Summary of Model 3

Continues on from Model 2.5 but taking into account new insights. Firstly, while we keep the energy models in place we are mostly looking at cases with all energies 0 except perhaps ϵ_b. This reflects a shift from looking at the known equilibrium effects like cooperative binding to non-equilibrium driven effects. Second when designing the rates we want to incorporate some mechanism so that rates of different processes have (by default) roughly the same scales so that they can then be brought out of this equality in a controlled manner. This reflects how before once we added an ϵ_b then we got factors of positive exponentials on conformational change meaning with all $r_? \sim 1$ we got wildly different rates. We also embrace more fully that we're really only looking at $C = 2$ at the moment so it won't be mentioned all the time.

1 Energies

As before, to summarize:

$$E(\{c_i\}, \{b_i\}) \simeq \sum_i E_M(c_i, b_i) + \frac{1}{2} \sum_i E_I(c_{i-1}, c_i) + E_I(c_i, c_{i+1}) \tag{1}$$

where E_M is a $C \times (B+1)$ matrix defining the energies of each individual monomer according to its conformation and number of bound ligands. And E_I is a $C \times C$ matrix defining the monomer interaction energies, specifically $E_I(c_1, c_2)$ is the energy cost of having a monomer of conformation c_1 to the left of one in conformation c_2, hence the particular ordering in eq. (1). The model is achiral if E_I is symmetric. There is one caviat to eq. (1) which is why the $\simeq$ symbol is used and that is the problem of boundaries. Specifically, are we considering a single chain of monomers or a loop that joins its ends, eq. (1) is correct for a loop and can be easily corrected for a chain configuration.

Notably, we can also write eq. (1) as

$$E(\{c_i\}, \{b_i\}) \simeq \sum_i E_M(c_i, b_i) + \frac{E_I(c_{i-1}, c_i) + E_I(c_i, c_{i+1})}{2} = \sum_i E_i \tag{2}$$

where E_i are the energies associated with each monomer and its state.

1.1 General case

In the general case we can write

$$E_M \leftrightarrow \begin{pmatrix} 0 & \epsilon_{1,1} & \epsilon_{1,2} & \cdots \\ \epsilon_{2,0} & \epsilon_{2,1} & \epsilon_{2,2} & \cdots \\ \epsilon_{3,0} & \epsilon_{3,1} & \epsilon_{3,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad E_I \leftrightarrow \begin{pmatrix} 0 & \epsilon_{b,1} & \epsilon_{b,2} & \cdots \\ \epsilon_{b,1} & 0 & \epsilon_{b,B+1} & \cdots \\ \epsilon_{b,2} & \epsilon_{b,B+1} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3}$$

Restricting ourselves to the $C = 2$, achiral case we can immediately simplify to

$$E_M \leftrightarrow \begin{pmatrix} 0 & \epsilon_{T,1} & \epsilon_{T,2} & \cdots \\ \epsilon_{R,0} & \epsilon_{R,1} & \epsilon_{R,2} & \cdots \end{pmatrix} \qquad E_I \leftrightarrow \begin{pmatrix} 0 & \epsilon_b \\ \epsilon_b & 0 \end{pmatrix} \tag{4}$$

borrowing the tense (T) and relaxed (R) conformation labels from haemoglobin models where presumably $\epsilon_{R,0} \geq 0$ and $\epsilon_{T,i} \geq \epsilon_{R,i}$ for most of $i \neq 0$.

