

Calculating DNA Properties

Biochemistry Boot Camp 2021

Session #8

Chris Johnson

cn.johnson@chemistry.misstate.edu

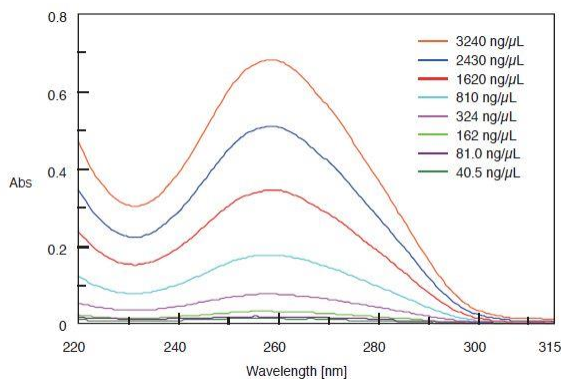
1

Nucleic Acid Extinction Coefficient

DNA Concentrations: Often measured in $\mu\text{g/mL}$ (or the equivalent $\text{ng}/\mu\text{L}$) 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

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



Source: www.jascoinc.com

2

2

DNA vs. Protein Absorbance

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

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

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

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

Which is more sensitive?

What are the implications?

3

3

Other Values for Long, Randomized Sequences

- Single-Stranded RNA: $0.025 (\mu\text{g/mL})^{-1} \text{ cm}^{-1}$
- Single-Stranded DNA: $0.030 (\mu\text{g/mL})^{-1} \text{ cm}^{-1}$
- For a pure nucleic acid, the 260/280 nm ratio should be approximately 1.8-2.0

4

4

Nucleic Acids – Smaller Molecules

OligoAnalyzer

Sequence: 5' MOD • INTERNAL • 3' MOD • MIXED BASES •

CGA AGA ACA GGA AGC GGA ATT TAA AGA AG

Bases 29

CLEAR SEQUENCE

Try the new batch mode here

Parameter sets

SpecSheet (Default)

Target type: DNA

Oligo Conc: 0.25 µM

Na⁺ Conc: 50 mM

Mg⁺⁺ Conc: 0 mM

dNTPs Conc: 0 mM

Instructions | Definitions | Feedback

ANALYZE

HAIRPIN

SELF-DIMER

HETERO-DIMER

NCBI-BLAST

TM MISMATCH

ADD TO ORDER

- IDT DNA Analyzer (extinction coefficient, T_m):
<https://www.idtdna.com/pages/tools/oligoanalyzer>
 – Need to log in, create an account (free)

5

5

Calculating Reverse Complement

Reverse Complement

Reverse Complement converts a DNA sequence into its reverse, complement, or reverse-complement counterpart. You may want to work with the reverse-complement of a sequence if it contains an ORF on the reverse strand.

Paste the raw or FASTA sequence into the text area below.

>Sample sequence
GGGGaaaaaaaaatttatatat

SUBMIT CLEAR

• Convert the DNA sequence into its reverse-complement counterpart.

[\[home\]](#)

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

Source: www.jascoinc.com

6

6

DNA Translation Tool

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

7

7

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

8

8

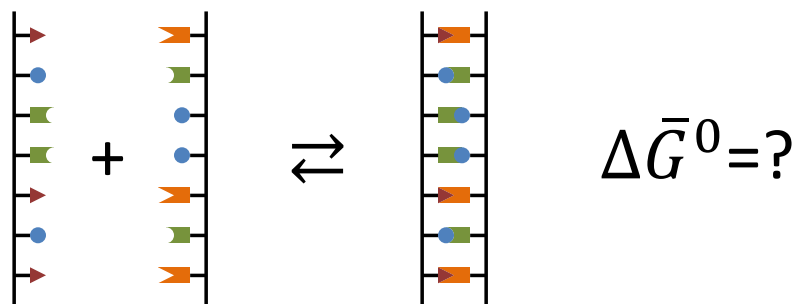
Think And Discuss

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

9

9

DNA “Melting”



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

10

Thermal Melts

- Adding heat favors highly random systems, DNA will separate at high temperature
 - Secondary and tertiary structure is lost, primary is maintained
- What will affect the melting temperature?

