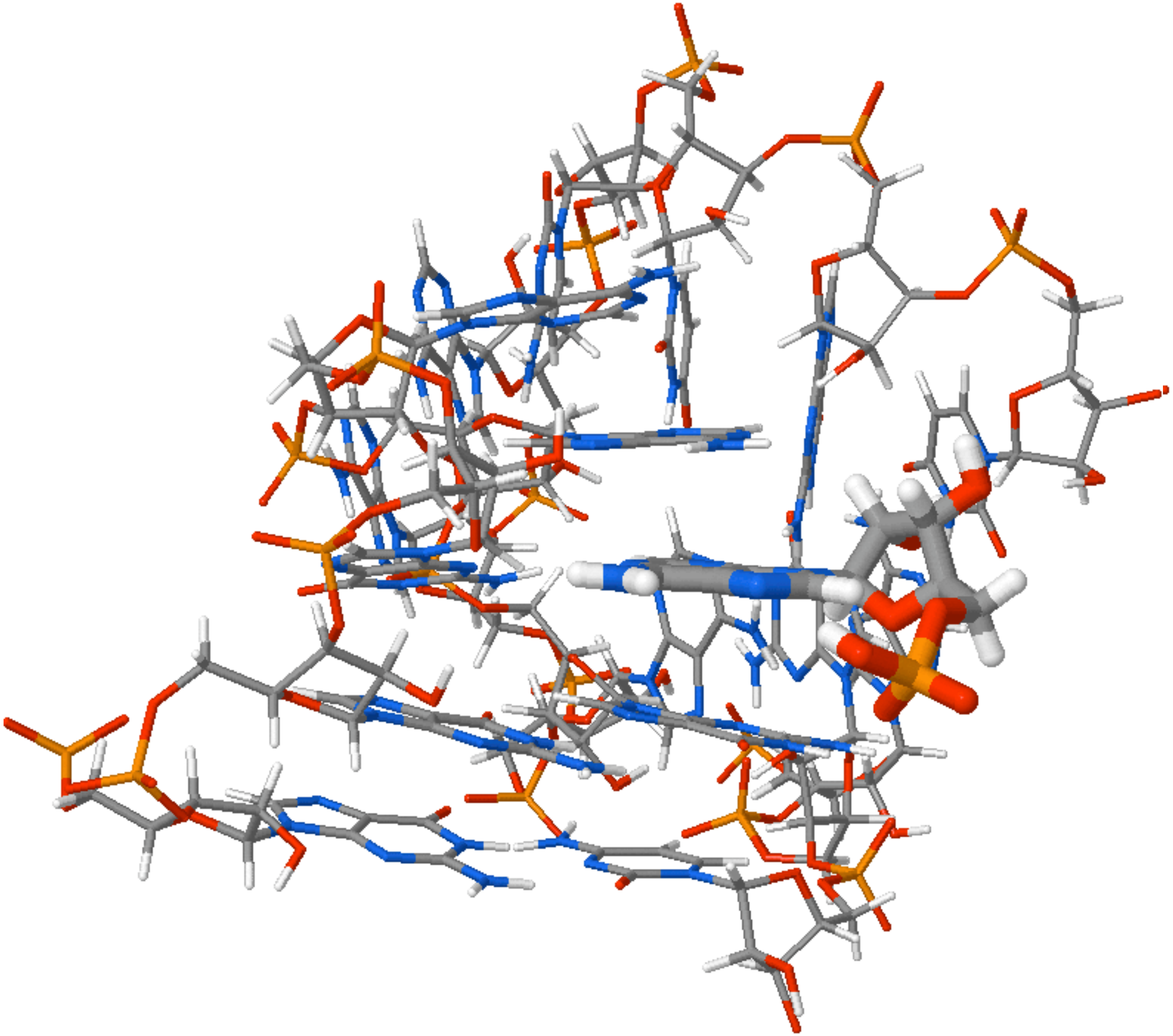


Hammerhead Ribozyme





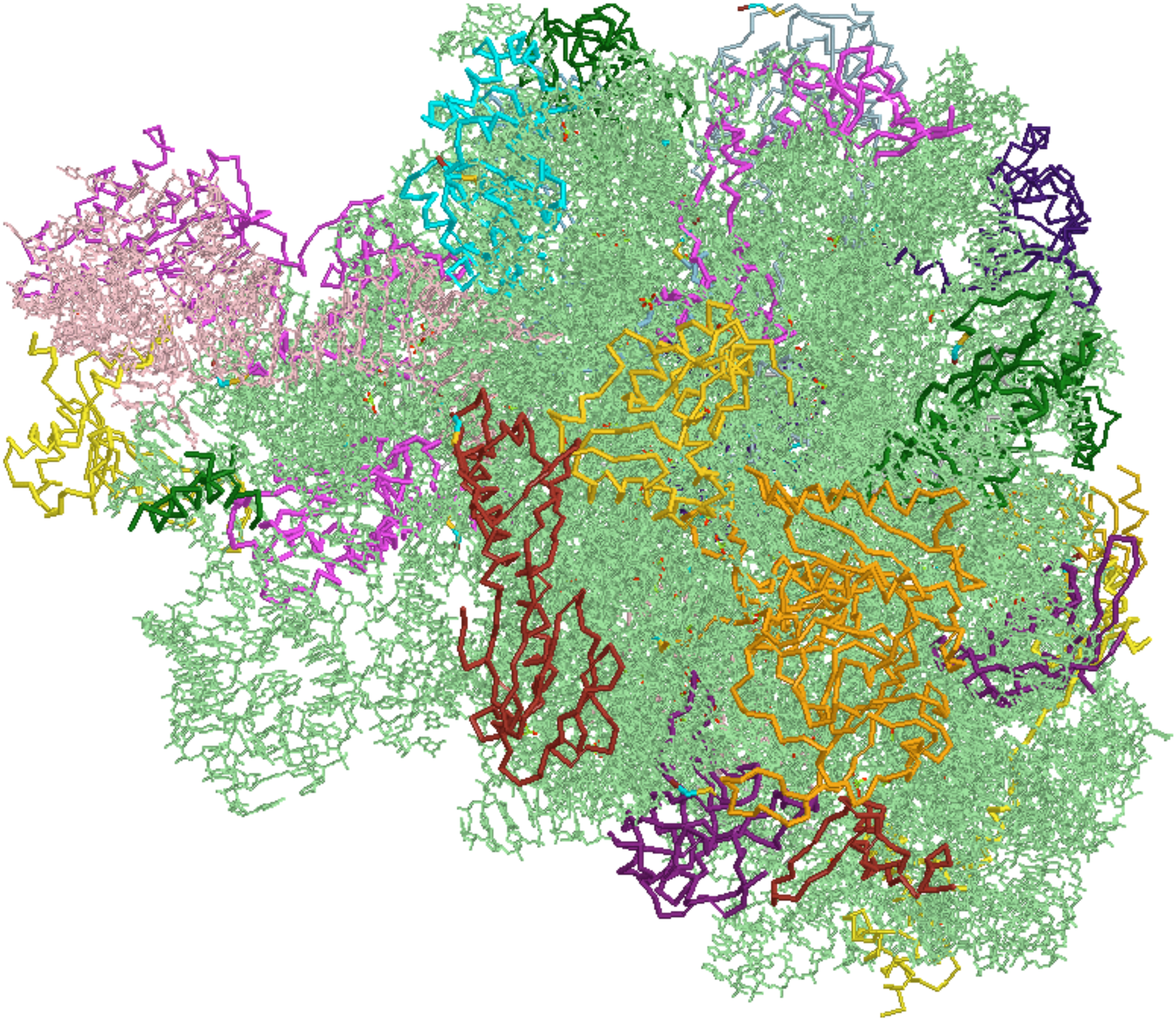
AMP Aptamer



◎ ☆

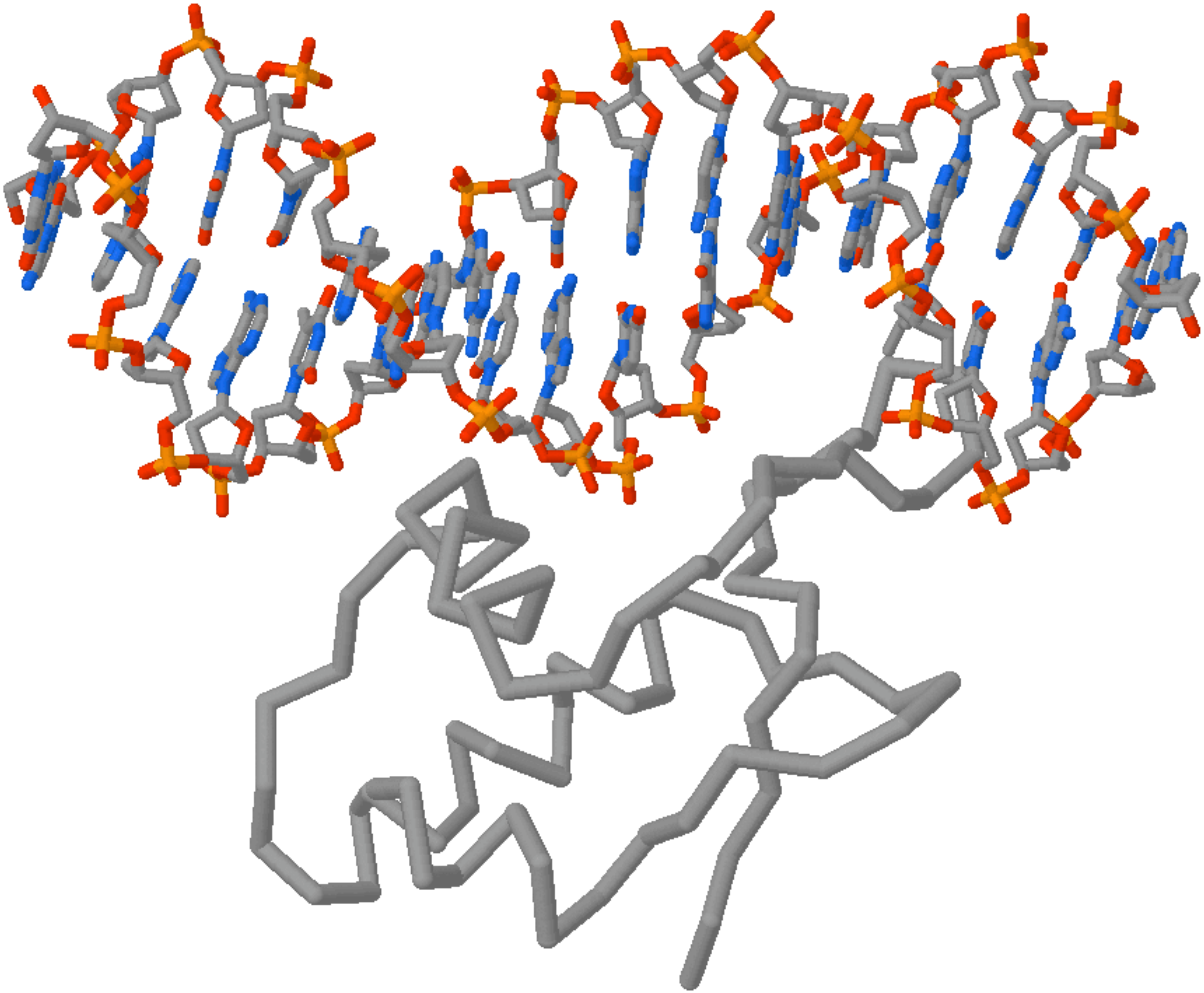
Ribosome

An RNA machine with protein cofactors



Winged Helix DNA Binding Domain

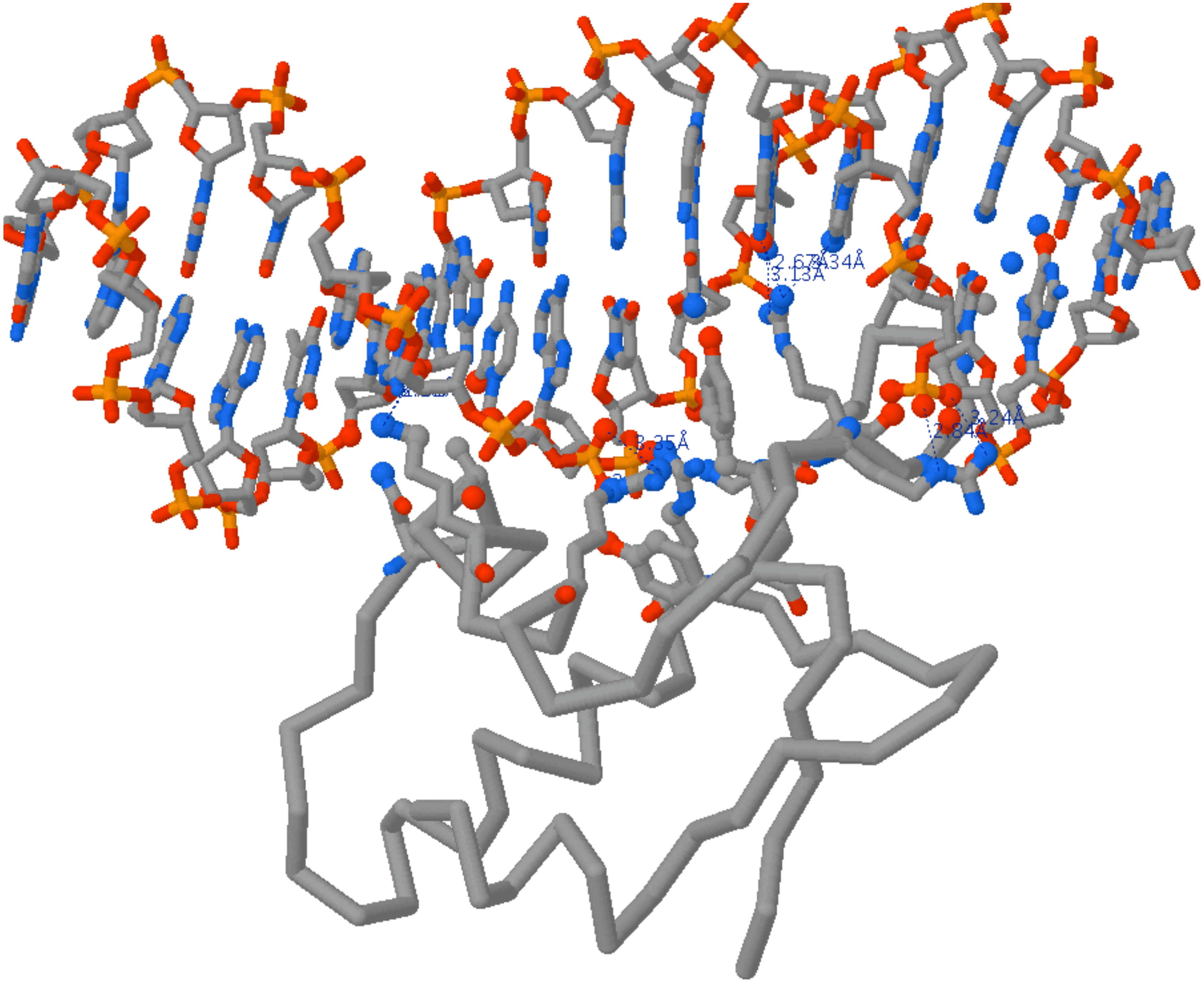
Classic helix-turn-helix



◎ ☆

