Solutions to problem set 5

Problem 1

Estimates of molecular forces.

a) C-O bond:
$E = 84$ kcal/mol $= 84$ kcal/mol $\cdot$ 4.184 kJ/kcal $\cdot$ 1000 J/kJ / ($6 \cdot 10^{23}$/mol)
$= 5.9 \cdot 10^{-19}$ J
$F = E \,/\, \Delta x = 5.9 \cdot 10^{-19}$ J / (10^{-10} m) $= 5.9 \cdot 10^{-9}$ N $= 5.9$ nN

S-S bond:
$E = 3.6 \cdot 10^{-19}$ J
$F = 3.6$ nN
i.e. the rupture forces for covalent bond are in the nN range. For more information, see Michel Crandbois, Martin Beyer, Matthias Hauke Clausen-Schaumann, Hermann E. Gaub, *How Strong Is a Covalent Bond?*, *Science* (1999)

b) Non-covalent bonds in biological systems have to be stronger than $E = 4$ pN·nm $= 10^{-21}$ J and have to withstand forces larger than ≈ 4 pN·nm/1 nm $= 4$ pN, otherwise thermal fluctuations would constantly break them. At the same time, they are considerably weaker than covalent bonds with energies in the range of $E \approx 10^{-19}$ J and forces ≈ 1 nN. Therefore, typical rupture forces for non-covalent bonds are 10-100 pN and typical energies 10-100 pN·nm $\approx$ 2-20 times $k_B T$.

Problem 2

DNA overstretching transition.

a) Van Marmeren *et al.* argue that DNA melts during the overstretching transition, i.e. that the double-stranded (ds) DNA is converted to two single strands (ss = single-stranded). Quoting from the paper (last paragraph of "Conclusion"):

In conclusion, we have unveiled that, independent of the details of strand attachment, DNA overstretching unambiguously comprises a gradual conversion of dsDNA to ssDNA.

b) There are several lines of evidence for DNA melting during overstretching presented in the paper:

1) From fluorescence imaging of DNA stretched in optical tweezers, they show that YOYO binding (linearly) decreases when the DNA is overstretched (Figure

2). YOYO is an *intercalator* that is known to bind to B-form DNA and not to ssDNA. The authors note themselves that this experiment is not conclusive, though, since it is not known whether YOYO binds to S-DNA or not.

2) Using the same assay, they observe that fluorescently labeled mitochondrial single-stranded binding protein (mtSSB) starts binding upon overstretching the DNA and that the amount bound corresponds to the amount of DNA overstretched (Figure 3).

3) Finally, they perform two-color experiments where they label the (B-form) dsDNA with intercalating dyes (YOYO or POPO) and the single-stranded parts with fluorescently labeled mtSSB or RPA (another single-stranded DNA binding protein) (Figure 4).

c) The experiments reported by van Marmeren *et al.* convincingly indicate that the overstretching transition can involve DNA melting. However, they do not rule out that S-DNA formation can also occur. First, as the authors note at least partially themselves, the measurements with intercalating dyes are inconclusive in this regard, since it is not know whether or two what extent intercalating dyes bind S-DNA. If intercalators bind strongly only to B-form DNA, then loss of fluorescence only means that overstretched DNA is no longer B-form, but it could still be either melting or S-form. Even the experiments with the single-stranded binding proteins do not rule out S-DNA formation upon overstretching. Strictly speaking, it is not known whether they bind S-form DNA or not; this seems unlikely, though, from what us known structurally. More importantly, even if we assume (similar to the authors' implicit assumptions) that the single-stranded binding proteins only bind single-stranded DNA, it is important to realize that adding a binding partner for ssDNA will shift the thermodynamic equilibrium in favor of single-stranded DNA. So whatever equilibrium between melting and S-DNA formation occurs upon overstretching, the equilibrium will be shifted towards more melting in the presence if the single-stranded binding proteins.

