

Stochastic thermodynamics of a chemical nanomachine: the channeling enzyme tryptophan synthase

Dimitri Loutchko,¹ Maximilian Eisbach,² and Alexander S. Mikhailov^{1,3}

¹*Department of Physical Chemistry, Fritz Haber Institute of the Max Planck Society, Faradayweg 4-6 14195 Berlin, Germany* ^{a)}

²*Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstraße 36, 10623 Berlin, Germany*

³*Department of Mathematical and Life Sciences, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526 Japan*

(Dated: 3 February 2017)

The enzyme tryptophan synthase is characterized by a complex pattern of allosteric interactions that regulate the catalytic activity of its two subunits and opening or closing of their ligand gates. As a single macromolecule, it implements 13 different reaction steps, with an intermediate product directly channeled from one subunit to another. Based on experimental data, a stochastic model for the operation of tryptophan synthase has been earlier constructed [J. Phys. Chem. B, 120, 2179 (2016)]. Here, this model is used to consider stochastic thermodynamics of such a chemical nanomachine. The Gibbs energy landscape of the internal molecular states is determined, the production of entropy and its flow within the enzyme are analyzed and the information exchange between the subunits resulting from allosteric cross-regulations and channeling is discussed.

Keywords: Nonequilibrium Phenomena, Single-Molecule Kinetics, Markov Networks, Entropy Production, Mutual Information

I. INTRODUCTION

Stochastic thermodynamics, using the concepts of random Markov processes, extends classical thermodynamics to small systems where fluctuations become strong¹⁻⁹. Because it is based on the Shannon entropy, there is an intimate connection between thermodynamics of small systems and information theory¹⁰⁻²⁴. A natural field for the application of stochastic thermodynamics is provided by the studies of molecular motors, protein machines and enzymes. When processes in individual biomolecules are considered, intrinsic fluctuations are essential and this theory has to be used²⁵⁻³³. Previously, stochastic thermodynamic descriptions have been employed for the rotary molecular motor F1-ATPase³⁴⁻³⁶ and for molecular walkers, such as myosin and kinesin³⁷⁻⁴⁰.

Here, stochastic thermodynamics of the channeling enzyme tryptophan synthase will be explored. This enzyme yields a paradigmatic example of a chemical nanomachine. Several reviews devoted to tryptophan synthase are available⁴¹⁻⁴⁴, with the authors characterizing it as an allosteric molecular factory⁴¹, a channeling nanomachine⁴² and even “a mine for enzymologists”⁴³. The enzyme consists of two subunits, each with its own active center, and, as a single molecule, catalyzes 13 different reaction steps.

A complex pattern of allosteric regulation is involved in its operation. The catalytic activity in a subunit is

enhanced or inhibited depending on the state of the other subunit. The gates controlling arrival and release of the ligands can become open or closed depending on the chemical states. Moreover, one intermediate product is directly channeled within the protein from one subunit to another, so that it is never released into the solution around it. The enzyme is important because it is used by all bacteria, yeast and plants to synthesize one of the essential amino acids, tryptophan.

A special feature of tryptophan synthase is that it has been extensively experimentally investigated⁴⁵⁻⁵¹. Its mechanism is completely known, all intermediate chemical states and reaction steps have been identified and rate constants for practically all transitions have been measured. The allosteric regulatory interactions have been determined and quantified. Moreover, protein conformations corresponding to different states have been determined by using X-ray diffraction methods.

Recently, a single-molecule stochastic description of tryptophan synthase has been constructed and the Markov network of this enzyme has been identified⁵². All its transition rate constants were deduced from the available experimental data and no arbitrary fitting parameters were employed. Using this model, stochastic single-molecule simulations have been performed. They have revealed the development of strong correlations between the subunits and synchronization of their internal states. Below, this stochastic description is used to explore thermodynamic aspects of the operation of this chemical nanomachine.

In the next section, the structure and the operation

^{a)}Electronic mail: loutchko@fhi-berlin.mpg.de

of tryptophan synthase are summarized. In Section III, its Markov transition network is introduced. This network is reversible and thus partially modified as compared to the previous publication⁵². We take into account the reverse binding of products, that can take place under physiological conditions for this enzyme, but has been typically negligible when kinetic experiments were performed. Moreover, additional calorimetric data is used to determine the rate constant for reverse channeling that has not been experimentally observed. The energetic aspects and the implications of the detailed balance are considered in Section IV. After that, in the following section, the analysis of entropy production and entropy flow within the enzyme in the nonequilibrium state corresponding to physiological conditions is performed. Section VI is focused on the information aspects of allosteric interactions between the two enzyme subunits and on the information effects of channeling events. The article ends with the conclusions and a discussion of obtained results.

The derivation of equations governing exchange of information between two subsystems in presence of transitions that affect simultaneously the states of both of them is given in the Appendix.

