Introduction into the Theory of Biological Nanomachines

Alexander S. Mikhailov/Holger Flechsig
Lectures 3 & 4
Molecular machines are single-molecule chemical engines! All of them represent various enzymes.

Therefore, we have to consider single-molecule kinetics and thermodynamics of enzymes.
Many reactions proceed only very slowly in their pure form. They can be however greatly accelerated when catalysts are added. The role of a catalyst is to lower the energy barrier for a reaction, thus strongly increasing its rate.

There are many inorganic catalysts. Any modern car is using the reaction of CO oxidation into CO$_2$ on a catalytic platinum surface.
Enzymes are single-molecule protein catalysts

The active center typically contains a metal ion. In its simplest version, the enzyme protein just provides mechanical support for the active ion and ensures an optimal local environment for the catalytic reaction event.

A catalytic reaction typically proceeds through formation of substrate-enzyme and product-enzyme complexes. There may be also several intermediate products, two different substrates, two different products etc. Protein conformations may also change depending on the ligand state.
Enzymes are single-molecule protein catalysts

The active center typically contains a metal ion. In its simplest version, the enzyme protein just provides mechanical support for the active ion and ensures an optimal local environment for the catalytic reaction event.

A catalytic reaction typically proceeds through formation of substrate-enzyme and product-enzyme complexes. There may be also several intermediate products, two different substrates, two different products etc. Protein conformations may also change depending on the ligand state.
Enzymes are single-molecule protein catalysts

The active center typically contains a metal ion. In its simplest version, the enzyme protein just provides mechanical support for the active ion and ensures an optimal local environment for the catalytic reaction event.

A catalytic reaction typically proceeds through formation of substrate-enzyme and product-enzyme complexes. There may be also several intermediate products, two different substrates, two different products etc. Protein conformations may also change depending on the ligand state.
Enzymes are single-molecule protein catalysts

The active center typically contains a metal ion. In its simplest version, the enzyme protein just provides mechanical support for the active ion and ensures an optimal local environment for the catalytic reaction event.

A catalytic reaction typically proceeds through formation of substrate-enzyme and product-enzyme complexes. There may be also several intermediate products, two different substrates, two different products etc. Protein conformations may also change depending on the ligand state.
Enzymes are single-molecule protein catalysts

The active center typically contains a metal ion. In its simplest version, the enzyme protein just provides mechanical support for the active ion and ensures an optimal local environment for the catalytic reaction event.

A catalytic reaction typically proceeds through formation of substrate-enzyme and product-enzyme complexes. There may be also several intermediate products, two different substrates, two different products etc. Protein conformations may also change depending on the ligand state.
Enzymes are single-molecule protein catalysts

The active center typically contains a metal ion. In its simplest version, the enzyme protein just provides mechanical support for the active ion and ensures an optimal local environment for the catalytic reaction event.

A catalytic reaction typically proceeds through formation of substrate-enzyme and product-enzyme complexes. There may be also several intermediate products, two different substrates, two different products etc. Protein conformations may also change depending on the ligand state.
Markov network models of enzymes

First we will consider a simple reaction scheme. A substrate binds to a free enzyme and forms a substrate-enzyme complex. Then, the substrate is converted into a product and thus a product-enzyme complex is formed. Finally, the product is released and the enzyme returns to its free form. All these reaction steps are generally reversible. Each transition is characterised by some probability rate.

\[ S + E \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P \]

There is only one enzyme molecule. There are many available substrate and product molecules and their number practically does not change when the enzyme operates.
At each time moment, only one state can be occupied. Sudden jumps between the states take place, with the jump probability per unit time determined by the respective reaction rate. There is no memory of the previous states.
At each time moment, only one state can be occupied. Sudden jumps between the states take place, with the jump probability per unit time determined by the respective reaction rate. There is no memory of the previous states.
At each time moment, only one state can be occupied. Sudden jumps between the states take place, with the jump probability per unit time determined by the respective reaction rate. There is no memory of the previous states.
At each time moment, only one state can be occupied. Sudden jumps between the states take place, with the jump probability per unit time determined by the respective reaction rate. There is no memory of the previous states.
At each time moment, only one state can be occupied. Sudden jumps between the states take place, with the jump probability per unit time determined by the respective reaction rate. There is no memory of the previous states.
Occupation probabilities of the states
\[ p_E(t), \ p_{ES}(t), \ p_{EP}(t), \ p_E(t) + p_{ES}(t) + p_{EP}(t) = 1 \]

Transition rates
\[ w_{E\rightarrow ES}, \ w_{ES\rightarrow EP}, \ w_{EP\rightarrow E} \]
\[ w_{E\leftarrow ES}, \ w_{ES\leftarrow EP}, \ w_{EP\leftarrow E} \]

\[ w_{E\rightarrow ES} = c_S v_{E\rightarrow ES}, \ w_{E\rightarrow EP} = c_P v_{E\rightarrow EP} \]
where \( c_S \) and \( c_P \) are (constant) substrate and product concentrations.