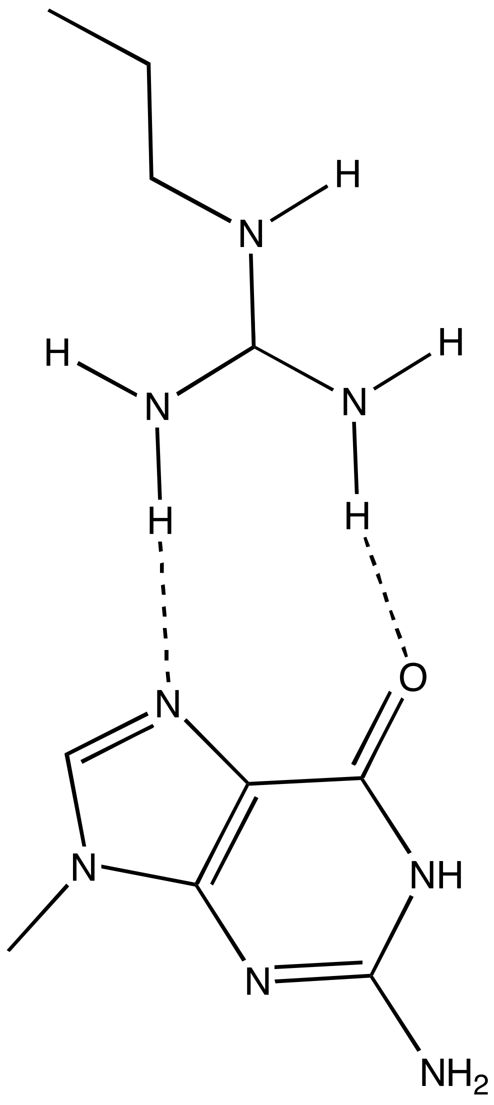
Winged Helix DNA Binding Domain

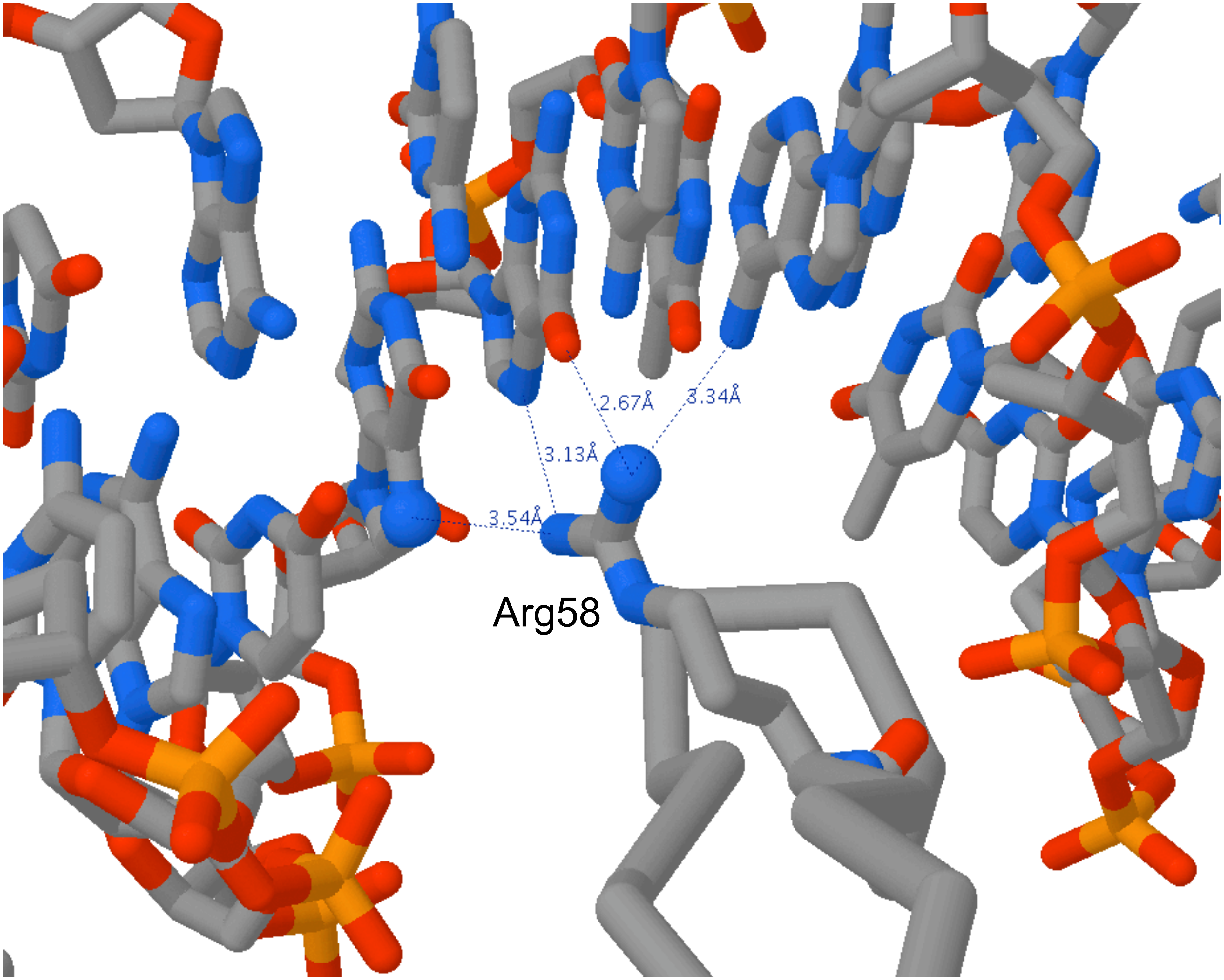
Classic helix-turn-helix



Winged Helix DNA Binding Domain

Classic helix-turn-helix





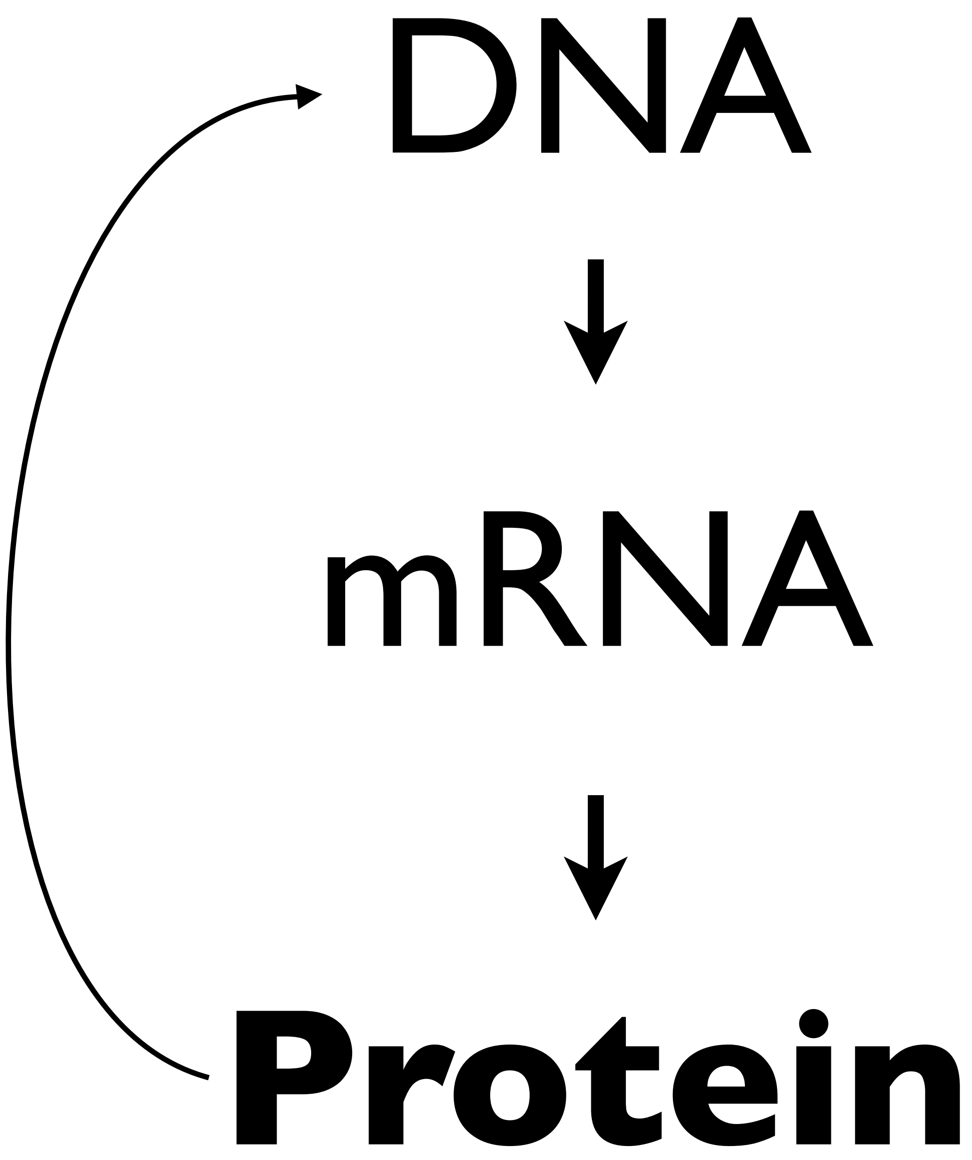
Hrfx1 bound to its X-box binding site

A Brave New World

revolutions at the turn of the century