It turns out that subsequent publications showed that both S-DNA formation and melting occur upon DNA overstretching and that the balance sensitively depends on solution conditions (salt concentration, temperature, etc.), GC content and pulling speed. See e.g. Bosaeus, *et al. PNAS* 2012 (`http://www.pnas.org/content/109/38/15179.full.pdf`), Zhang, *et al. PNAS* 2012 (`http://www.pnas.org/content/109/21/8103.full.pdf`), King *et al. PNAS* 2013 (`http://www.pnas.org/content/110/10/3859.full.pdf`)

Problem 3

Force-extension relationship for the 1D freely-jointed chain. We consider the 1D FJC model, with a two-state variable σ that takes on the value $\sigma_i = +1$ for each segment that points "forward" in the z-direction, along the external applied force, or

$\sigma_i = -1$ for segments that point "backwards", against the external force. The total extension is then given by

$$z = b \cdot \sum_{i=1}^{N} \sigma_i \tag{1}$$

To derive an expression for the average extension $\langle z \rangle$, we take the ensemble average, averaging over "states of the world" j, weighting the value that z takes on in each state, z_j by the probability of the state to occur p_j:

$$\langle z \rangle = \sum_j p_j \cdot z_j = \sum_{\sigma_1 = \pm 1} ... \sum_{\sigma_N = \pm 1} p(\sigma_1, ..., \sigma_N) \cdot z \tag{2}$$

The probability for a state with energy E_j to occur is given by its Boltzmann factor, properly normalized:

$$p_j = p(\sigma_1, ..., \sigma_N) = \frac{e^{-E_j/(k_B\,T)}}{Z} = \frac{e^{-(-f \cdot z)/(k_B\,T)}}{Z} = \frac{e^{(f \cdot b \cdot \sum_{i=1}^{N} \sigma_i)/(k_B\,T)}}{Z} \tag{3}$$

where the normalization Z is the partition function (i.e. the sum over all Boltzmann factors) and we have used the expression for the extension z from Equation 1.
Inserting the expression for the probabilities and for the length z into Equation 2, we get

$$\langle z \rangle = \sum_{\sigma_1 = \pm 1} ... \sum_{\sigma_N = \pm 1} \left(\frac{e^{(f \cdot b \cdot \sum_{i=1}^{N} \sigma_i)/(k_B\,T)}}{Z} \right) \cdot \left(b \cdot \sum_{i=1}^{N} \sigma_i \right) \tag{4}$$

which can be written short hand by using the "logarithm trick" (you can verify this by simply doing the derivative):

$$\langle z \rangle = k_B\,T \frac{\partial}{\partial f} \ln \left(\sum_{\sigma_1 = \pm 1} ... \sum_{\sigma_N = \pm 1} e^{(f \cdot b \cdot \sum_{i=1}^{N} \sigma_i)/(k_B\,T)} \right) \tag{5}$$

We notice that the argument of the logarithm is just the product of N independent and identical factors:

$$\langle z \rangle = k_B\,T \frac{\partial}{\partial f} \ln \left(\left(\sum_{\sigma_1 = \pm 1} e^{(f \cdot b \cdot \sigma_1)/(k_B\,T)} \right) \cdot ... \cdot \left(\sum_{\sigma_N = \pm 1} e^{(f \cdot b \cdot \sigma_N)/(k_B\,T)} \right) \right) \tag{6}$$

This allows us to write it as a simply product and "pull down" the factor N:

$$\langle z \rangle = k_B\,T \frac{\partial}{\partial f} \ln \left(e^{(f \cdot b)/(k_B\,T)} + e^{(-f \cdot b)/(k_B\,T)} \right)^N = k_B\,TN \frac{\partial}{\partial f} \ln \left(e^{(f \cdot b)/(k_B\,T)} + e^{(-f \cdot b)/(k_B\,T)} \right) \tag{7}$$

Finally, we carry out the derivative with respect to f; to make the results look "pretty", we can additionally use a trigonometric identity:

$$\langle z \rangle = N \cdot b \frac{e^{(f \cdot b)/(k_B\,T)} - e^{(-f \cdot b)/(k_B\,T)}}{e^{(f \cdot b)/(k_B\,T)} + e^{(-f \cdot b)/(k_B\,T)}} = N \cdot b \cdot \tanh(f \cdot b / k_B\,T) \tag{8}$$