Module 1: Basics of DNA and DNA self-assembly

CSE590: Molecular programming and neural computation
Double-stranded DNA

Biological DNA stores hereditary information

Width of the double helix: 2nm
Distance between base pairs: 0.34 nm

Carsonella rudii (smallest non-viral genome): 160,000 bp
Human genome: 3.2 Billion bp
Lungfish (largest vertebrate genome): 130 Billion bp
DNA nucleotides

- Phosphate group
- Sugar
- Nitrogenous base (adenine)

- Adenine
- Guanine
- Cytosine
- Thymine

Purines
Pyrimidines
DNA directionality

(a)

5’ end

Phosphodiester bond

3’ end

Phosphodiester bond

A DNA strand

= 5’-CAG-3’

The strand sequence
Watson Crick base pairing

Adenine (A) and Thymine (T) form a base pair because they have opposite charges and dissolve in water.

Guanine (G) and Cytosine (C) form a base pair because they have opposite charges and dissolve in water.

Each phosphate group links the 3' carbon of one sugar to the 5' carbon of the next sugar along the backbone.

The strands both run in a 5'-to-3' direction—they are antiparallel.

Plains of complementary bases form hydrogen bonds that hold the two strands of the DNA double helix together.

A-T pairs have two hydrogen bonds.

C-G pairs have three hydrogen bonds.
The double helix consists of **two DNA strands with complementary sequences** (base pairs: A:T, C:G) and with opposite orientation.
DNA can be commercially synthesized

**Single-stranded DNA with any sequence can be commercially synthesized**
length up to ~200 bases, cost: ~50 cents/base, 1 nMole \((10^{15})\) per order, same-day synthesis

**DNA sequence 1:**
5’ - ATTCAGATCCACCCAAAGAG-3’

**DNA sequence 2:**
5’ - CTCTTTGGGTTCCCAAATGT-3’

**DNA sequence 3:**
5’ - ACATTTGGGAGGATCTGAAT-3’
Find the reverse complement of the following sequences:

5′ -AAAAA-3′

5′ -AACCC-3′

5′ -CTGGACTAGAATT-3′
DNA hybridization (binding)

What happens when these strands are mixed in a test tube?

1: 5′-CACACACA-3′
2: 5′-TTTTTTGTGTGTG-3′
3: 5′-GTGTGTGT-3′
DNA hybridization (binding)

What happens when these strands are mixed in a test tube?

1: 5′-TTTTTT-3′

2: 5′-CACACACANNNNNNAGAGAGAG-3′
DNA hybridization (binding)

What happens when these strands are mixed in a test tube?

1: 5’-ATTCAGATCCACCCAAAAGAG-3’

2: 5’-CTCTTTGGGTTCACAAATGT-3’

3: 5’-ACATTTGGGAGGATCTGAAT-3’
DNA hybridization (binding)

test tube with reaction buffer (water + salt)
DNA hybridization (binding)

test tube with reaction buffer (water + salt)

Single-stranded DNA is flexible
Complementary single-stranded domains bind (hybridize) to each other. Formation of base pairs is energetically favorable and drives the reaction forwards.
DNA is programmable: sequence determines interactions

test tube with water + salt, $10^9$ copies of each molecule
Are we really doing this by hand?

NUPACK nucleic acid package

NUPACK is a growing software suite for the analysis and design of nucleic acid systems.

The NUPACK web application currently enables:

- **Analysis**: thermodynamic analysis of dilute solutions of interacting nucleic acid strands (demos).
- **Design**: single-state and multi-state sequence design for interacting nucleic acid strands (demos).
- **Utilities**: evaluation, display, and annotation of equilibrium properties of a complex of nucleic acid strands (demos).

NUPACK algorithms are formulated in terms of nucleic acid secondary structure. In most cases, pseudoknots are excluded from the structural ensemble.

You are welcome to use NUPACK for your research. Please cite NUPACK algorithms and the NUPACK web application appropriately.
**NUPACK**

**nucleic acid package**

**Analysis**

**Input**

Nucleic acid type:  
- RNA
- DNA

Temperature: 25°C

Number of strand species: 3

Maximum complex size: 3 strands

**Strand species**

<table>
<thead>
<tr>
<th>Strand</th>
<th>Sequence</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>strand1</td>
<td>ATTCAGATCCAGCAGAACAGC</td>
<td>1 μM</td>
</tr>
<tr>
<td>strand2</td>
<td>CTCCTTGGGTTCGAAATGT</td>
<td>1 μM</td>
</tr>
<tr>
<td>strand3</td>
<td>ACATTGGCAGATCTGAAAT</td>
<td>1 μM</td>
</tr>
</tbody>
</table>

**Advanced options**

Email address: [Enter email]
Ensemble pair fractions

Equilibrium concentrations

- strand1-strand2-strand3: 0.88 μM
- strand1-strand3: 0.12 μM
- strand2: 0.12 μM
- strand2-strand2: 0.0011 μM

Histogram filters

Change strand concentrations
MFE structure at 25.0 C

Free energy of secondary structure: -38.48 kcal/mol
Free energy of ordered complex (-kT log Q): -39.90 kcal/mol
**Amounts and concentrations**

<table>
<thead>
<tr>
<th>Properties</th>
<th>Amount Of Oligo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tm (50mM NaCl)</td>
<td>60.1 °C</td>
</tr>
<tr>
<td>GC Content</td>
<td>50. %</td>
</tr>
<tr>
<td>Molecular Weight</td>
<td>8,129.1</td>
</tr>
<tr>
<td>Dimensions</td>
<td>2.50 mg</td>
</tr>
<tr>
<td>nmoles/OD260</td>
<td>4.0</td>
</tr>
<tr>
<td>ug/OD260</td>
<td>32.6</td>
</tr>
<tr>
<td>Ext. Coefficient</td>
<td>249,100 L/(mole-cm)</td>
</tr>
</tbody>
</table>

**Secondary Structure Calculations**

- Lowest folding free energy (kcal/mole): -0.05 at 25 °C
- Strongest Folding Tm: 25.9 °C
- Secondary structure should not affect yield or purity for this oligo.

**Oligo Base Types**

- RNA bases: 19
- Chimeric DNA bases: 7

**Modifications And Services**

- RNase Free HPLC Purification: 1

**Shipped To**

GEORGE SEEUG
UNIVERSITY OF WASHINGTON-BIOCHEM
1959 NE PACIFIC ST
SEATTLE, WA 98195
USA

**Disclaimer**

See on reverse page notes (I), (II) & (III) for usage, label license, and product warranties.

The calculated amount of Oligo is **11.6 nMoles**.
What’s a Mole (unit)? $6.02 \times 10^{23}$ molecules

$11.6$ nMoles $= 11.6 \times 10^{-9} \times 6.02 \times 10^{23}$ molecules

DNA reactions occur in an aqueous solution and it is convenient to think about concentrations rather than amounts.

Concentration = Number/Volume (Unit: M, Molar=Mole/liter)

How much water do you need to add to $11.6$ nMoles of DNA to get a $100$ uM (micro Molar) concentration?
Outlook: Designing DNA structures

So far, we analyzed sequence that were already given to us. But how can we design the sequences that correspond to a target structure?