How the occupation probability of the substrate-enzyme state will change within a short time because of transitions between the states E and ES?

\[ \Delta p_{ES} = (w_{E\rightarrow ES} \Delta t)p_E - (w_{E\leftarrow ES} \Delta t)p_{ES} \Rightarrow \frac{\Delta p_{ES}}{\Delta t} = w_{E\rightarrow ES}p_E - w_{E\leftarrow ES}p_{ES} \]

This system represents the \textbf{master equation} for the considered Markov network.

\[ \frac{dp_{ES}}{dt} = w_{E\rightarrow ES}p_E - w_{E\leftarrow ES}p_{ES} + w_{ES\leftarrow EP}p_{ES} - w_{ES\rightarrow EP}p_{ES} \]

\[ \frac{dp_{EP}}{dt} = w_{ES\rightarrow EP}p_{ES} - w_{ES\leftarrow EP}p_{EP} + w_{E\rightarrow EP}p_E - w_{EP\rightarrow E}p_{EP} \]

\[ p_E(t) = 1 - p_{ES}(t) - p_{EP}(t) \]
\[ \begin{align*}
\frac{dp_{ES}}{dt} &= w_{E\rightarrow ES} p_E - w_{E\leftarrow ES} p_{ES} + w_{ES\leftarrow EP} p_{EP} - w_{ES\rightarrow EP} p_E \\
\frac{dp_{EP}}{dt} &= w_{ES\rightarrow EP} p_{ES} - w_{ES\leftarrow EP} p_{EP} + w_{EP\leftarrow E} p_E - w_{EP\rightarrow E} p_{EP} \\
p_E(t) &= 1 - p_{ES}(t) - p_{EP}(t)
\end{align*} \]

Suppose \( w_{E\rightarrow ES} = 1 \text{ ms}^{-1}, w_{E\leftarrow ES} = 0.1 \text{ ms}^{-1}, w_{ES\rightarrow EP} = 1 \text{ ms}^{-1}, w_{ES\leftarrow EP} = 5 \text{ ms}^{-1}, w_{EP\rightarrow E} = 1 \text{ ms}^{-1}, w_{EP\leftarrow E} = 0.1 \text{ ms}^{-1} \)

and that at the initial time moment \( t=0 \) the enzyme is in the state \( E \), that is \( p_E(0) = 1, \quad p_{ES}(0) = p_{EP}(0) = 0 \).

The system evolves to a steady state.

\[ \bar{p}_E = 0.17, \quad \bar{p}_{ES} = 0.71, \quad \bar{p}_{EP} = 0.12 \]

Occupation probabilities in the steady state give relative times that the enzyme stay in different chemical states.
But we can also determine other statistical properties of the enzyme.

For example, what is the probability distribution for its cycle times?

To determine this distribution, we assume that, after product release, the enzyme goes into a non-reactive state $E^*$.

\[
S + E \leftrightarrow ES \leftrightarrow EP \rightarrow E^* + P
\]

With the initial conditions $p_E(0) = 1$, $p_{ES}(0) = p_{EP}(0) = p_{E^*}(0) = 0$, the solution $p_{E^*}(T)$ gives us the probability that the enzyme has finished its cycle by time $T$.

Then, the probability that it will finish it turnover cycle within time $\Delta T$ will be $\rho(T)\Delta T$, where

\[
\rho(T) = \frac{dp_{E^*}}{dT}
\]

is the distribution of turnover times.
But we can also determine other statistical properties of the enzyme.

For example, what is the probability distribution for its cycle times?

To determine this distribution, we assume that, after product release, the enzyme goes into a non-reactive state $E^*$. 

$$S + E \Leftrightarrow ES \Leftrightarrow EP \rightarrow E^* + P$$

With the initial conditions $p_E(0)=1$, $p_{ES}(0)=p_{EP}(0)=p_{E^*}(0)=0$, the solution $p_{E^*}(T)$ gives us the probability that the enzyme has finished its cycle by time $T$. Then, the probability that it will finish it turnover cycle within time $\Delta T$ will be $\rho(T)\Delta T$, where

$$\rho(T) = \frac{dp_{E^*}}{dT}$$

is the distribution of turnover times.

