DNA Structure and Properties

Biochemistry Boot Camp 2019
Session #7
Siddik Alom
ma1889@msstate.edu

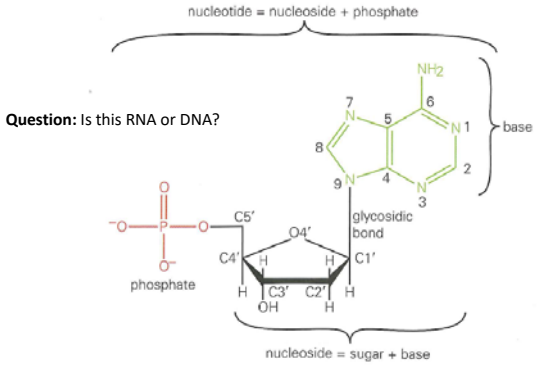
1

DNA

- DNA- a polymer of deoxyribonucleotides
- Found in chromosomes, mitochondria and chloroplasts
- Carries the genetic information

2

Nucleic Acid Structure



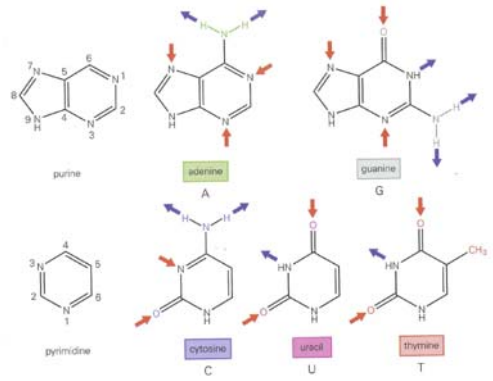
Question: Is this RNA or DNA?

Molecules of Life, pp. 15

3

3

Nucleic Acid Bases



Molecules of Life, pp. 20

4

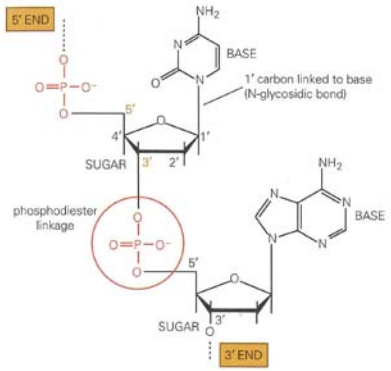
4

Nomenclature

	Base	Nucleoside	Nucleotide	Nucleic Acid
Purine	Adenine	Adenosine	Adenylate	RNA
		Deoxyadenosine	Deoxyadenylate	DNA
	Guanine	Guanosine	Guanylate	RNA
		Deoxyguanosine	Deoxyguanylate	DNA
Pyrimidines	Cytosine	Cytidine	Cytidylate	RNA
		Deoxycytidine	Deoxycytidylate	DNA
	Thymine	Thymidine	Thymidylate	
		Deoxythymidine	Deoxythymidylate	DNA
	Uracil	Uridine	Uridylate	RNA

5

Nucleic Acids Are Also Polymers



DNA & RNA Polymerase: Build up DNA and RNA from nucleoside triphosphates (5′ → 3′ synthesis)

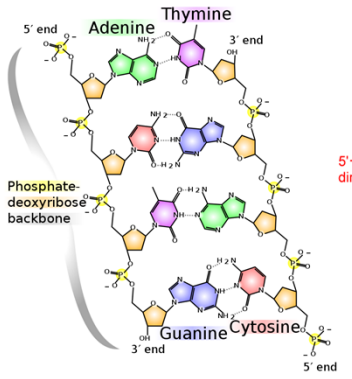
Convention: RNA/DNA typically is read from 5′ to 3′ direction (e.g. 5′-ATTGCAAC-3′)

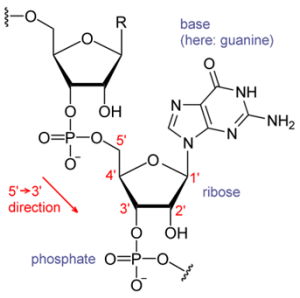
Molecules of Life, pp. 21

6

6

DNA vs RNA





- DNA less reactive
- RNA is easily attacked by enzymes

7

Nucleic Acid Primary Structure

- **Just like proteins:** the sequence of bases

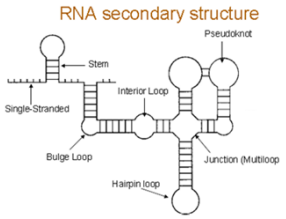
5′-dAdGdTdTdCdAdCdCdC-3′ (DNA)
AGTTCACCC

