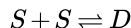


21.3: DNA Hybridization

To illustrate the use of statistical thermodynamics to describe binding, we discuss simple models for the hybridization or melting of DNA. These models are similar to our description of the helix–coil transition in their approach. These do not distinguish the different nucleobases, only considering nucleotides along a chain that are paired (bp) or free (f).

Consider the case of the pairing between self-complementary oligonucleotides.



S refers to any fully dissociated ssDNA and D to any dimer forms that involve two strands which have at least one base pair formed. We can then follow expressions for monomer–dimer equilibria above. The equilibrium constant for the association of single strands is

$$K_a = \frac{c_D}{c_S^2} \quad (21.3.1)$$

This equilibrium constant is determined by the concentration-dependent free-energy barrier for two strands to diffuse into contact and create the first base pair. If the total concentration of molecules present is either monomer or dimer, the form is

$$C_{TOT} = c_S + 2c_D \quad (21.3.2)$$

then the fraction of the DNA strands in the dimer form is

$$\theta_D = \frac{2c_D}{C_{tot}} \quad (21.3.3)$$

and eq. (10) leads to

$$\theta_D = 1 + (4K_a C_{tot})^{-1} - \sqrt{(1 + (4K_a C_{tot})^{-1})^2 - 1} \quad (21.3.4)$$

We see that at the total concentration, which results in a dimer fraction $\theta_D = 0.5$, the association constant is obtained from $K_a = (9C_{tot})^{-1}$. This is a traditional description of the thermodynamics of a monomer–dimer equilibrium.

We can calculate K_a from the molecular partition functions for the S and D states:

$$K_a = \frac{q_D}{q_S^2}$$

Different models for hybridization will vary in the form of these partition functions. For either state, we can separate the partition function into contributions from the conformational degrees of freedom relevant to the base-pairing and hybridization, and other degrees of freedom, $q_i = q_{i,conf} q_{i,ext}$. Assuming that the external degrees of freedom will be largely of an entropic nature, we neglect an explicit calculation and factor out the external degrees of freedom by defining the variable γ :

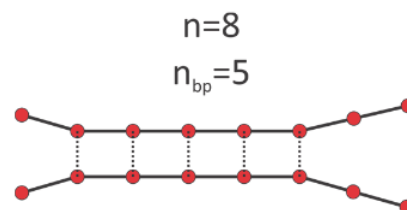
$$\gamma = \frac{q_{D,ext} C_{tot}}{q_{S,ext}^2} \quad (21.3.5)$$

then

$$\theta_D = 1 + \frac{q_{S,int}^2}{4\gamma q_{D,int}} - \sqrt{\left(1 + \frac{q_{S,int}^2}{4\gamma q_{D,int}}\right)^2 - 1} \quad (21.3.6)$$

Short Oligonucleotides: The Zipper Model

For short oligonucleotide hybridization, a common (and reasonable) approximation is the single stretch model, which assumed that base-pairing will only occur as a single continuous stretch of base pairs. This is reasonable for short oligomers ($n < 20$) where two distinct helical stretches separated by a bubble (loop) are unlikely given the persistence length of dsDNA. The zipper model refers to the single-stretch case with “perfect matching”, in which only pairing between the bases in precisely sequence-



aligned DNA strands is counted. As a result of these two approximations, the only dissociated base pairs observed in this model appear at the end of a chain (fraying).

The number of bases in a single strand is n and the number of bases that are paired is n_{bp} . For the dimer, we consider all configurations that have at least one base pair formed. The dimer partition function can be written as