11

11

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

12

12

Predicting Melting Temperatures

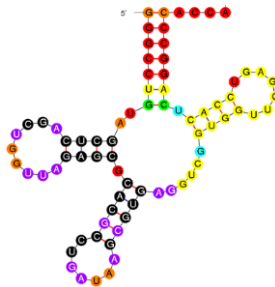
- IDT OligoAnalyzer:
<https://www.idtdna.com/pages/tools/oligoanalyzer>
- **Input:** Your DNA sequence of interest, salt concentration
- **Output:** T_m , extinction coefficient, %GC content

13

13

Predicting Secondary Structure

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



14

Example: HIV TAR RNA

- Trans-Activation Response Element – Binds with a protein (Tat) to promote viral transcription
- Sequence:

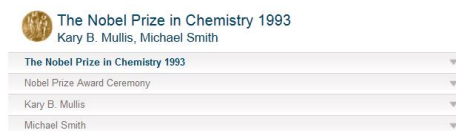
GGGUCUCUCUGGUUAGACCAGAUCUGAGCCUGGGAGCUCUCU
GGCUAACUAGGGAACCCAC

15

15

Why is this Useful?

- Site-Directed Mutagenesis
- Good Primers:
 - $T_m > 78^\circ\text{C}$ (2 mM MgCl_2 , 50 mM NaCl)
 - GC content > 40%
 - No secondary structure (< 50 bp)
 - End with G or C



Kary B. Mullis

Michael Smith

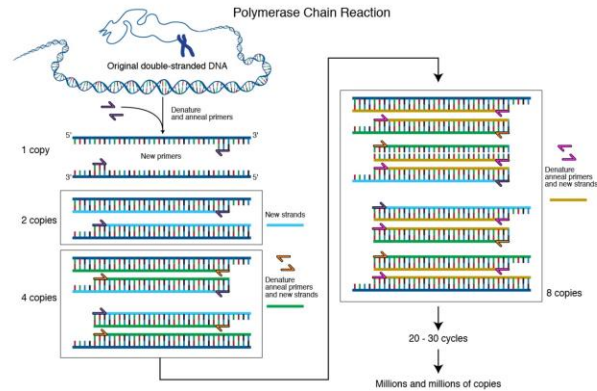
The Nobel Prize in Chemistry 1993 was awarded "for contributions to the developments of methods within DNA-based chemistry" jointly with one half to Kary B. Mullis "for his invention of the polymerase chain reaction (PCR) method" and with one half to Michael Smith "for his fundamental contributions to the establishment of oligonucleotide-based, site-directed mutagenesis and its development for protein studies".

Photos: Copyright © The Nobel Foundation

16

16

Polymerase Chain Reaction (PCR)



- **Key consideration:** Temperature for primer annealing (computational tools)!

<https://www.genome.gov/genetics-glossary/Polymerase-Chain-Reaction>

17

17

Example: Designing Primers

The screenshot shows the Agilent Technologies Genomics website. The navigation bar includes links for Home, Products & Services, Applications, and Resources & Support. The main content area is titled "QuikChange Primer Design" and includes a "Help" button. The text describes the QuikChange® Primer Design Program, which supports mutagenic primer design for QuikChange mutagenesis experiments. It mentions that the program calculates/designs appropriate primer sequences with optimal melting temperature. A link to "Expand Help" is provided. A numbered list starts with: "1. It is recommended that you clear this form prior to loading each new sequence:"