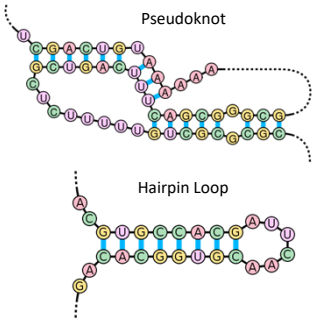
5′-AGUUCACCC-3′ (RNA)

8

8

Secondary Structure





- Base pairing motifs

Source: Wikipedia, "RNA Secondary Structure," "Nucleic Acid Secondary Structure"

9

9

Tertiary Structure

A Form DNA
Due to dehydration

B Form DNA
Most naturally occurring DNA

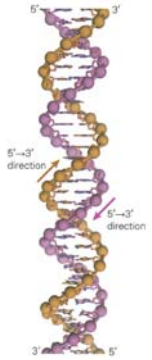
Z Form DNA
Chemical modification like methylation

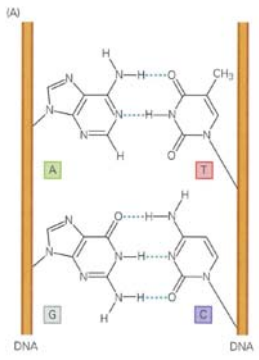
Source: Steven Carr, www.mun.ca

10

10

Watson-Crick Base Pairing in an (Antiparallel) Double Helix





Molecules of Life, pp. 23

11

11

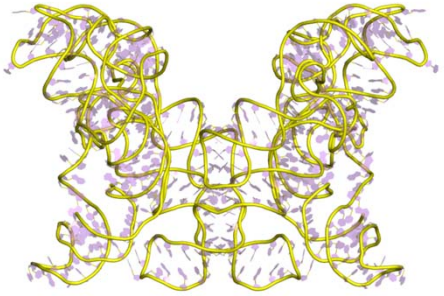
Reverse Complement

- Watson-Crick base pairing
  - A pairs with T (or U in RNA)
  - G pairs with C
- RNA can "hybridize" with DNA, forming mixed strands
- **Example:** What's the reverse complement to AUCCGCCTT?

12

12

Tertiary & Quaternary Structure



Ribozyme: An RNA capable of catalyzing a chemical reaction

The ribosome contains a significant amount of RNA as well as proteins

Macromolecules can perform incredibly diverse structures!
(And we haven't even mentioned lipids and sugars.)

Wikipedia, "Group I Catalytic Intron." Accessed 8/23/2012.

13

13

Think and Discuss

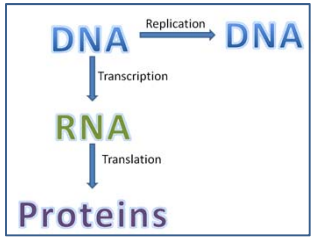
What are the major differences between DNA and protein structures? What are the similarities?

14

14

Central Dogma

- DNA → mRNA "Transcription"
  - Synthesized RNA Polymerase
  - RNA formed from 5' to 3'
- mRNA → Protein "Translation"
  - Synthesized by ribosome
  - New proteins formed from NT to CT



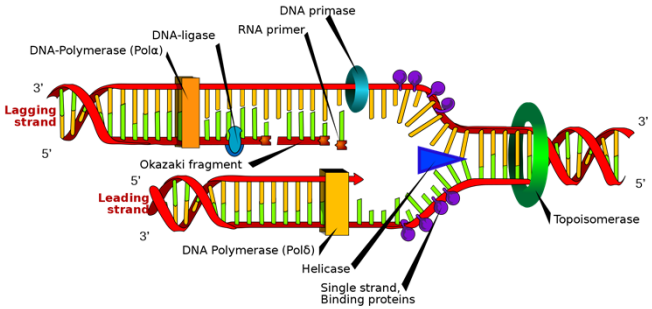
Trick: Reading the DNA in the "standard way", one can easily identify the codons for peptide synthesis.

