Cooperative Transitions I: DNA hybridization

\[ K = \frac{[AA']}{[A][A']} = \frac{C_{\text{dim}}}{C_A C_A'} \]

**Example**

5'-ATAGCA-3' + 5'-TGCTAT-3'

How do we predict \( T_m \)? At melting \( T \) (\( T_m \)), \( \frac{1}{2} \) oligomers are hybridized.

Know from experiments (and logic) that process can be broken into 2 steps:

1. **Initiation** - formation of 1st bond

   \[ T^\circ \rightarrow E^\circ \quad \Delta G_{\text{init}}^\circ \]

2. **Pairing** (sequentially) of remaining bonds

   \[ \Delta G_1 \rightarrow \Delta G_2 \rightarrow \Delta G_3 \rightarrow \Delta G_4 \]
How to Calculate ΔG° (and thus obtain "K")?

Presume that

\[ ΔG° = ΔG°_{init} + Σ ΔG°_{bonds} \]

Step 1: sum of each bond in

Step 2: A-T

Since we know all possible base pairs, G-C can we make a table of ΔG° values for each pair?

Possible ways to parse ΣΔG° bonds

1. Each pair has a ΔG° value independent of environment

\[
\begin{align*}
\text{A} & \quad \text{T} & \quad \text{A} & \quad \text{G} & \quad \text{C} \\
\text{A} & \quad \text{T} & \quad \text{G} & \quad \text{C} \\
\text{T} & \quad \text{A} & \quad \text{T} & \quad \text{C} & \quad \text{G} \\
\text{T} & \quad \text{A} & \quad \text{C} & \quad \text{T} & \quad \text{G} \\
\text{A} & \quad \text{T} & \quad \text{C} & \quad \text{T} & \quad \text{G} \\
\end{align*}
\]

\[ ΔG° = ΔG°_{init} + ΔG°_{A-T} + ΔG°_{A-C} + ΔG°_{A-G} \]

2. Experimental validation

ΔG° for a pair depends on the "nearest neighbor" pair formed just before it

See handout with values for the 10 possible pairs at 37°C, 1M NaCl

\[ ΔG° = 8.1 - 3.7 - 2.4 - 5.4 - 9.3 - 6.0 = -18.7 \text{kJ} \text{mol}^{-1} \]
From each $\Delta G^0$, can calculate $\Delta S^0$, $\Delta H^0$ (see Table)

For this example

\[
\Delta H^0 = 0.8 \cdot 30.2 - 30.2 - 33.7 - 41 - 35.6 = -168.9 \frac{KJ}{mol}
\]

\[
\Delta S^0 = -23.4 - 85.4 - 89.2 - 87.9 - 102.2 - 95 = -48.3 \frac{J}{mol-K}
\]

Note $\Delta S^0 \equiv \frac{J}{mol-K} \not\equiv \frac{KJ}{mol-K}$

How does the equilibrium depend on $T$?
It is reasonable (and backed up by experiment) to assume that $\Delta H^0$ and $\Delta S^0$ are independent of $T$ \(\Rightarrow\) can use values from Table for $T \neq 37^\circ C$.

Then $\Delta G^0 = \Delta G^0$ at $T$ of interest

\[
\Delta G^0 = \Delta H^0 - T \Delta S^0 = -RT \ln K = -RT \ln \frac{C_{A_{\text{dim}}}}{C_{A} C_{A'}}
\]

What we usually want is to know the fraction dimerized as function of $T$ (melting curve).

Some notation

$C_{A,0} = \text{initial concentration of } A$

$C_{A',0} = \text{initial concentration of } A'$

If $C_{A0} \neq C_{A',0}$, designate $C_{A,0}$ as "limiting"

(i.e. $C_{A0} < C_{A',0}$)

Then maximum number of dimers, $C_{\text{dim}, \text{max}}$, is

$C_{\text{dim}, \text{max}} = C_{A0}$
call the fraction of total possible dimers \( f \), i.e.
\[
f = \frac{\text{Cd}}{\text{C}_{\text{dim.,max.}}} = \frac{\text{Cd}}{\text{C}_{\text{a,0}}}
\]

and note \( \text{C}_4 = \text{C}_{a,0} - \text{Cd} \), \( \text{C}_{4i} = \text{C}_{a,0} - \text{C}_{\text{dim}} \)

We can thus write
\[
K = \frac{\text{Cd}}{[\text{C}_{a,0} - \text{Cd}][\text{C}_{a,0} - \text{Cd}]} = \frac{\text{Cd}}{\text{C}_{a,0}^2 \left[ 1 - \frac{1}{\text{C}_{a,0}} \right] \left[ \frac{\text{C}_{a,0}}{\text{C}_{a,0}} - \frac{\text{Cd}}{\text{C}_{a,0}} \right]}
\]

\[
K = \frac{f}{\text{C}_{a,0} \left[ 1 - f \right] \left[ \frac{\text{C}_{a,0}}{\text{C}_{a,0}} - f \right]}
\]

Special but common case
\( \text{C}_{a,0} = \text{C}_{a,0} \Rightarrow \) equal oligos at start

More notation (used in the literature)
\( \text{C}_T = \text{C}_{a,0} + \text{C}_{a,0} = 2\text{C}_{a,0} \)
\( \text{C}_{\text{dim.,max.}} = \frac{1}{2}\text{C}_T \)
Then
\[
K = \frac{f^{1/2}}{\text{C}_T \left( \frac{1}{2} - \frac{f}{2} \right)^2}
\]

and
\[
\Delta G^0_T = \Delta H^0 - T \Delta S^0 = -RT \ln \left( \frac{f^{1/2}}{\text{C}_T \left( \frac{1}{2} - \frac{f}{2} \right)^2} \right)
\]

We can solve to get \( T(f) \), Temp at which a given \( f \) occurs.

\[
T_f = \frac{\Delta H^0}{\Delta S^0 + RT \ln \left( \frac{f^{1/2}}{\text{C}_T \left( \frac{1}{2} - \frac{f}{2} \right)^2} \right)} = \frac{\Delta H^0}{\Delta S^0 + RT \ln \left( \frac{f^{1/2}}{\text{C}_T \left( \frac{1}{2} - \frac{f}{2} \right)^2} \right)}
\]
Define $T_m$ as $T$ where $f = 0.5$

Analysis of $T(f)$ for our example compared to oligomers twice as long shows that longer oligomers have higher, sharper $T_m$.

<table>
<thead>
<tr>
<th>Oligomer</th>
<th>$\Delta H^\circ$ KJ/mol</th>
<th>$\Delta S^\circ$ KJ/mol-K</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATAGC TATCGA &quot;Short&quot;</td>
<td>-168.9</td>
<td>-0.483</td>
</tr>
<tr>
<td>ATAGCATAAGC TATCGATATGC &quot;Long&quot;</td>
<td>-336.2</td>
<td>-0.919</td>
</tr>
</tbody>
</table>

Melting Curves for Short and Long DNA Oligomers ($C_T = 1$ mM)

Figure by MIT OCW.
Melting Curve Analysis To Detect Mismatches

Figure by MIT OCW.
Mismatch has very high $\Delta G^\circ$ cost

Figure by MIT OCW.