Unusually long cycles
Now we will change the notations. We enumerate the states

"E" = 1, "ES" = 2, "EP" = 3

and denote the transition rates as

\[ w_{E \rightarrow ES} = k_{12}, w_{E \leftarrow ES} = k_{12}, w_{ES \rightarrow EP} = k_{32}, w_{ES \leftarrow EP} = k_{23}, w_{EP \rightarrow E} = k_{31}, w_{EP \leftarrow E} = k_{31} \]

We arrive at state 2 from state 1

\[
\frac{dp_1}{dt} = k_{12}p_2 - k_{21}p_1 + k_{13}p_3 - k_{31}p_1
\]
\[
\frac{dp_2}{dt} = k_{21}p_1 - k_{12}p_2 + k_{23}p_3 - k_{32}p_2
\]
\[
\frac{dp_3}{dt} = k_{32}p_2 - k_{23}p_3 + k_{31}p_1 - k_{13}p_3
\]

This is the general form of the master equation for Markov networks.
Detailed Balance conditions

The master equation can be also written as

\[
\frac{dp_i}{dt} = \sum_j (k_{ij}p_j - k_{ji}p_i) \quad \Rightarrow \quad \frac{dp_i}{dt} = \sum_j J_{ij}
\]

where \( J_{ij} = k_{ij}p_j - k_{ji}p_i \) is the probability flux between the states \( i \) and \( j \).

Suppose that the enzyme is in a small closed compartment. Then, the numbers of substrate and product molecules will change until the thermal equilibrium is reached. In this thermal equilibrium, the following conditions will be satisfied:

1. All probability fluxes vanish, \( \overline{J}_{ij} = 0 \).
2. The occupation probability of each state \( i \) is determined only by its Gibbs free energy \( g_i \) and the Boltzmann distribution holds,

\[
\overline{p}_i = \frac{1}{Z} \exp \left( -\frac{g_i}{k_B T} \right)
\]

What are the consequences of these conditions?
Detailed Balance conditions

Let us first consider a transition that does not involve binding or release of ligands (substrate or product). The detailed balance then implies that

\[ \overline{J}_{ij} = k_{ij} \bar{p}_j - k_{ji} \bar{p}_i = 0 \]

\[ \Rightarrow \frac{k_{ij}}{k_{ji}} = \frac{\bar{p}_j}{\bar{p}_i} = \exp \left( -\frac{g_i - g_j}{k_B T} \right) \]

and therefore

\[ \frac{k_{ij}}{k_{ji}} = \exp \left( -\frac{g_i - g_j}{k_B T} \right) \]

where \( g_i \) and \( g_j \) are Gibbs free energies of internal molecular states \( i \) and \( j \)

This condition is **general**, it holds not only at thermal equilibrium. The rates of forward and backward transitions between different molecular states **are not independent**!
Detailed Balance conditions

Now we consider the transitions that involve binding or release of a substrate or product,

$$E + S \rightleftharpoons ES$$

Because one free substrate molecule disappears after binding, the detailed balance condition in this case is

$$\frac{w_{E \rightarrow ES}}{w_{E \leftarrow ES}} = \exp \left( -\frac{g_{ES} - g_E - \mu_s}{k_BT} \right)$$

where $\mu_s$ is the chemical potential of free substrate molecules.

In a similar way, we obtain

$$\frac{w_{EP \rightarrow E}}{w_{EP \leftarrow E}} = \exp \left( -\frac{g_E + \mu_p - g_{EP}}{k_BT} \right)$$

where $\mu_p$ is the chemical potential of free product molecules.
Detailed Balance conditions

\[
g_E + \mu_p - g_{EP} = -k_B T \ln \left( \frac{w_{EP \rightarrow E}}{w_{EP \leftarrow E}} \right)
\]

\[
g_{EP} - g_{ES} = -k_B T \ln \left( \frac{w_{ES \rightarrow EP}}{w_{ES \leftarrow EP}} \right)
\]

\[
g_{ES} - g_E - \mu_s = -k_B T \ln \left( \frac{w_{E \rightarrow ES}}{w_{E \leftarrow ES}} \right)
\]

Summing these three equations, we get a thermodynamic identity:

\[
\mu_s - \mu_p = k_B T \ln \left( \frac{w_{EP \rightarrow E} w_{ES \rightarrow EP} w_{E \rightarrow ES}}{w_{EP \leftarrow E} w_{ES \leftarrow EP} w_{E \leftarrow ES}} \right)
\]

This is the special case of the general Schnakenberg identity for chemical Markov networks.
The energy landscape