opportunities for the 21st century

In the beginning…



DNA

↓

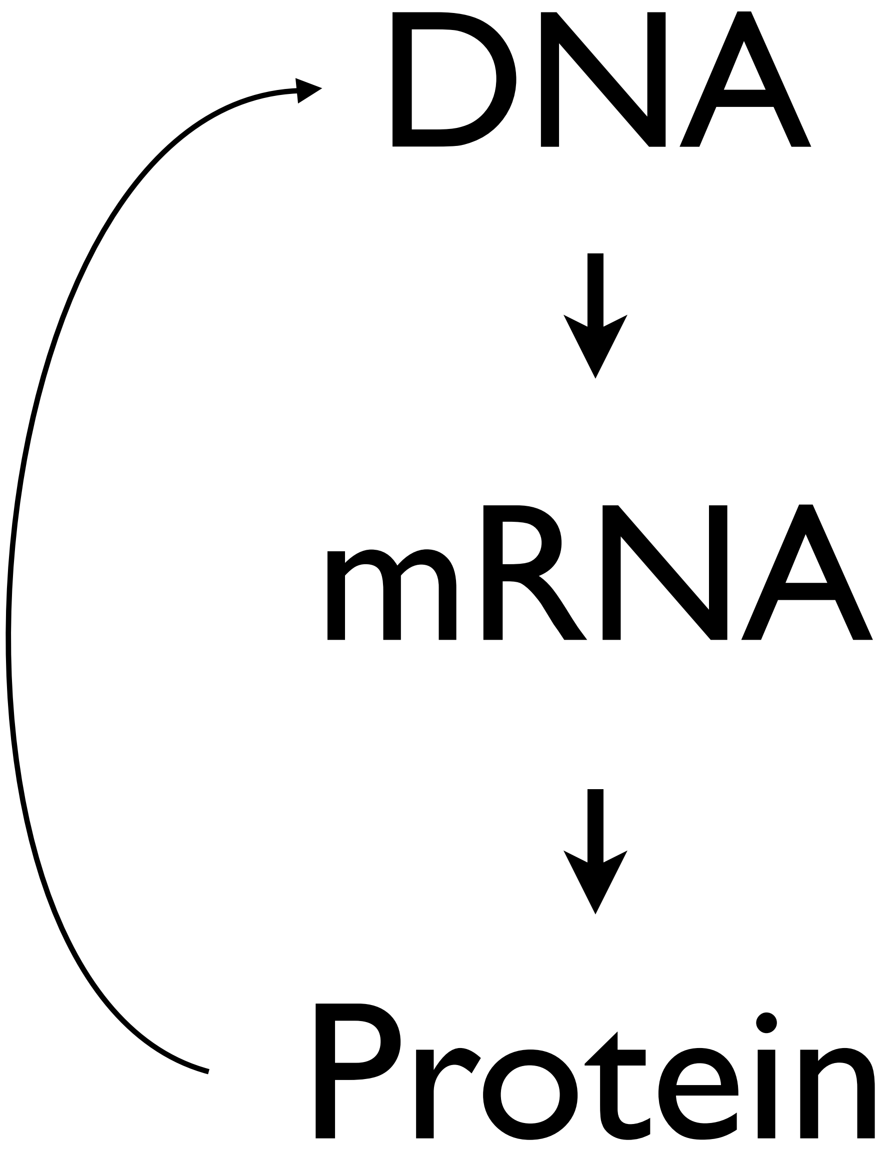
mRNA

↓

Protein

DNA	Archival information storage
mRNA	Transient information storage
Protein	Catalysis, structure, regulation, et al.

Chicken & Egg?



DNA

Archival
information storage

↓

mRNA