II. TRYPTOPHAN SYNTHASE AS A CHEMICAL NANOMACHINE

Tryptophan synthase is an enzyme that performs synthesis of the essential amino acid tryptophan. Higher organisms do not possess this enzyme, because they take up tryptophan with their food. However, all bacteria, plants and yeast are using it. The complete reaction scheme is shown in Fig. S1. There are two substrates, serine (ser) and indole glycerole phosphate (IGP), and they are converted into the products, tryptophan (trp) and glycerol-3-phosphate (G3P). The synthesis is complicated and consists of 13 elementary reaction steps.

It should be noted that one of the substrates, IGP, is scarce inside the cell and therefore the synthesis has to be performed in an economic way. Moreover, an intermediate product, indole, is hydrophobic and can easily escape through the cell membrane; therefore, it should not be released into the cytoplasm. To satisfy such limitations, Nature has found an elegant solution. The entire synthesis is performed inside a single biomolecule that has two subunits, each catalyzing a sequence of chemical reaction steps. The intermediate indole is never released into the solution, but channeled directly from one subunit to the other through a tunnel inside the enzyme.

The molecular structure of tryptophan synthase is shown in Fig. S2. The protein consists of α and β -subunits, both with their own catalytic sites. The

active centers of the subunits are connected by a 25 Å long intramolecular tunnel for transport of indole. Each catalytic site has its own gate controlling the release and uptake of substrates and products. The operation of tryptophan synthase involves a complex pattern of allosteric cross-regulation controlling the reactions and the configuration of the gates. It is known that the enzyme adopts two different conformational states: the catalytically inactive state with open gates (“open conformation”) and the state with enhanced catalytic activity and closed gates (“closed conformation”) depending on its chemical state^{51,53–55}. The switching between the two conformational states is mediated by the COMM subdomain of the β -subunit.

A simplified scheme of the catalytic cycle of tryptophan synthase with several omitted states is displayed in Fig. 1. Here, the α -subunit is shown in green and the β -subunit in blue. The chemical states have the same notations as in Fig. S1. The catalytic cycle begins with the enzyme in the state where both sites are empty and the gates are open. Then, the substrate IGP binds to the α -subunit and serine to the β -subunit, where it is quickly converted to the serine quinoline intermediate (Q_1). IGP activates the formation of the α -aminoacrylate (A-A) and the enzyme adopts the closed conformation, as schematically shown in Fig. 1c. In the state (IGP,A-A) where both gates are closed, A-A activates the cleavage of IGP to produce G3P and indole. Indole is then channeled to the β -site where it reacts with A-A to give the tryptophan quinoline intermediate (Q_3) that is converted to tryptophan (Aex₂ is the external aldimine of tryptophan in the β -subunit). In the state (G3P,Aex₂) the gates open and the products tryptophan and G3P are released. Thus the enzyme returns to the initial conformation (empty,empty) and is ready to start the next cycle.

The functioning of tryptophan synthase can be also illustrated in a different way (Fig. 2). Each subunit undergoes stochastic transitions that either represent internal chemical transformations or correspond to binding and release of substrates and products. Both subunits are coupled to chemical reservoirs where constant substrate and product concentrations are maintained. These reservoirs can also be considered as *chemostats*. Generally, there is a difference of chemical potentials between the substrate and the product chemostats. 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 \rightleftharpoons IGP and G3P \rightleftharpoons empty (magenta) in the α -site are blocked (i.e., the gate in the α -subunit is closed) in the states A-A, A-A + indole and Q_3 of the β -site. The transitions IGP \rightleftharpoons indole+G3P (light and dark blue) in the α -site are blocked in the states empty, Q_1 , Aex₂ of