1.2 Simplest case

However, to further reduce the number of parameters we use

$$E_M \leftrightarrow \begin{pmatrix} 0 & \epsilon_t & 2\epsilon_t & \cdots \\ \Delta\epsilon_r & \epsilon_t - \Delta\epsilon_r & 2\epsilon_t - \Delta\epsilon_r & \cdots \end{pmatrix} \qquad E_I \leftrightarrow \begin{pmatrix} 0 & \epsilon_b \\ \epsilon_b & 0 \end{pmatrix} \tag{5}$$

where ϵ_t sets the overall energy of binding additional ligands and $\Delta\epsilon_r$ is a measure of how different the R state is.
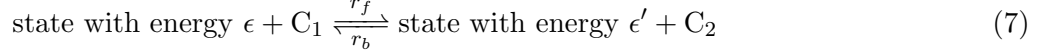
2 Equilibrium/Boltzmann Statistics

Firstly, defining our system as the polymer only (not any ligands or other chemicals floating around) its clear we are working in a Grand Canonical Ensamble. Thus for each microstate we are interested in what its energy is and how many ligands are bound in that microstate, denote these as E_α and N_α. Then the probabilities of microstates being occupied is given by their Gibbs factors so that

$$p_\alpha \propto \exp(-\beta(E_\alpha - \mu N_\alpha)) \tag{6}$$

with μ being the chemical potential of the ligand. This is a slightly problematic quantity as I'm not too sure how this fits in with the next section, however I suspect it should be kept as a separate thing as long as possible.

3   Full Transition Rates

We once again modify the same recipe from before, however I will go through it in the more specific case here.

$$\text{state with energy } \epsilon + \mathrm{C}_1 \underset{r_b}{\overset{r_f}{\rightleftharpoons}} \text{state with energy } \epsilon' + \mathrm{C}_2 \tag{7}$$

we get

$$\frac{r_f}{r_b} = \exp\big(\beta(\epsilon + \mu_1 - \epsilon' - \mu_2)\big) = \frac{c_1}{c_2} \exp\big(\beta(\epsilon + \epsilon_1 - \epsilon' - \epsilon_2)\big) \tag{8}$$

with each $\mu_X = \epsilon_X + \mathrm{k_B} T \ln(c_X)$. The methods from earlier then follows by splitting the factors to get

$$r_f = r c_1 \exp\big(\beta(\theta_f(\epsilon + \epsilon_1) - (1 - \theta_b)(\epsilon' + \epsilon_2))\big) \tag{9}$$
$$r_b = r c_2 \exp\big(\beta(\theta_b(\epsilon' + \epsilon_2) - (1 - \theta_f)(\epsilon + \epsilon_1))\big) \tag{10}$$

so that higher concentrations of any chemicals increase the rates of those reactions using them (reasonable) but we can split the energetic contributions to the rates between the forward and backward rates using the dimensionless $\theta_?$ parameters.

The issue we were encountering with this was that once we had positive energetic contributions the rates rapidly increased and dominated everything else thus the general reaction rate scale r parameter didn't really hold up to its name. In order to combat this the idea is to take the above eqs. (9) and (10) and divide both by some equal measure not to affect detailed balance but to make sure both rates are effectively capped. This is both for convenient theory so that comparable r lead to comparable rates but also for physical reasons as in the real world reaction rates have a cap, the reaction cannot happen faster than a certain speed.

Going from eqs. (9) and (10) we simply divide both by the sum of those excluding the r. This has a very clear interpretation, it effectively sets the average of the two rates to be $\frac{r}{2}$ and is convenient to use (certainly much more than using their max). Thus we end up with

$$r_f = \frac{r}{d} c_1 \exp\big(\beta(\theta_f(\epsilon + \epsilon_1) - (1 - \theta_b)(\epsilon' + \epsilon_2))\big) \tag{11}$$
$$r_b = \frac{r}{d} c_2 \exp\big(\beta(\theta_b(\epsilon' + \epsilon_2) - (1 - \theta_f)(\epsilon + \epsilon_1))\big) \tag{12}$$
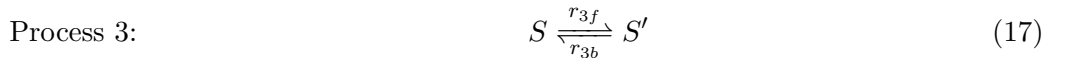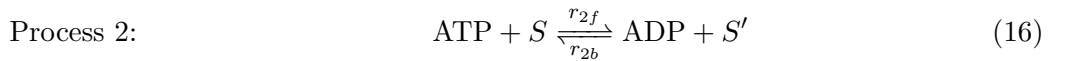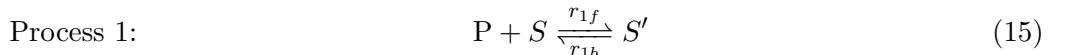$$\tag{13}$$