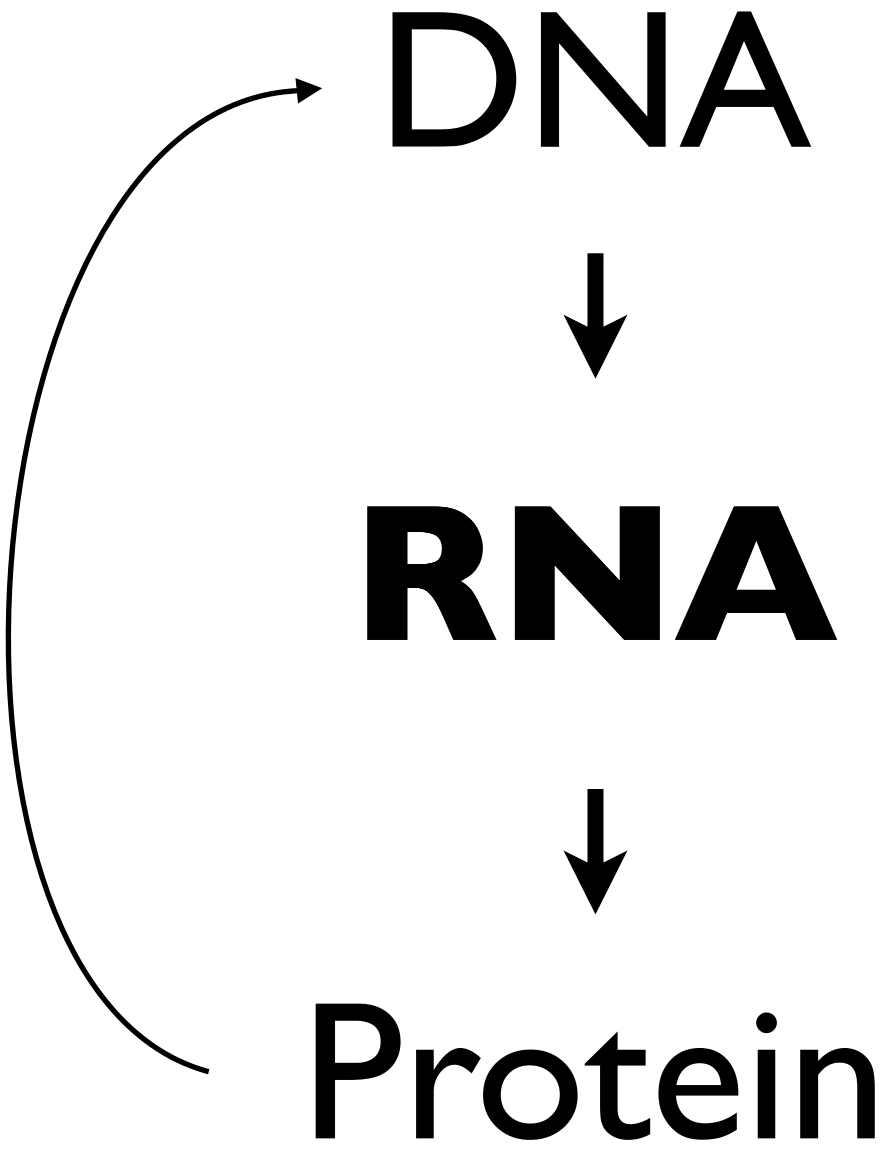
Transient
information storage

↓

Protein

Catalysis, structure,
regulation, et al.

RNA can do everything



DNA

Archival
information storage

↓

RNA

Transient
information storage
Catalysis!

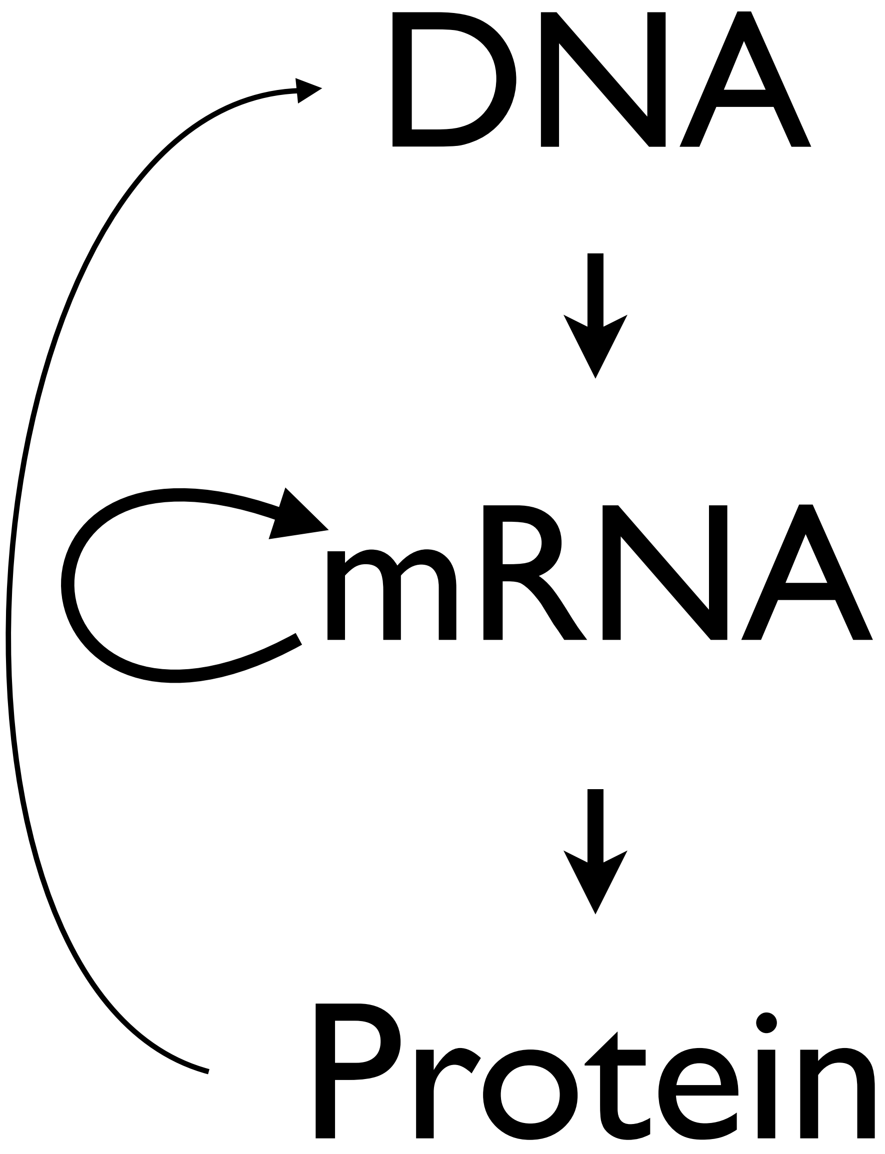
1980-2000

↓

Protein

Catalysis, structure,
regulation, et al.