Source: Wikipedia, "Ribosome"

15

15

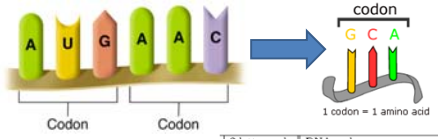
Replication: Duplicating DNA



16

16

Genetic Code



- But 1 amino acid= multiple codons (with exception)

3-letter code	DNA codons
Ile	ATT, ATC, ATA
Leu	CTT, CTC, CTA, CTG, TTA, TTG
Val	GTT, GTC, GTA, GTG
Phe	TTT, TTC
Met (start)	ATG
Cys	TGT, TGC
Ala	GCT, GCC, GCA, GCG
Gly	GGT, GGC, GGA, GGG
Pro	CCT, CCC, CCA, CCG
Thr	ACT, ACC, ACA, ACG
Ser	TCT, TCC, TCA, TCG, AGT, AGC
Tyr	TAT, TAC
Trp	TGG
Gln	CAA, CAG
Asn	AAT, AAC
His	CAT, CAC
Glu	GAA, GAG
Asp	GAT, GAC
Lys	AAA, AAG
Arg	CGT, CGC, CGA, CGG, AGA, AGG
termination	TAA, TAG, TGA

Source: Wikipedia, "Genetic Code"

17

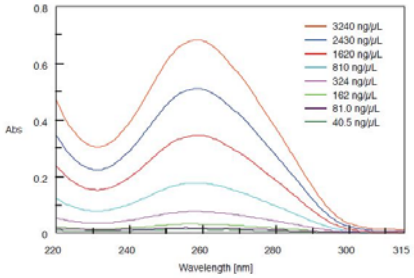
17

Nucleic Acid Extinction Coefficient

DNA Concentrations: Often measured in µg/mL (or the equivalent ng/µL) instead of M, mM, etc. Also sequence isn't exactly known in many cases.

Rule of Thumb: For double-stranded, plasmid DNA, the extinction coefficient at 260 nm is

0.020 $(\mu g/mL)^{-1}\,cm^{-1}$



Source: www.jascoinc.com

18

18

DNA vs. Protein Absorbance

DNA Concentrations: At 260 nm, double-stranded DNA has an extinction coefficient of

0.020 $(\mu g/mL)^{-1}\,cm^{-1}$

Protein Concentrations: At 280 nm, the GB3 protein has an extinction coefficient (in equivalent units) of

0.0016 $(\mu g/mL)^{-1}\,cm^{-1}$

Which is more sensitive?

What are the implications?

19

19

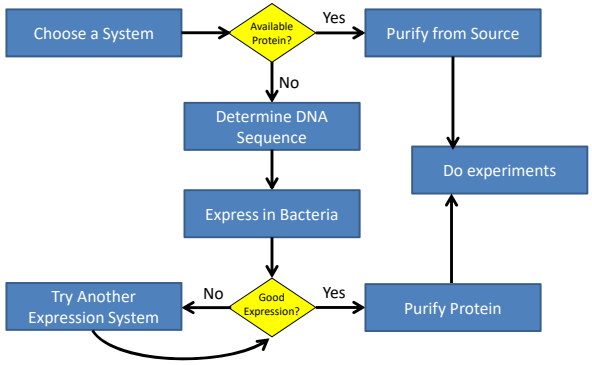
Think and Discuss

Our biochemistry experiments are normally done in aqueous buffer. Is this a good model for the inside of a cell?

20

20

Biochemistry Research Flow Chart

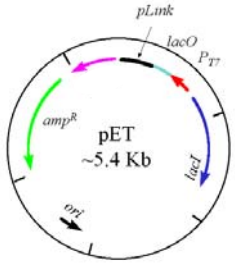


21

21

Bacterial Expression Vectors

- pET Plasmid Genes
  - Origin of replication
  - Lac repressor (lacI)
  - RNA Pol promoter (P_{T7})
  - Lac Operator (lacO)
  - Polylinker – where your DNA sequence goes (pLink)
  - Ampicillin resistance (amp^R)
- Is this plasmid persistent?