If all reaction rates are known, we can use detailed balance conditions to reconstruct the energy landscape of the enzyme. After simple calculation, we find

\[
g_{EP} = (g_E + \mu_p) + k_B T \ln \left( \frac{w_{EP \rightarrow E}}{w_{EP \leftarrow E}} \right)
\]

\[
g_{ES} = (g_E + \mu_p) + k_B T \ln \left( \frac{w_{EP \rightarrow E}}{w_{EP \leftarrow E}} \right) + k_B T \ln \left( \frac{w_{ES \rightarrow EP}}{w_{ES \leftarrow EP}} \right)
\]

\[
\mu_s - \mu_p = k_B T \ln \left( \frac{w_{EP \rightarrow E}}{w_{EP \leftarrow E}} \right) + k_B T \ln \left( \frac{w_{ES \rightarrow EP}}{w_{ES \leftarrow EP}} \right) + k_B T \ln \left( \frac{w_{E \rightarrow ES}}{w_{E \leftarrow ES}} \right)
\]

Suppose \( v_{E \rightarrow ES} = 0.1 \text{ ms}^{-1} \text{mMol}^{-1} \), \( w_{E \rightarrow ES} = 0.1 \text{ ms}^{-1} \), \( w_{ES \rightarrow EP} = 1 \text{ ms}^{-1} \), \( w_{EP \rightarrow E} = 5 \text{ ms}^{-1} \), \( v_{EP \rightarrow E} = 1 \text{ ms}^{-1} \), \( v_{EP \leftarrow E} = 0.1 \text{ ms}^{-1} \text{mMol}^{-1} \), \( c_s = 10 \text{ mMol} \), \( c_p = 3 \text{ mMol} \),

\[
w_{E \rightarrow ES} = c_s v_{E \rightarrow ES} = 1 \text{ ms}^{-1} , w_{EP \leftarrow E} = c_p v_{EP \leftarrow E} = 3 \text{ ms}^{-1}
\]

\[
g_{EP} = (g_E + \mu_p) + 1.204 k_B T
\]

\[
g_{ES} = (g_E + \mu_p) - 0.405 k_B T
\]

\[
(g_E + \mu_s) = (g_E + \mu_p) + 1.897 k_B T
\]

\[
\mu_s - \mu_p = 1.897 k_B T
\]
The energy landscape

Suppose \( \nu_{E\rightarrow ES} = 0.1 \text{ ms}^{-1}\text{mMol}^{-1}, w_{E\leftarrow ES} = 0.1 \text{ ms}^{-1}, w_{ES\rightarrow EP} = 1 \text{ ms}^{-1}, w_{ES\leftarrow EP} = 5 \text{ ms}^{-1}, w_{EP\rightarrow E} = 1 \text{ ms}^{-1}, \nu_{EP\leftarrow E} = 0.1 \text{ ms}^{-1}\text{mMol}^{-1}, c_s = 10 \text{ mMol}, c_p = 3 \text{ mMol}, w_{E\rightarrow ES} = c_s \nu_{E\rightarrow ES} = 1 \text{ ms}^{-1}, w_{EP\leftarrow E} = c_p \nu_{EP\leftarrow E} = 3 \text{ ms}^{-1}\)

\[
\begin{align*}
g_{ES} - (g_E + \mu_p) &= 0.405 k_B T \\
g_{EP} - (g_E + \mu_p) &= 1.204 k_B T \\
\mu_s - \mu_p &= 1.897 k_B T
\end{align*}
\]

Chemical potential of a weak solutions of substrate and product molecules are

\[
\begin{align*}
\mu_s &= \mu_s^0 + k_B T \ln c_s \\
\mu_p &= \mu_p^0 + k_B T \ln c_p
\end{align*}
\]

where \( \mu_s^0 \) and \( \mu_p^0 \) are the chemical potentials at unit concentrations.

\[
\mu_s^0 - \mu_p^0 = \mu_s - \mu_p - k_B T \ln \frac{c_s}{c_p} = 0.693 k_B T
\]

\( S + E \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P \)
Entropy at nanoscales

Entropy is microscopically defined as (it is then also known as Shannon or information entropy)

\[ S = - \sum_i p_i \ln p_i \]

Microscopic entropy is measured in bits, 1 bit = ln2

What is the rate of change of the entropy?