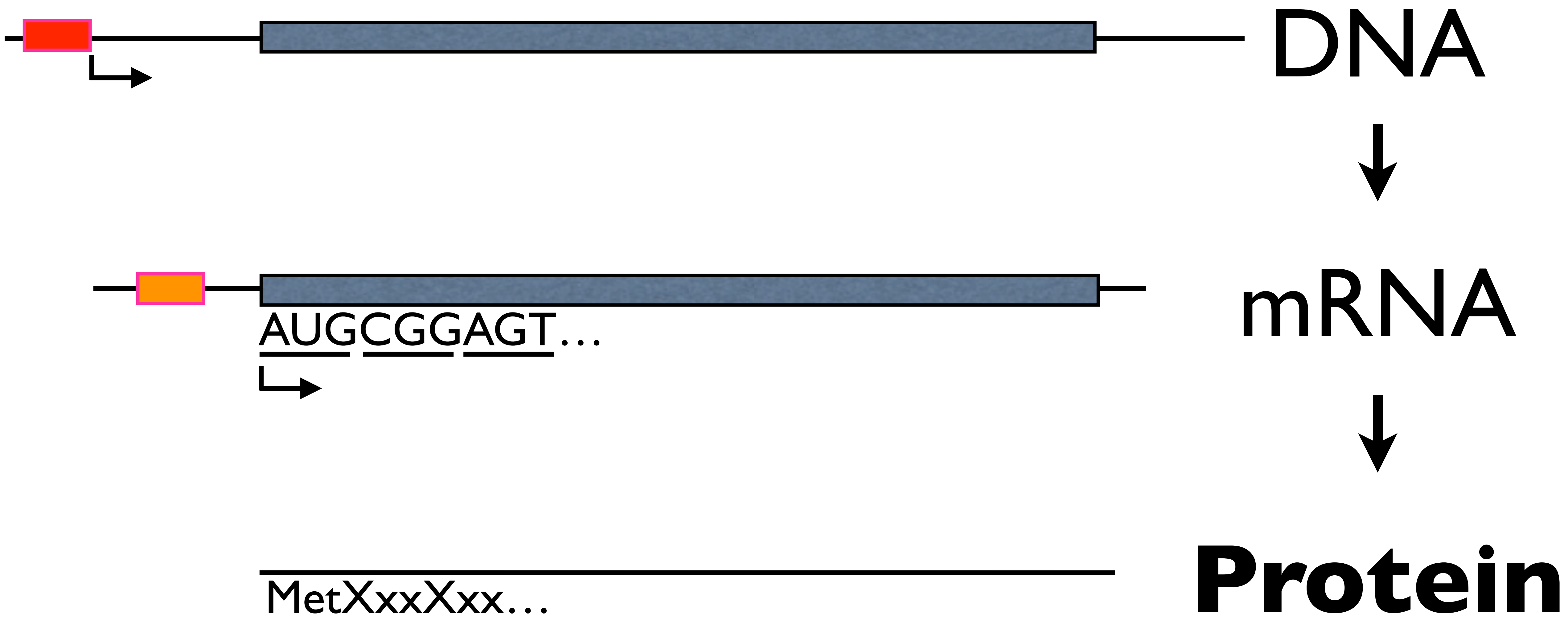
RNA - primordial molecule



So we accepted that RNA was probably the first, primitive do-everything biomolecule.

But **proteins** came along to supplant everything and make the world, evolutionarily, what it is today. All hail the protein!

Seems simple…



DNA

mRNA

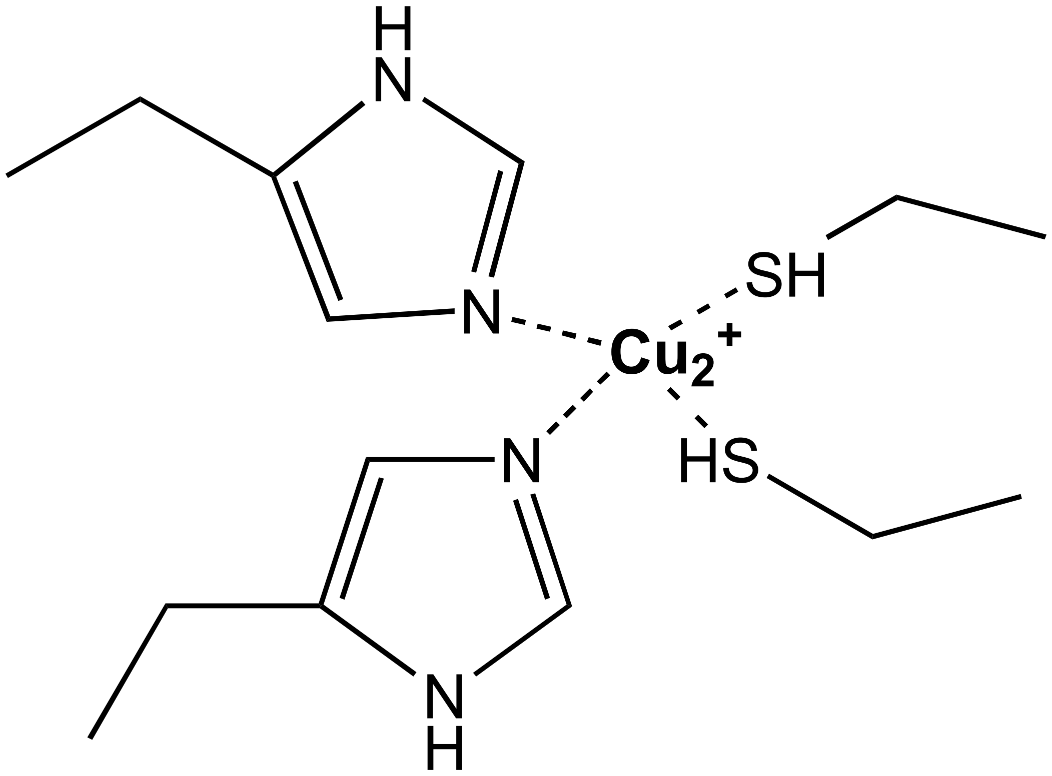
AUGCGGAGT…

Protein

MetXxxXxx…

Simple?

A personal aside



Cu_A - cytochrome *c* oxidase

PhD Dissertation 1984

ENDOR and EPR evidence for a model in which two absolutely conserved Cys residues coordinate copper

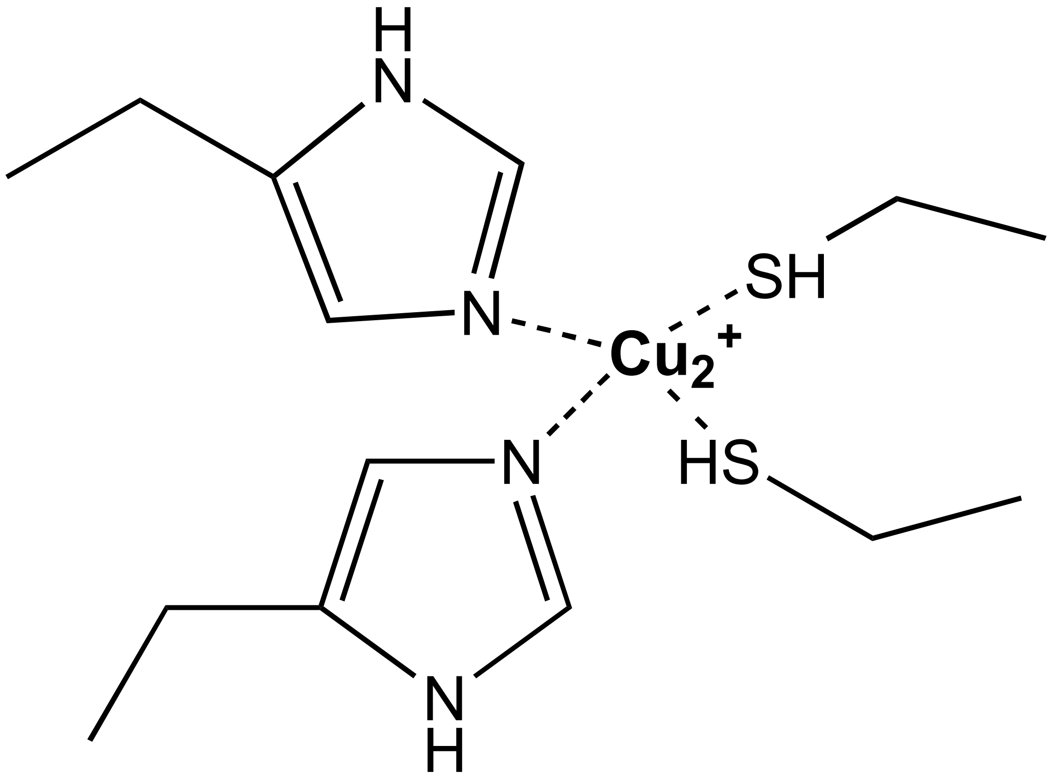
1983

Gene sequencing reveals that wheat cytochrome oxidase has one of the two conserved Cys replaced by Arg

Aaaarrggh!

Simple?