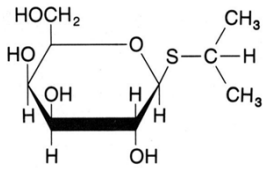
Source: Mike Blaber, BCH5425 Course Notes

22

22

Inducible Expression

- **IPTG:** Turns on protein expression without being hydrolyzed
- Protein expression can be switched on when desired

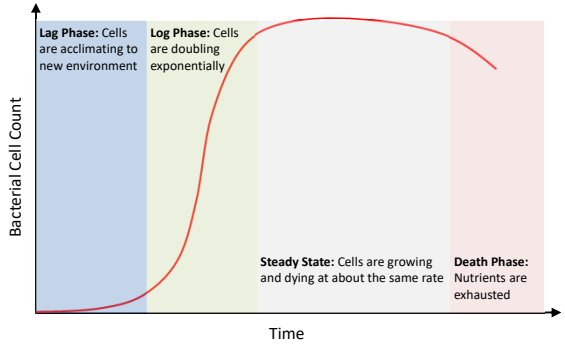


Isopropyl-β-D-thiogalactoside (IPTG)

23

23

When Should I Induce?

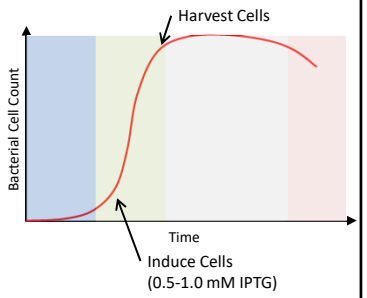


24

24

When Should I Induce?

- Protein expression is greatest during log phase
- Inducing at lag phase may unnecessarily cripple your cells
- Typically, induce at an OD_{600} of 0.5-0.6
- Always follow your lab's protocols!



25

25

Think and Discuss

Why is Ampicillin resistance necessary for the function of the pET vector system?

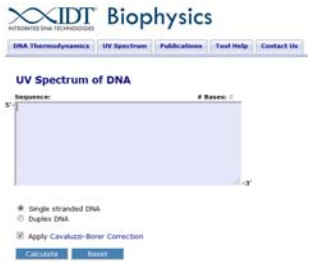
26

Summary

- DNA structure is as varied as protein structure, and nucleic acids can catalyze chemical reactions ("ribozymes")
- Bacterial and animal cells store and process DNA slightly differently, although both use similar ribosomes and the same genetic code
- Modern molecular biology allows us to express virtually any gene using bacterial expression systems

27

Nucleic Acids – Smaller Molecules



- IDT DNA Calculator: http://biophysics.idtdna.com/UVSpectrum.html

Source: www.jascoinc.com

28

28

Calculating Reverse Complement

- Bioinformatics.org Calculator (no-frills): http://bioinformatics.org/sms/rev_comp.html

Source: www.jascoinc.com

29

29

DNA Translation Tool

- **Site:** http://web.expasy.org/translate/
- **Input:** DNA or RNA sequence (5' → 3' orientation)
- **Output:** All six possible translation frames

30

30

Other Databases

- NCBI Databases work for DNA sequences, too (reference sequences start with NM_)
- PDB also houses a number of RNA/DNA structures in addition to proteins

31

31

Putting it Together:

SDSC Biology Workbench

SDSC
SAN DIEGO SUPERCOMPUTER CENTER
Biology WorkBench

- **Site:** http://workbench.sdsc.edu/
- **Exercise:** Create an account, try to examine some of the tools. What looks familiar?

32

32

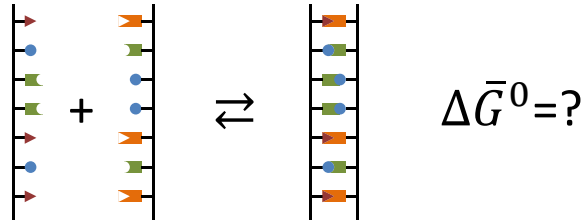
Think And Discuss

How can these databases be used to make your lab work easier? What are some practical examples

33

33

DNA "Melting"



- Two strands come together:
  - How much work can be done?
  - Which side of the reaction does temperature favor?