<http://www.genomics.agilent.com/primerDesignProgram.jsp>

18

18

PRODUCTS & SERVICES CROSSLAB SOLUTIONS SUPPORT LIBRARY TRAINING & EVENTS BUY

1. It is recommended that you clear this form prior to loading each new sequence:

[Clear Input](#)

2. Select QuikChange® mutagenesis kit that you are using:

QuikChange® II [Help in choosing a mutagenesis kit](#)

3. Find your DNA sequence by pressing

Load a file from your hard-drive: [Browse...](#) Or, paste plain text or FASTA-formatted DNA sequence in the box below:

No file selected [Browse...](#) atagatagacagacataccagacatacaggaggagatacacag
acatacagacaga

4. Load it.

[Upload New](#) or [Upload Translated](#)

or specify a DNA region to translate from to [Upload Translated Region](#)

5. Select up to seven nucleotides that you want to change

DNA: change nucleotide(s) to: Site 1 Site 2 Site 3 Site 4 Site 5 Site 6 Site 7

g

or

☐ Delete a region between two checked nucleotides (Note: two checked nucleotides will not be deleted)

or

☐ Insert between two checked nucleotides

☐ 1 a ☐ 2 t ☐ 3 a ☐ 4 g ☐ 5 a ☐ 6 t ☐ 7 a ☐ 8 t ☐ 9 a

☐ 10 g ☐ 11 a ☐ 12 c ☐ 13 a ☐ 14 g ☐ 15 a ☐ 16 c ☐ 17 a ☐ 18 t

☐ 19 a ☐ 20 c ☐ 21 a ☐ 22 t ☐ 23 c ☒ 24 a ☐ 25 g ☐ 26 a ☐ 27 c

☐ 28 a ☐ 29 t ☐ 30 a ☐ 31 t ☐ 32 a ☐ 33 c ☐ 34 a ☐ 35 g ☐ 36 g

☐ 37 a ☐ 38 g ☐ 39 g ☐ 40 a ☐ 41 g ☐ 42 g ☐ 43 a ☐ 44 t ☐ 45 a

☐ 46 c ☐ 47 a ☐ 48 c ☐ 49 a ☐ 50 g ☐ 51 a ☐ 52 c ☐ 53 a ☐ 54 t

☐ 55 c ☐ 56 a ☐ 57 t ☐ 58 a ☐ 59 c ☐ 60 a ☐ 61 g ☐ 62 c ☐ 63 a

☐ 64 g ☐ 65 a

Finally, obtain your designed primer sequences.

[Design Primers](#)

Clear this form to load a new sequence:

[Clear Input](#)

19

PRODUCTS & SERVICES CROSSLAB SOLUTIONS SUPPORT LIBRARY TRAINING & EVENTS BUY

☐ Insert between two checked nucleotides

☐ 1 a ☐ 2 t ☐ 3 a ☐ 4 g ☐ 5 a ☐ 6 t ☐ 7 a ☐ 8 t ☐ 9 a

☐ 10 g ☐ 11 a ☐ 12 c ☐ 13 a ☐ 14 g ☐ 15 a ☐ 16 c ☐ 17 a ☐ 18 t

☐ 19 a ☐ 20 c ☐ 21 a ☐ 22 t ☒ 24 a ☐ 25 g ☐ 26 a ☐ 27 c

☐ 28 a ☐ 29 t ☐ 30 a ☐ 31 t ☐ 32 a ☐ 33 c ☐ 34 a ☐ 35 g ☐ 36 g

☐ 37 a ☐ 38 g ☐ 39 g ☐ 40 a ☐ 41 g ☐ 42 g ☐ 43 a ☐ 44 t ☐ 45 a

☐ 46 c ☐ 47 a ☐ 48 c ☐ 49 a ☐ 50 g ☐ 51 a ☐ 52 c ☐ 53 a ☐ 54 t

☐ 55 c ☐ 56 a ☐ 57 t ☐ 58 a ☐ 59 c ☐ 60 a ☐ 61 g ☐ 62 c ☐ 63 a

☐ 64 g ☐ 65 a

Finally, obtain your designed primer sequences.