\[
\frac{dS}{dt} = - \sum_i \frac{dp_i}{dt} \ln p_i - \sum_i p_i \frac{1}{p_i} \frac{dp_i}{dt} = - \sum_i \frac{dp_i}{dt} \ln p_i = - \sum_{ij} J_{ij} \ln p_i = - \frac{1}{2} \left( \sum_{ij} J_{ij} \ln p_i + \sum_{ij} J_{ji} \ln p_j \right) = \frac{1}{2} \sum_{ij} J_{ij} \left( \ln p_j - \ln p_i \right)
\]

\[
\frac{d}{dt} \sum_i p_i = 0 \quad \frac{dp_i}{dt} = \sum_j J_{ij} \quad J_{ji} = -J_{ij} \\
J_{ij} = k_{ij} p_j - k_{ji} p_i
\]
Second Law of Thermodynamics for microscopic open systems

\[
\frac{dS}{dt} = \frac{1}{2} \sum_{ij} J_{ij} \ln \frac{p_j}{p_i} = \frac{1}{2} \sum_{ij} \left[ J_{ij} \ln \left( \frac{k_{ji} p_j}{k_{ij} p_i} \right) - J_{ij} \ln \left( \frac{k_{ij}}{k_{ji}} \right) \right] = \sigma - h
\]

Entropy production

Entropy export
Entropy production and flow in a Markov network

\[ \frac{dS}{dt} = \sigma - h \]

The Second Law of thermodynamics for a microscopic system

\[ \sigma = \frac{1}{2} \sum_{ij} \sigma_{ij}, \quad h = \frac{1}{2} \sum_{ij} h_{ij} \]

Entropy production in transition \( i \nsim j \)

\[ \sigma_{ij} = J_{ij} \ln \left( \frac{k_{ij} p_j}{k_{ji} p_i} \right) \quad \sigma_{ij} \geq 0 \]

Entropy export (or import) in transition \( i \nsim j \)

\[ h_{ij} = J_{ij} \ln \left( \frac{k_{ij}}{k_{ji}} \right) \]

In a non-equilibrium state, each transition is accompanied by entropy production and by export or import of entropy. At thermal equilibrium, all fluxes vanish - therefore entropy is not produced, imported or exported in this state.

In the steady state \( \frac{dS}{dt} = 0 \) and therefore \( \sigma = h \). However, \( \sigma_{ij} \neq h_{ij} \)!
Entropy production and flow in an enzyme

\[ S + E \iff ES \iff EP \iff E + P \]

We can determine entropy production and its export in the steady state.

It can be shown that the flux is constant for all transitions in the steady state of this system.

\[ \bar{J}_{E,ES} = \bar{J}_{ES,EP} = \bar{J}_{EP,E} = J \quad J = w_{E \rightarrow ES} \overline{p}_E - w_{E \leftarrow ES} \overline{p}_{ES} \]

The flux yields the mean turnover rate of the enzyme.

\[
\begin{align*}
\frac{dS}{dt} &= \sigma - h = 0 \\
\sigma &= \sigma_{E,ES} + \sigma_{ES,EP} + \sigma_{EP,E} \\
h &= h_{E,ES} + h_{ES,EP} + h_{EP,E} \\
\sigma &= h
\end{align*}
\]

\[
\begin{align*}
\sigma_{E,ES} &= J \ln \left( \frac{w_{E \rightarrow ES} \overline{p}_E}{w_{E \leftarrow ES} \overline{p}_{ES}} \right), \\
\sigma_{ES,EP} &= J \ln \left( \frac{w_{E \rightarrow ES} \overline{p}_{ES}}{w_{ES \leftarrow EP} \overline{p}_{EP}} \right), \\
\sigma_{EP,E} &= J \ln \left( \frac{w_{EP \rightarrow E} \overline{p}_{EP}}{w_{EP \leftarrow E} \overline{p}_E} \right) \\
h_{E,ES} &= J \ln \left( \frac{w_{E \rightarrow ES}}{w_{E \leftarrow ES}} \right), \\
h_{ES,EP} &= J \ln \left( \frac{w_{ES \rightarrow EP}}{w_{ES \leftarrow E}} \right), \\
h_{EP,E} &= J \ln \left( \frac{w_{EP \rightarrow E}}{w_{EP \leftarrow E}} \right)
\end{align*}
\]
Entropy production and flow in an enzyme

\[ S + E \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P \]

Suppose \( \nu_{E \rightarrow ES} = 0.1 \text{ ms}^{-1} \text{mMol}^{-1} \), \( w_{E \rightarrow ES} = 0.1 \text{ ms}^{-1} \), \( w_{ES \rightarrow EP} = 1 \text{ ms}^{-1} \), \( w_{ES \rightarrow EP} = 5 \text{ ms}^{-1} \), \( w_{EP \rightarrow E} = 1 \text{ ms}^{-1} \), \( \nu_{EP \leftarrow E} = 0.1 \text{ ms}^{-1} \), \( c_s = 10 \text{ mMol} \), \( c_p = 3 \text{ mMol} \)