A personal aside



Cu$_A$ - cytochrome *c* oxidase

1984

Gene sequencing reveals that wheat cytochrome oxidase has one of the two conserved Cys replaced by Arg

1985-1995

Deamination of cytosine converts C to U, changing Arg codon back to that for Cys!

The RNA is edited!

A sequence-specific and precisely controlled process

Simple?

A personal aside

12 subunits

Subunits I, II, III
- major subunits

Subunits IV-XII
- minor subunits

Cytochrome *c* oxidase

1984

Trypanosomes completely lack subunit III

1985

U's are post-transcriptionally added to the pre-mRNA, generating full-length subunit III

Two thirds of the final gene is added U's!

A sequence-specific and precisely controlled process

Seems simple?

DNA

AUGCGGAGT…

pre-mRNA

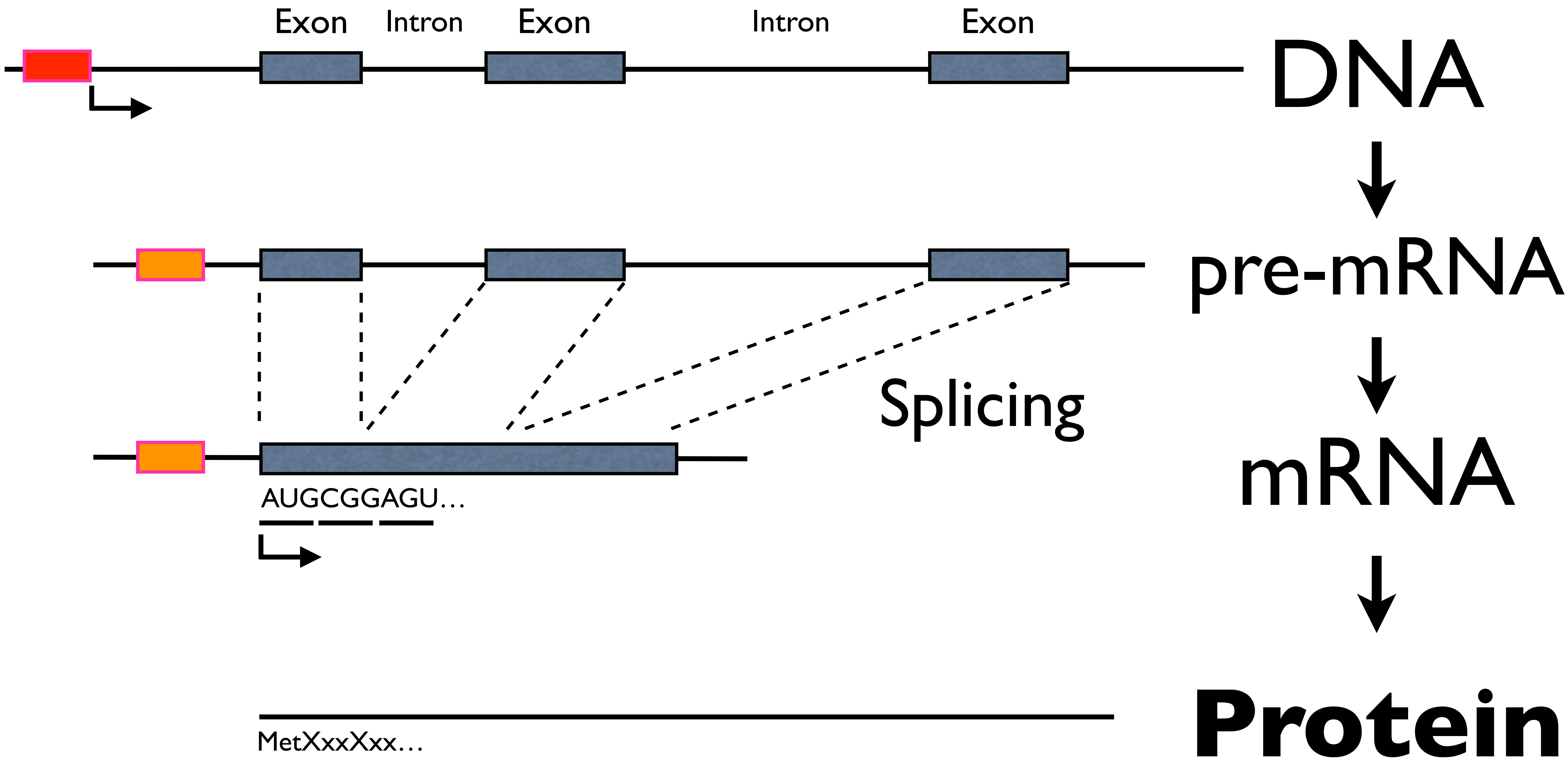
AUGUUCGUGAGC…

mRNA

MetZzzXxxWww…

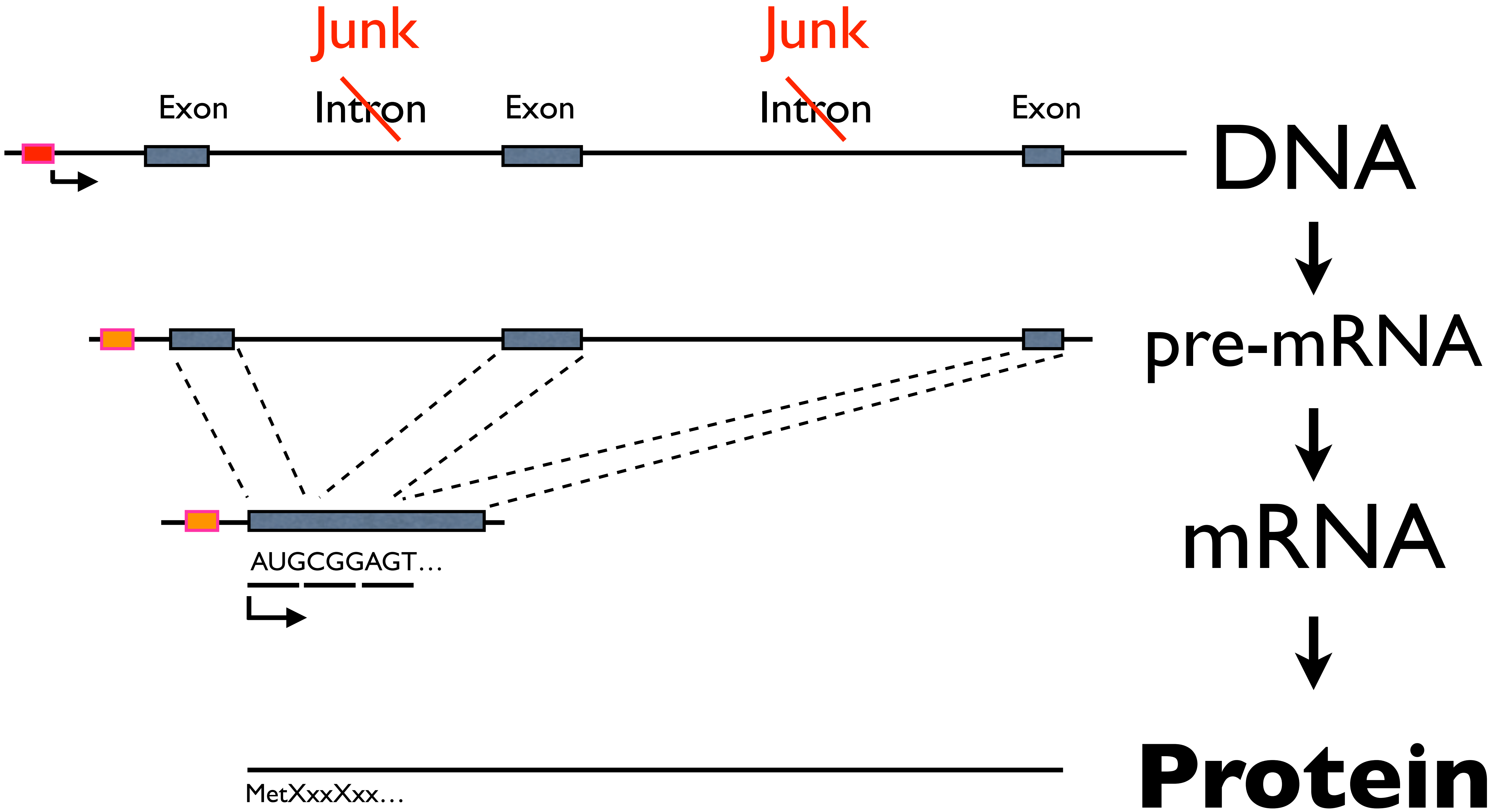
Protein

Even more complicated…



Reality…

and our perception of it



Junk Junk

Exon ~~Intron~~ Exon ~~Intron~~ Exon

DNA

pre-mRNA

mRNA

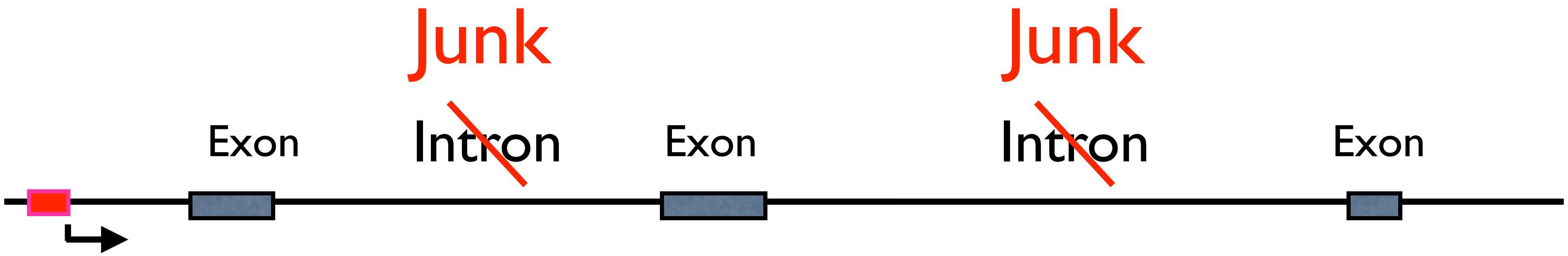
AUGCGGAGT…

Protein

MetXxxXxx…

Reality…

and our perception of it



Genome project goals

Identify and characterize the proteins.
What are their structures?