$$q_{D,int}(n) = \sigma \sum_{n_{bp}=1}^n g(n, n_{bp}) s^{n_{bp}}$$

$$= \sigma \sum_{n_{bp}=1}^n (n - n_{bp} + 1) s^{n_{bp}}$$

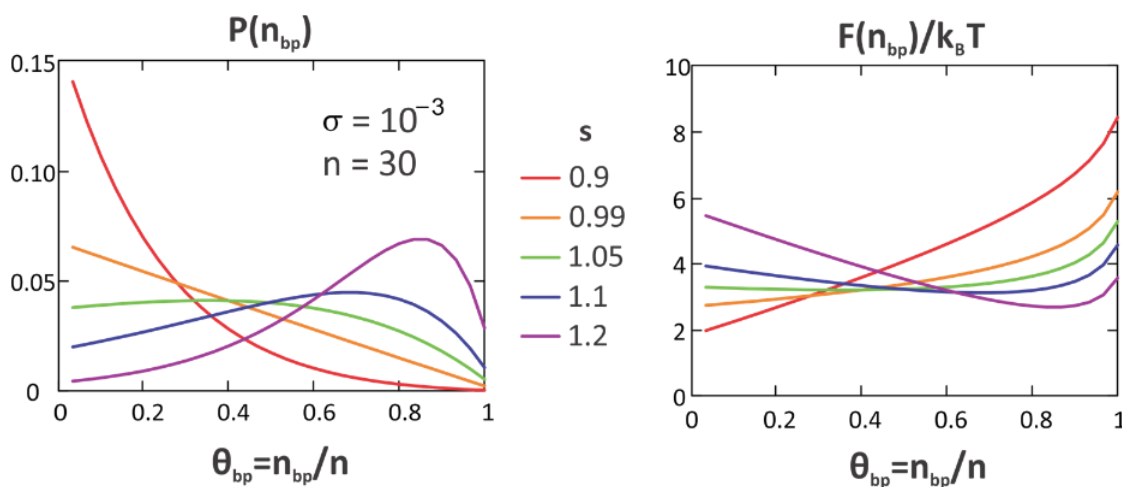
Here g is the number of ways of arranging n_{bp} continuous base pairs on a strand with length n ; σ is the statistical weight for nucleating the first base pair; and s is the statistical weight for forming a base pair next to an already-paired segment: $s = e^{-\Delta\epsilon_{bp}/k_B T}$. Therefore, in the zipper model, the equilibrium constant in eq. (23) between ssDNA and dimers involving at least one intact base pair is: $K_{zip} = \sigma s$. In the case of homogeneous polynucleotide chains, in which sliding of registry between chains is allowed: $q_{D,int}(n) = \sigma \sum_{n_{bp}=1}^n (n - n_{bp} + 1) s^{n_{bp}}$. The sum in eq. (27) can be evaluated exactly, giving

$$q_{D,int}(n) = \frac{\sigma s}{(s-1)^2} [s^{n+1} - (n+1)s + n] \quad (21.3.7)$$

In the case that $s > 1$ ($\Delta\epsilon_{bp} < 0$) and $n \gg 1$, $q_{D,int} \rightarrow \sigma s^n$. Also, the probability distribution of helical segments is

$$P_{bp}(n, n_{bp}) = \frac{(n - n_{bp} + 1) \sigma s^{n_{bp}}}{q_{D,int}} \quad 1 \leq n_{bp} \leq n \quad (21.3.8)$$

The plot below shows illustrations of the probability density and associated energy landscape for a narrow range of s across the helix-coil transition. These figures illustrate a duplex state that always has a single free-energy minimum characterized by frayed configurations.



In addition to the fraction of molecules that associate to form a dimer, we must also consider the fraction of contacts that successfully form a base pair in the dimer state

$$\theta_{bp} = \frac{\langle n_{bp} \rangle}{n}$$

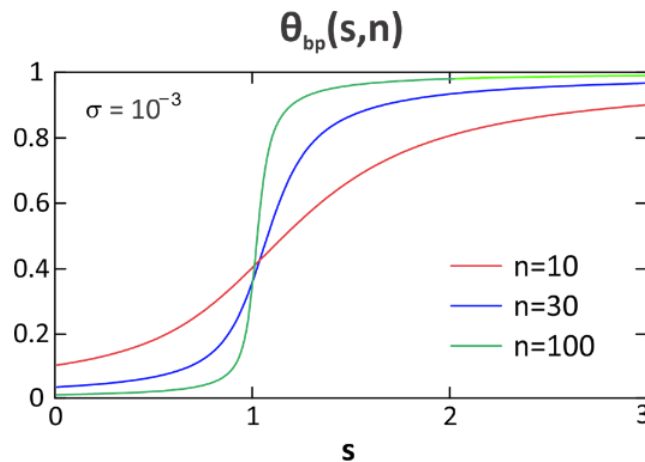
We can evaluate this using the identity

$$\langle n_H \rangle = \frac{s}{q} \frac{\partial q}{\partial s}$$

Using eq. (28) we have

$$\theta_{bp} = \frac{ns^{n+2} - (n+2)s^{n+1} + (n+2)s - n}{n(s-1)(s^{n+1} - s(n+1) + n)}$$

Similar to the helix–coil transition in polypeptides, θ_{bp} shows cooperative behavior with a transition centered at $s = 1$, which gets steeper with increasing n and decreasing σ .



Finally, we can write the total fraction of nucleobases that participate in a base pair as the product of the fraction of the DNA strands that are associated in a dimer form, and the average fraction of bases of the dimer that are paired.

$$\theta_{tot} = \theta_D \theta_{bp}$$

This page titled [21.3: DNA Hybridization](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Andrei Tokmakoff](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.