Turnover rate \( J = 0.081 \text{ ms}^{-1} \), turnover time \( T = 1 / J = 12.3 \text{ ms} \)

This enzyme generates (and exports) entropy at the rate \( \sigma = 0.223 \text{ bits/ms} \), or \( \sigma / J = 2.73 \text{ bits per its turnover time} \).
Entropy production and flow in an enzyme

\[ S + E \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P \]

Suppose \( \nu_{E \rightarrow ES} = 0.1 \text{ ms}^{-1} \text{mMol}^{-1} \), \( w_{E \leftarrow ES} = 0.1 \text{ ms}^{-1} \), \( w_{ES \rightarrow EP} = 1 \text{ ms}^{-1} \), \( w_{ES \leftarrow EP} = 5 \text{ ms}^{-1} \), \( w_{EP \rightarrow E} = 1 \text{ ms}^{-1} \), \( \nu_{EP \leftarrow E} = 0.1 \text{ ms}^{-1} \), \( c_s = 10 \text{ mMol} \), \( c_p = 100 \text{ mMol} \)

Under such conditions, the enzyme operates, on the average, in the opposite direction (converts product into substrate)

Flux \( J = -0.057 \text{ ms}^{-1} \), turnover time \( T = 1 / |J| = 17.5 \text{ ms} \)

This enzyme generates (and exports) entropy at the rate \( \sigma = 0.133 \text{ bits/ms} \), or \( \sigma / J = 2.32 \text{ bits per its turnover time} \).
Direction of enzymic reaction

\[ S + E \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P \]

\[ \mu_s > \mu_p \]
Direction of enzymic reaction

\[ S + E \Leftrightarrow ES \Leftrightarrow EP \Leftrightarrow E + P \]

Substrate chemostat \( \mu_s \) \hspace{5cm} Product chemostat \( \mu_p \)

\( \mu_s < \mu_p \)

\[ \mu_s = \mu_s^0 + k_B T \ln c_s \]
\[ \mu_p = \mu_p^0 + k_B T \ln c_p \]

Reaction direction can be controlled by changing substrate and/or product concentrations.

The direction is reversed (the flux vanishes, \( J = 0 \)) when

\[ \mu_s = \mu_p \implies \mu_s^0 + k_B T \ln c_s = \mu_p^0 + k_B T \ln c_p \]

\[ \ln \frac{c_s}{c_p} = \frac{\mu_p^0 - \mu_s^0}{k_B T} \]
The energy landscape

Suppose $v_{E\rightarrow ES} = 0.1 \text{ ms}^{-1}\text{mMol}^{-1}$, $w_{E\leftarrow ES} = 0.1 \text{ ms}^{-1}$, $w_{ES\rightarrow EP} = 1 \text{ ms}^{-1}$, $w_{ES\leftarrow EP} = 5 \text{ ms}^{-1}$, $w_{EP\rightarrow E} = 1 \text{ ms}^{-1}$, $v_{EP\leftarrow E} = 0.1 \text{ ms}^{-1}\text{mMol}^{-1}$, $c_s = 10 \text{ mMol}$, $c_p = 3 \text{ mMol}$, $w_{E\rightarrow ES} = c_s v_{E\rightarrow ES} = 1 \text{ ms}^{-1}$, $w_{EP\leftarrow E} = c_p v_{EP\leftarrow E} = 3 \text{ ms}^{-1}$

S + E $\rightleftharpoons$ ES $\rightleftharpoons$ EP $\rightleftharpoons$ E + P

$g_{ES} - (g_E + \mu_p) = 0.405 k_B T$

$g_{EP} - (g_E + \mu_p) = 1.204 k_B T$

$\mu_s^0 - \mu_p^0 = 0.693 k_B T$

The reaction direction becomes reversed at the concentration ratio

$$\frac{\tilde{c}_p}{\tilde{c}_s} = \exp\left(\frac{\mu_s^0 - \mu_p^0}{k_B T}\right) = e^{0.693} = 1.99$$

But suppose that the difference of chemical potentials of products and substrates (at unit concentrations) is large. Then the required reversion ratio becomes extremely large.

$$\mu_s^0 - \mu_p^0 = 20 k_B T \Rightarrow \frac{\tilde{c}_p}{\tilde{c}_s} = \exp\left(\frac{\mu_s^0 - \mu_p^0}{k_B T}\right) = e^{20} = 4.85 \times 10^8$$

Such reactions are practically irreversible.
This enzyme synthesises an essential amino acid tryptophan from serine and indole glycerol phosphate (IGP). It is employed by all bacteria, plants, yeast and molds, but not by the higher organisms (they get this amino acid with food). Its substrate IGP is present only in small amounts in a cell, and high efficiency is required. Its intermediate product indole is hydrophobic and can easily escape through the cell membrane. Therefore, its release into cytoplasm must be avoided.