34

Predicting Melting Temperatures

- To calculate T_m, add 4 °C for each G-C pair, and 2 °C for each A-T
  - Not terribly accurate
- **Example:** GCCCTGAAGGTCAAGTCCCCC
  - 14 G-C = 56 °C
  - 7 A-T = 14 °C
  - Prediction is 70

35

35

Predicting Melting Temperatures

- IDT OligoAnalyzer: http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/
- **Input:** Your DNA sequence of interest, salt concentration
- **Output:** T_m, extinction coefficient, %GC content

36

36

Predicting Secondary Structure

- `mfold` Web Server: http://mfold.rna.albany.edu/?q=mfold
- **Input:** RNA/DNA sequence
- **Output:**

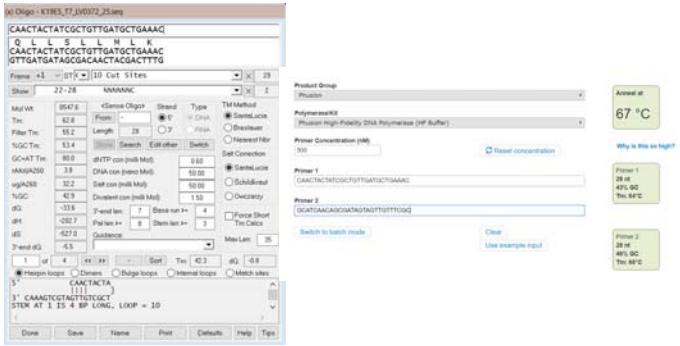
37

General Primer Design Principles

- **PCR Steps:** Denature (98 °C), anneal (~ 60 °C), extend (72 °C)
- Considerations:
  - **Melting Temperature:** Should be 52-58 °C
  - **GC Content:** 40-60%
  - **Length:** ~30 bp (but longer can be okay)
  - **Secondary Structure:** Avoid if possible
- Lots of software exists (some costs $$$). For more information (some trial and error here):
  https://goo.gl/4EwMG3 (Life Technologies)
  http://www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html

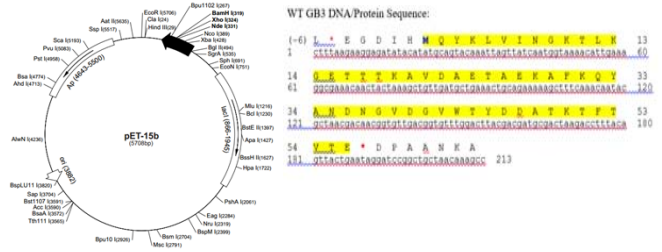
38

Tm Calculators



39

pET-15b Plasmid and WT GB3 Sequence



Slide Courtesy: Dr. Fitzkee

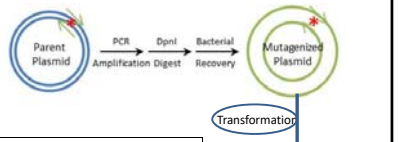
40

40

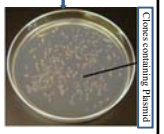
Designing of Primers for PCR: Site-Directed Mutagenesis

- Phusion DNA polymerase enzyme
- Annealing tem. 53°C

Primer Name	Primer Sequence (5' to 3')
K19I-Fwd	CAACTACTATCGCTGTTGATGCTGAAAC
K19I-Rev	GCATCAACAGCGATAGTAGTTGTTTCGC
K19C -Fwd	CAACTACTGACGCTGTTGATGCTGAAAC
K19C -Rev	GCATCAACAGCGTCAGTAGTTGTTTCG
K19R -Fwd	CAACTACTCGTGCTGTTGATGCTGAAAC
K19R -Rev	GCATCAACAGCACGAGTAGTTGTTTCG



Transformation



Description	Primer Duplex
K19I -Fwd/ Template	5′-CAACTACTATCGCTGTTGATGCTGAAAC 3′-TAACTTTCCGCTTTGTTGATGATTTCGACAACTACGACTTTGACGT
K19I -Rev/ Template	CATTGAAAGGCGAAACAACTACTAAAGCTGTTGATGCTGAAACTGCAGAAA CGCTTTGTTGATGATAGCGACAACTACG
K19I -Fwd/ K19I -Rev	CAACTACTATCGCTGTTGATGCTGAAAC CGCTTTGTTGATGATAGCGACAACTACG

41

41

Think And Discuss

Compared to DNA, why is it harder to calculate melting temperature and dimerization for proteins?

42

42

Think and Discuss

What problems could arise when introducing new mutations in to a known sequence?

43

Summary

- Advanced computational tools for nucleic acids depend on two things:
  - The simplicity of DNA primary structure (4 bases)
  - The regularity of Watson-Crick base pairing
- Combining DNA and protein tools makes it possible to perform very advanced sequence analysis

44