What do they do?
How do they interact?

20th Century View

Polymerases

Kinases

Hydrogenases

Receptors

Oxygenases

Proteases

Nucleic acids → Trash

21st Century News

Genome project

Number of protein-encoding genes in the human genome: 25,000

Number of protein-encoding genes in the 1,000 cell *C elegans* genome: 19,500

Number of protein-encoding genes in the corn genome: 40,000

21st Century News

Genome project

Number of protein-encoding genes in the human genome: 25,000

Number of different proteins: ≫25,000

How? RNA editing and alternative splicing

- Up to 3/4 of all human genes are subject to alternative editing
- The prevalence of alternative editing appears to increase with an organism’s complexity

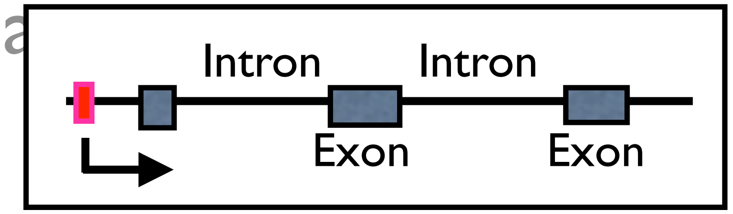
21st Century News

Genome project

How? RNA editing and alternative splicing

Up to 3/4 of all human genes are subject to alternative editing

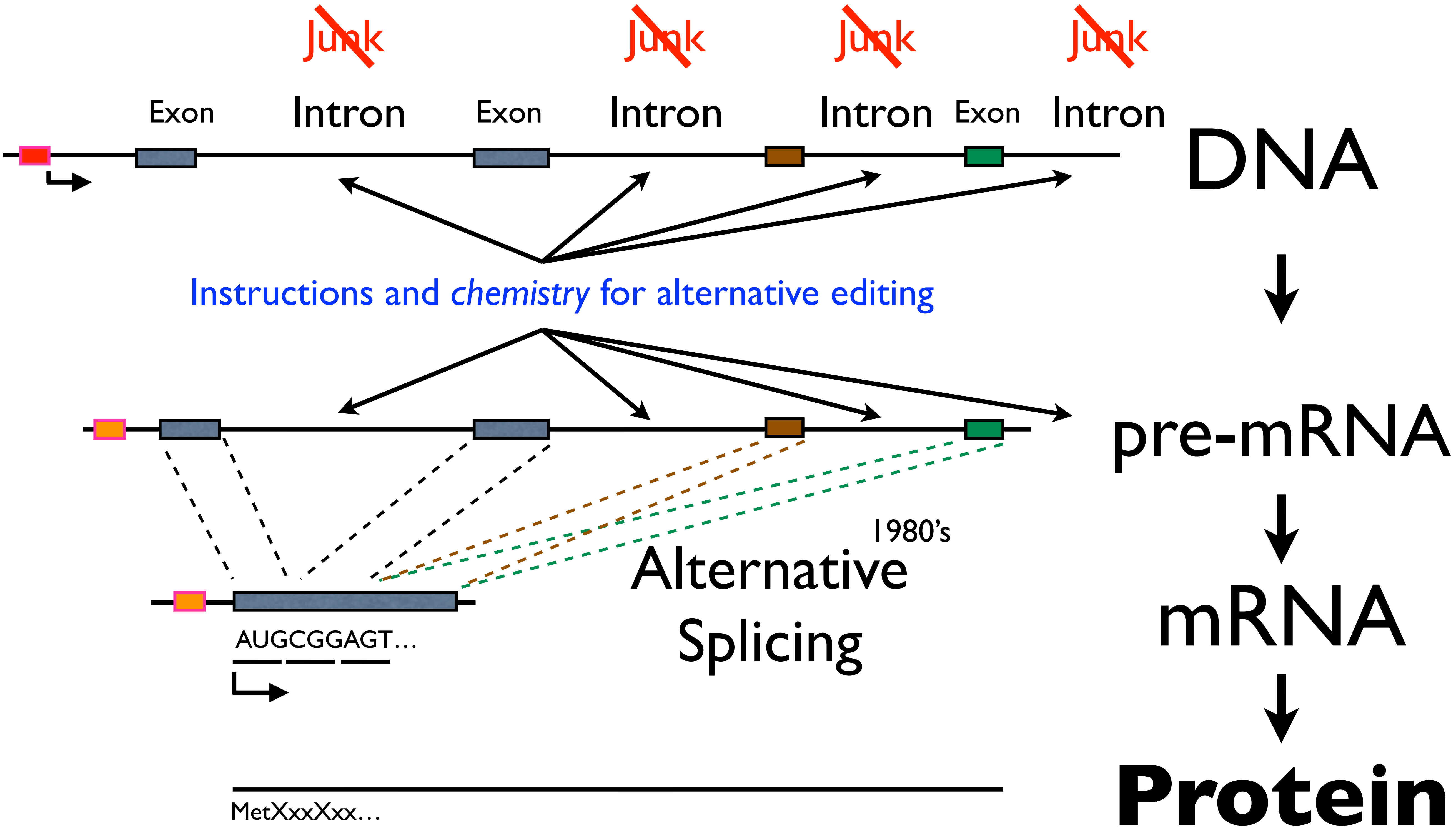
The prevalence of alternative editing appea
increase with an organism's complexity



Humans have the highest number of introns per gene of any organism

At least 15% of of the gene mutations that produce genetic diseases and cancers do so by effecting pre-mRNA editing