Nature has found an elegant solution for these constraints. The entire synthesis with 13 elementary reaction steps is performed within a single enzyme molecule with two catalytic active centers, and the intermediate indole is channeled within the protein from one center to another.

This enzyme has been extensively experimentally investigated. Several reviews are available, with the authors describing tryptophan synthase as "an allosteric molecular factory" and "a channeling nanomachine". Recently, we have constructed a complete microscopic kinetic theory and investigated stochastic thermodynamics of such chemical nanomachine.

Figure 1. Tryptophan synthase as a chemical nanomachine. (A) Structure of the enzyme from Salmonella typhimurium (PDB code: 2J9X)\(^{22}\). (B) Reduced reaction scheme. The internal aldimine E(Ain) corresponds to the PLP cofactor without any bound ligand. It is therefore referred to as empty throughout the text. (C) Operation of the machine: Once substrates are bound (a) at both catalytic sites, IGP activates (b) the formation of A-A and the enzyme adopts the closed conformation. A-A activates (c) the cleavage of IGP and indole is channeled (d) to the β-site where it reacts (e) with A-A to give Q_3. Q_3 undergoes (f) further transformations that return of the enzyme to the open conformation where tryptophan and G3P are released (g).
Markov network of tryptophan synthase

Joint chemical states of two subunits

Reaction networks in individual subunits
Numerical stochastic simulations

Tryptophan synthase is a very slow enzyme. Its mean turnover time is 0.15 s. The cycles of duration of almost a second are also occasionally taking place.
Synchronization of intramolecular processes tryptophan synthase

How to characterise correlations in the states of two subunits?

For each state $i$ or $j$ of two subunits, we introduce binary variables $X_i(t)$ and $Y_j(t)$ such that

$X_i(t) = 1$ if $\alpha$-subunit is in state $i$ at time $t$, $X_i(t) = 0$ otherwise

$Y_j(t) = 1$ if $\beta$-subunit is in state $j$ at time $t$, $Y_j(t) = 0$ otherwise

The Pearson correlation coefficients for the states of the two subunits are determined as

$$c_{ij} = \frac{\langle X_i(t)Y_j(t) \rangle - \langle X_i(t) \rangle \langle Y_j(t) \rangle}{\sqrt{\langle X_i^2(t) \rangle - \langle X_i(t) \rangle^2} \sqrt{\langle Y_j^2(t) \rangle - \langle Y_j(t) \rangle^2}}$$

These coefficients take the maximum value

$c_{ij} = +1$ if $X_i(t) \equiv Y_j(t)$ (complete correlation)

and the minimum value

$c_{ij} = -1$ if $X_i(t) \equiv 1 - Y_j(t)$ (complete anti-correlation)

If there is no correlation, $c_{ij} = 0$.

**Pearson correlation coefficients between the states of two subunits**

**Strong intramolecular correlations develop, processes in two subunits become synchronised.**
There is also another way to characterise synchronisation.

Suppose that both subunits are initially (at \( t = 0 \)) in the empty states. We can run a stochastic simulation and determine times \( t_{IGP} \) and \( t_{Q_1} \) when substrates arrive into the subunits. By running many simulations, the mean square-root difference of these times

\[
\sigma_{IGP,Q_1} = \sqrt{\langle (t_{IGP} - t_{Q_1})^2 \rangle}
\]

can be found.

In a similar way, we can find the mean square-root difference

\[
\sigma_{\text{indole+G3P,A-A}} = \sqrt{\langle (t_{\text{indole+G3P}} - t_{A-A})^2 \rangle}
\]

of the times when both subunits are in the pre-channeling state.

Statistical analysis of simulation data yields

\[
\sigma_{IGP,Q_1} = 22 \text{ ms}, \quad \sigma_{\text{indole+G3P,A-A}} = 2.4 \text{ ms}
\]

While binding of substrates occurs at largely different rimes, both subunits almost simultaneously arrive at the pre-channeling states.
Energy landscape of tryptophan synthase

\[
\left(\mu_{\text{serine}}^0 + \mu_{\text{IGP}}^0\right) - \left(\mu_{\text{G3P}}^0 + \mu_{\text{tryptophan}}^0\right) = 20.46 \ k_B T
\]

The reaction is practically irreversible!

