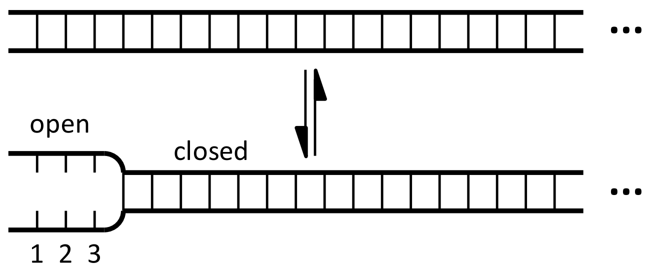
8

DNA can be modeled as two parallel polymer strands with links between the strands called *base pairs*. Each base pair can be in a closed state with energy 0 or in an open state with energy ε.

Consider a DNA molecule with N base pairs in thermal equilibrium at temperature T, as shown below. Thermal fluctuations can cause each base pair to open, leading to separation of the two strands. The molecule can open only from the left end, and only in *sequential order* (i.e. base pair s can open only if 1, 2... s–1 to the left of it are already open).

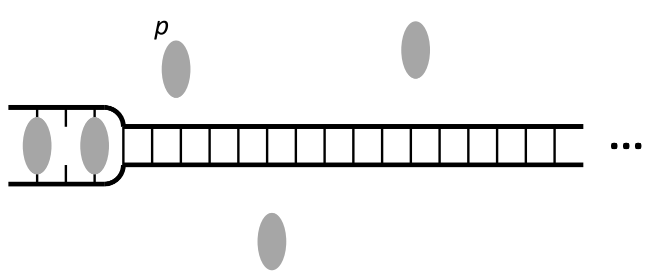


(a) Show that the partition function Z for this system is given by the following expression:

a. $Z = \frac{1-e^{-(N+1)\varepsilon/k_BT}}{1-e^{-\varepsilon/k_BT}}$

(b) In the limit that $N \to \infty$, determine the mean number $\langle n \rangle$ of open base pairs. Approximate your expression separately in the limits that $k_BT \gg \varepsilon$, and that $k_BT \ll \varepsilon$.

Next, consider the same DNA molecule now surrounded by a protein p at concentration c. Protein p can bind to the DNA *only at a site that is open*, as shown below. Assume each protein p can occupy no more than one base pair. The chemical potential for p is $\mu = \Delta + k_BT\ln(c/c_0)$, where c_0 and Δ are constants ($c_0 > 0$).



(c)

(d) Using the Grand canonical ensemble, in the limit that $N \to \infty$, determine the mean number of DNA-binding proteins $\langle p \rangle$ per DNA at a temperature T for a given concentration c.