[Design Primers](#)

Clear this form to load a new sequence:

[Clear Input](#)

Primer sequences:

Primer Name	Primer Sequence (5' to 3')
s24c_	5'-atctctctctgatatgctgcgtagtgctgctctatc-3' 5'-gatatagacagacataccagacatacaggaggat-3'

Oligonucleotide information:

Primer Name	Length (nt.)	Tm	Duplex Energy at 68 °C	Energy Cost of Mismatches
s24c_	41	78.60°C	-45.55 kcal/mole	3.05%
	41	78.60°C	-49.02 kcal/mole	3.85%

Primer-template duplexes:

Primer Name	Primer-Template Duplex
s24c_	gatatagacagacataccagacatacaggaggagatac 3'-ctatatctgtctgtatgttaggtgtatgtctctctcta-5' 5'-gatatagacagacataccagacatacaggaggat-3' ????ctatatctgtctgtatgttaggtgtatgtctctctctatgt

20

General Primer Design Principles

- **PCR Steps:** Denature (95 °C), anneal (60 °C), extend (70 °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

21

Example: General Approach

Exact and Single Base Mismatch DNA Thermodynamics

Primary Sequence: 5' to 3'; Target Sequence: 3' to 5'

Additional Target Base

5'- CGAAGAACAGGAAGCGGAATTTAAAGAAG
 |||||
 3'- GCTTCTTG T CCTTCGCCTTAAATTCTTC

Additional Target Base

5'-
 3'-

A

Hybridization Temperature

☒ Use Exact Complement T_m
☐ User Defined

Target Concentration

IDT OligoAnalyzer Mismatch Mode – estimate T_m cost of non-matching bases

22

22

Other Calculators: GeneRunner

- Download (free) from <http://generunner.net/>
- Analysis → Oligo brings up window to the right
- Suggested options shown to the right
 - SantaLucia temperature & salt recommended
 - Check your specific dNTP, DNA, Mg²⁺ and salt concentrations

GeneRunner Oligo analysis window showing sequence input and calculated parameters. The 'Tm' value is circled in red.

Parameter	Value
Mol Wt	8547.6
Tm	62.8
Filter Tm	52.2
%GC Tm	53.4
GC+AT Tm	60.0
nMol(A260)	3.8
ug(A260)	32.2
%GC	42.9
dG	-33.6
dH	-262.7
dS	-527.0
3'-end dG	-5.5
dNTP con (milli Mol)	0.60
DNA con (nano Mol)	50.00
Salt con (milli Mol)	50.00
Divalent con (milli Mol)	1.50
3'-end len	7
Pal len	8
Stem len	3
Guidance	
Max Len	35
dG	-0.8

23

23

Other Calculators: Phusion Calculator

Phusion Calculator interface showing input fields and calculated values.

Field	Value
Product Group	Phusion
Polymerase/Kit	Phusion High-Fidelity DNA Polymerase (HF Buffer)
Primer Concentration (nM)	500
Primer 1	CAACTACTATCGCTGTTGATGCTGAAAC
Primer 2	GCATCAACAGCGATAGTAGTTGTTTCQ
Anneal at	67 °C
Primer 1	28 nt, 48% GC, Tm: 64 °C
Primer 2	28 nt, 48% GC, Tm: 66 °C

- Some polymerases have their own calculator specific to buffers, recommendations
- Phusion Polymerase found at NEB website (<https://tmcalculator.neb.com/#!/main>)

24

24

Think And Discuss

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

27

27

Example: Sequence Analysis of SH3 Mutants

- Step 1: Design Primers (for T22G)
 - Agilent Web Program (we'll do this)
- Step 2: Do experiments, get sequence of result
- Step 3: Check sequence to see if mutation was successful (we'll do this)

28

Think and Discuss

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

29

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

30