21st Century News



~~Junk~~ ~~Junk~~ ~~Junk~~ ~~Junk~~

Exon Intron Exon Intron Intron Exon Intron

DNA

Instructions and *chemistry* for alternative editing

pre-mRNA

1980's

Alternative Splicing

AUGCGGAGT…

mRNA

MetXxxXxx…

Protein

Back in the lab…

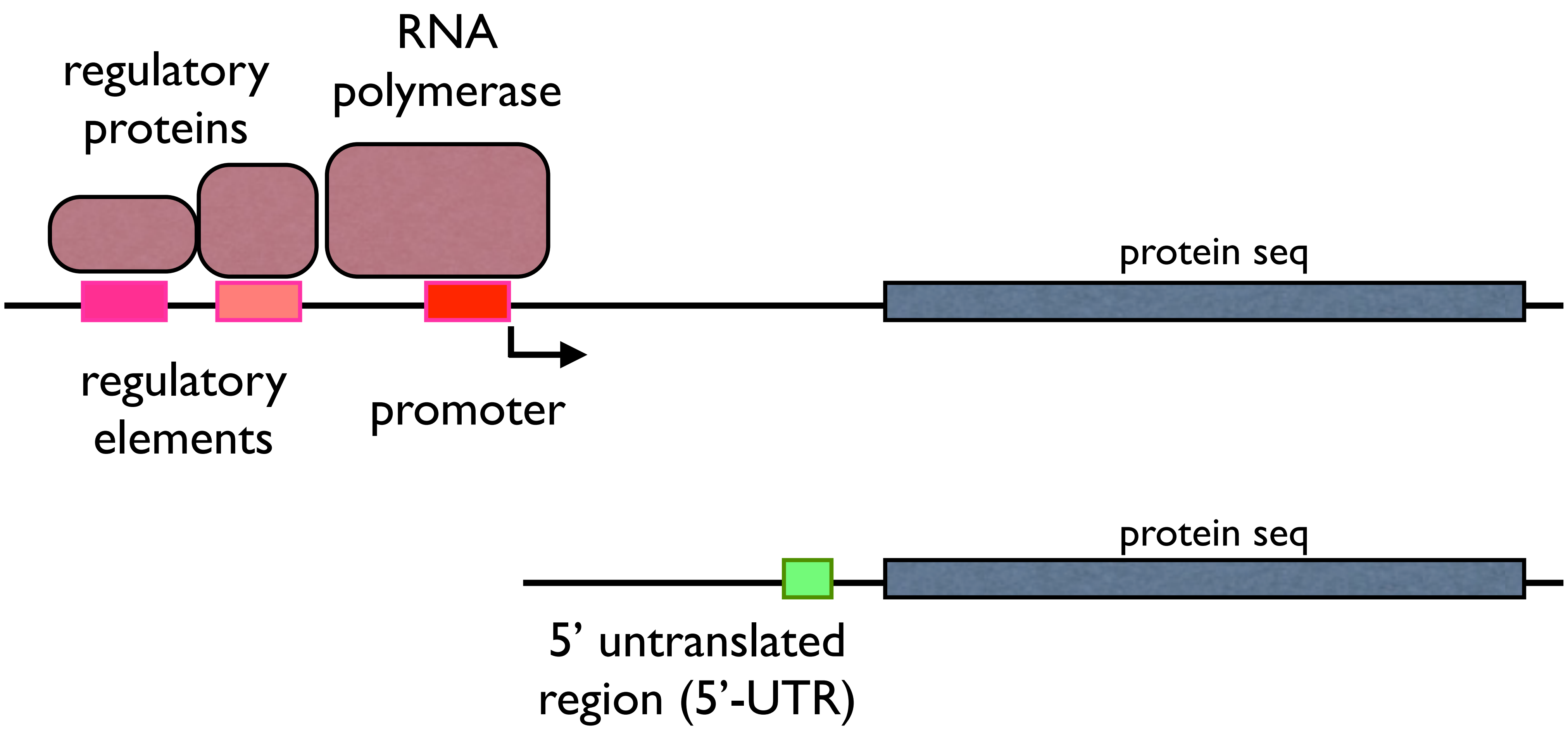
With the advent of PCR, nucleic acids have been recognized as extremely powerful combinatorial tools in the test tube

Aptamers can be selected that bind to "your favorite molecule"

Can create riboswitches

Ron Breaker:
"if it's so easy for us, I'll bet nature exploits this"

More gene regulation



regulatory
proteins

RNA
polymerase

protein seq

regulatory
elements

promoter

protein seq

5’ untranslated
region (5’-UTR)

DNA
↓
mRNA
↓
Protein

Ron: search for genes with no known protein regulator and which have a highly conserved 5’ UTR

Does RNA from that conserved 5’ UTR bind the product or substrate of the encoded enzyme? ↘ One week: a Nature paper!

2004

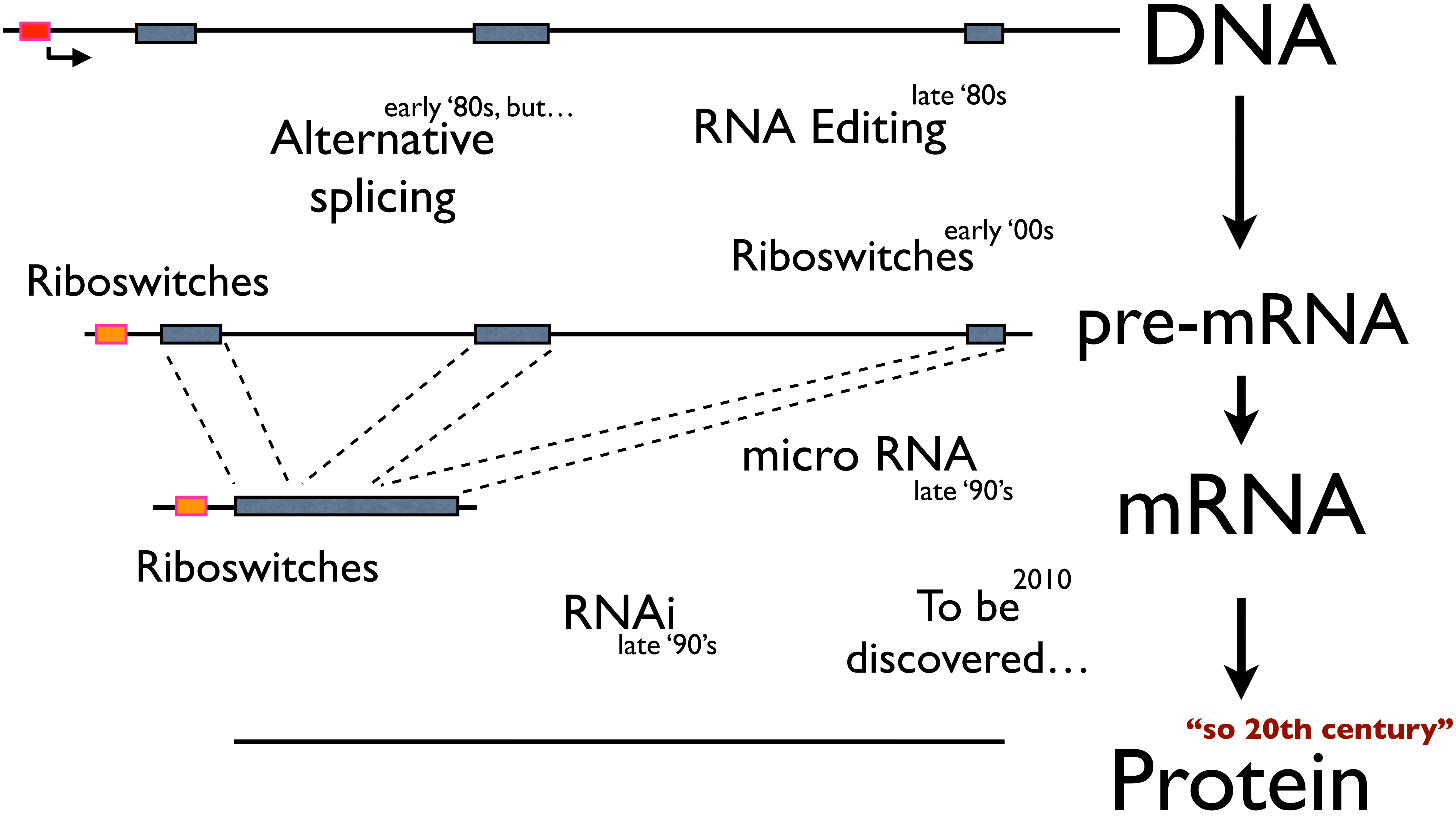
Project Encode (2007)

(More) rewriting of textbooks

June 2007, published in Nature

- Some regions of DNA far from protein-coding genes (extreme "junk?") are nevertheless highly conserved
- Most of both strands of the DNA is transcribed (far beyond that required for protein-coding genes)

21st Century Opportunities



DNA

early ‘80s, but…

Alternative splicing

RNA Editing late ‘80s

Riboswitches early ‘00s

Riboswitches

pre-mRNA

micro RNA late ‘90’s

mRNA

Riboswitches

RNAi late ‘90’s

To be discovered… 2010

“so 20th century”

Protein

3rd Annual Meeting of the Oligonucleotide Therapeutics Society
October 4-6, 2007, Langenbeck-Virchow-Auditorium, Berlin

Thursday, October 4th, 2007

8:30
START OF MEETING
WELCOME AND ANNOUNCEMENTS

Karl-Hermann Schlingensiepen, Antisense Pharma, Germany
The TGF-B2 inhibitor AP 12009 for recurrent or refractory high grade glioma: Results of a Phase IIb study.
25 min

Martin Gleave, OncoGenex Technologies, Canada
Targeting the cytoprotective chaperone, Clusterin, to enhance chemotherapy in advanced prostate cancer.
25 min

Michael Lahn, Eli Lilly and Company, USA
Clinical development of LY2181308 for cancer
25 min

Rosanne Crooke, Isis Pharmaceuticals, USA
Clinical development of ISIS 301012 for dyslipidemia
25 min

OLIGONUCLEOTIDE THERAPEUTICS SOCIETY

October 4-6, 2007

3rd Annual Meeting

Thomas Tuschl, Rockefeller University, USA
Mechanisms of Small RNA Mediated Mammalian Gene Silencing
30 min

Dinshaw J. Patel, Memorial Sloan Kettering Cancer Center, USA
Structural Biology of RNA Silencing
30 min

Ingo Roehl, Alnylam Europe, Germany
Analytical approaches to determine quality, pharmacokinetic properties and biodistribution of siRNAs
30 min

COFFEE BREAK 30 MIN

Peter Linsley, Rosetta Inpharmatics, USA
Widespread siRNA "off-target" transcript silencing mediated by seed region sequence Complementarity
30 min

Craig P. Hunter, Harvard University, USA
Systemic RNAi and intracellular transport
30 min

Markus Stoffel, ETH Zurich, Switzerland
Therapeutic strategies using the microRNA pathway
30 min