Because of such a difference, the enzyme is out of equilibrium and can operate as a chemical nanomachine. There is an extensive pattern of allosteric cross-regulation between the two subunits. The transitions empty ^-site G3P ^-site are blocked (i.e., the gate in the ^-subunit is closed) in the states A-A, A-A+indole, and Q^3 of the ^-site. The transitions IGP ^-site indole+G3P ^-site are blocked in the states empty, Q^1, Aex^2 of the ^-site. The rate of the transition IGP ^-site indole+G3P ^-site is enhanced by a factor of 27.7 in the state A-A of the ^-site.

The transitions Q^1 ^-site A-A and Q^3 ^-site Aex^2 (green) in the ^-site are blocked in the state empty of the ^-site. The transition Q^1 ^-site A-A (light green) in the ^-site is enhanced by a factor of 9.7 in the state IGP of the ^-site. The changes indole+G3P ^-site G3P and A-A ^-site indole+A-A corresponding to indole channeling from the ^-site to the ^-site occur simultaneously and represent a single stochastic transition.

To simplify the notations, the states of both subunits will be enumerated below. The assignment of integer numerical variables a and b to different states of the ^- and ^-subunits is given in Table I and introduced in Fig. 2.

FIG. 3. The kinetic Markov network of tryptophan synthase with numerical values of all transition rates in units of s^-1. The kinetic Markov network for tryptophan synthase has been constructed in the previous publication^52 where all transition rates have been determined from the available experimental data. In that publication, we have modeled, however, a typical experimental situation where product concentrations remain vanishingly small. Therefore, product binding events are displayed. The states within the colored boxes correspond to the closed conformation of the enzyme where the gates are closed and the ligands cannot arrive or be released. The numbers next to the arrows give the respective transition rates in units of s^-1. Note that the bottom and upper states (1, 1) in Fig. 3 are identical; they are shown separately only for convenience in the displayed network.

The values of entropy production are given in units of bit s^-1 next to the links between the states. Additionally, color coding of the links according to the corresponding entropy production is used.

Tryptophan synthase produces 27.79 bits of entropy per turnover cycle.

Because of such a difference, the enzyme is out of equilibrium and can operate as a chemical nanomachine. There is an extensive pattern of allosteric cross-regulation between the two subunits. The transitions empty \( \rightarrow \) IGP and G3P \( \rightarrow \) empty (magenta) in the -site are blocked (i.e., the gate in the -subunit is closed) in the states A-A, A-A+indole, and Q3 of the -site. The transitions IGP \( \rightarrow \) indole+G3P (light and dark blue) in the -site are blocked in the states empty, Q1, Aex2 of the -site. The rate of the transition IGP \( \rightarrow \) indole+G3P (light blue) in the -site is enhanced by a factor of 27.7 in the state A-A of the -site. The transitions Q1 \( \rightarrow \) A-A and Q3 \( \rightarrow \) Aex2 (green) in the -site are blocked in the state empty of the -site. The transition Q1 \( \rightarrow \) A-A (light green) in the -site is enhanced by a factor of 9.7 in the state IGP of the -site. The changes indole+G3P \( \rightarrow \) G3P and A-A \( \rightarrow \) indole+A-A (red) corresponding to indole channeling from the -site to the -site occur simultaneously and represent a single stochastic transition.

To simplify the notations, the states of both subunits will be enumerated below. The assignment of integer numerical variables \( a \) and \( b \) to different states of the - and -subunits is given in Table I and introduced in Fig. 2.

III. THE KINETIC MARKOV NETWORK

The stochastic single-molecule operation of tryptophan synthase can be seen as random wandering over a Markov network of its internal states. In this network, each binary internal state \((a, b)\) with \( a = 1, \ldots, 4 \) and \( b = 1, \ldots, 6 \) represents a possible combination of the individual states \( a \) and \( b \) of the - and -subunits. The network is shown in Fig. 3. The nodes of the network are different chemical states and arrows indicate the transitions between them. Additionally, ligand binding and release are displayed. The states within the colored boxes correspond to the closed conformation of the enzyme where the gates are closed and the ligands cannot arrive or be released. The numbers next to the arrows give the respective transition rates in units of s\(^{-1}\). Note that the bottom and upper states \((1, 1)\) in Fig. 3 are identical; they are shown separately only for convenience in the displayed network.

The kinetic Markov network for tryptophan synthase has been constructed in the previous publication where all transition rates have been determined from the available experimental data. In that publication, we have modeled, however, a typical experimental situation where product concentrations remain vanishingly small. Therefore, product binding events...