where

$$d = c_1 \exp\big(\beta(\theta_f(\epsilon + \epsilon_1) - (1 - \theta_b)(\epsilon' + \epsilon_2))\big) + c_2 \exp\big(\beta(\theta_b(\epsilon' + \epsilon_2) - (1 - \theta_f)(\epsilon + \epsilon_1))\big) \tag{14}$$

3.1   The concrete transition rates

So firstly, recap the reactions we have

$$\text{Process 1:} \qquad\qquad \mathrm{P} + S \underset{r_{1b}}{\overset{r_{1f}}{\rightleftharpoons}} S' \tag{15}$$

$$\text{Process 2:} \qquad\qquad \mathrm{ATP} + S \underset{r_{2b}}{\overset{r_{2f}}{\rightleftharpoons}} \mathrm{ADP} + S' \tag{16}$$

$$\text{Process 3:} \qquad\qquad S \underset{r_{3b}}{\overset{r_{3f}}{\rightleftharpoons}} S' \tag{17}$$

4

where the Ss are different in each process and denote different microstates of our polymer.

We then apply the general process above with the following summarized nuances taken from Model 2.5. For the energy of the system we only use the part of the energy that is associated with the monomer that is being affected, and only that part which can be changed. This means we consider only the independent monomer energies for phosphorylation and both that and the NN interaction energies for conformation changes. That means these energies will always be just an element of the E_M matrix along with an interaction contribution of 0, ϵ_b or $2\epsilon_b$. Additionally, we allow the overall reaction rates $r_?$ for the phosphorylation reactions to depend on the conformation of the affected monomer.

Finally, for the new divisors added in this model we don't explicitly write them out as they are very large but just denote them as $d(\cdots)$

Thus we arrive at

$$r_{1f}(c_i) = r_1(c_i)c_P \exp\Big(\beta(\theta_{1f}(\epsilon_i^S + \epsilon_P) - (1-\theta_{1b})\epsilon_i^{S'})\Big)/d(\cdots) \tag{18}$$

$$r_{1b}(c_i) = r_1(c_i)\exp\Big(\beta(\theta_{1b}\epsilon_i^{S'} - (1-\theta_{1f})(\epsilon_i^S + \epsilon_P))\Big)/d(\cdots) \tag{19}$$

$$r_{2f}(c_i) = r_2(c_i)c_{ATP}\exp\Big(\beta(\theta_{2f}(\epsilon_i^S + \epsilon_{ATP}) - (1-\theta_{2b})(\epsilon_i^{S'} + \epsilon_{ADP}))\Big)/d(\cdots) \tag{20}$$

$$r_{2b}(c_i) = r_2(c_i)c_{ADP}\exp\Big(\beta(\theta_{2b}(\epsilon_i^{S'} + \epsilon_{ADP}) - (1-\theta_{2f})(\epsilon_i^S + \epsilon_{ATP}))\Big)/d(\cdots) \tag{21}$$

$$r_{3f} = r_3\exp\Big(\beta(\theta_{3f}\epsilon_i^S - (1-\theta_{3b})\epsilon_i^{S'})\Big)/d(\cdots) \tag{22}$$

$$r_{3b} = r_3\exp\Big(\beta(\theta_{3b}\epsilon_i^{S'} - (1-\theta_{3f})\epsilon_i^S)\Big)/d(\cdots) \tag{23}$$

where the $r_?, c_?, \epsilon_P, \epsilon_{ATP}, \epsilon_{ADP}$ and $\theta_?$ are parameters.