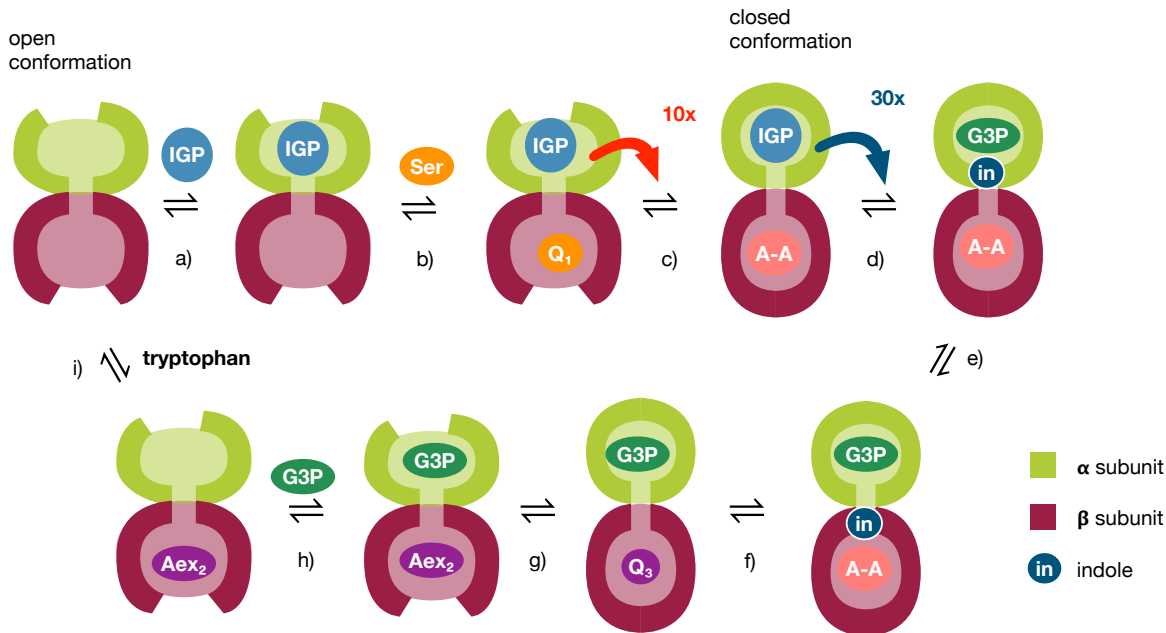


FIG. 1. Schematic operation of tryptophan synthase. Aex₂ is the state of the β -subunit with tryptophan present inside it.

TABLE I. Enumeration of chemical states of α - and β -subunits by variables a and b .

α -site	empty	IGP	indole+G3P	G3P		
Variable a	1	2	3	4		
β -site	empty	Q_1	A-A	indole+A-A	Q_3	Aex ₂
Variable b	1	2	3	4	5	6

the β -site. The rate of the transition IGP \rightarrow indole+G3P (light blue) in the α -site is enhanced by a factor of 27.7 in the state A-A of the β -site. The transitions $Q_1 \rightleftharpoons$ A-A and $Q_3 \rightleftharpoons$ Aex₂ (green) in the β -site are blocked in the state empty of the α -site. The transition $Q_1 \rightarrow$ A-A (light green) in the β -site is enhanced by a factor of 9.7 in the state IGP of the α -site. The changes indole+G3P \rightleftharpoons G3P and A-A \rightleftharpoons indole+A-A (red) corresponding to indole channeling from the α - to the β -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 α - and β -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 $a = 1, \dots, 4$ and

$b = 1, \dots, 6$ represents a possible combination of the individual states a and b of the α - and the β -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 box 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 publication⁵² 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 were not included into the scheme. Moreover, the reverse indole channeling transitions were not taken into account, because such reaction events have never been experimentally observed. Additionally, the reverse reaction in the beta-subunit was not included. Because of this, the previously constructed network has been partially irreversible.

In contrast to this, the network shown in Fig. 3 is fully reversible. The enzyme is now assumed to operate under typical physiological conditions where products are present in substantial concentrations and product binding events can therefore take place. Moreover,

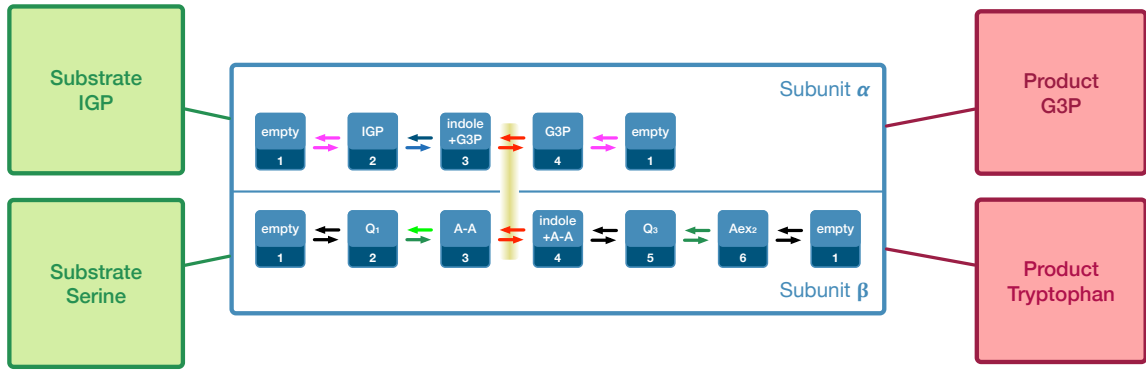


FIG. 2. Allosteric cross-regulation and channeling in tryptophan synthase. Magenta: transitions blocked in the states A-A, A-A + indole and Q_3 of the β -site. Green (light and dark): blocked in the state empty of the α -site. Light green: enhanced by a factor of 9.7 in the state IGP of the α -site. Blue (light and dark): blocked in the states empty, Q_1 , Aex_2 of the β -site. Light blue: enhanced by a factor of 27.7 in the state A-A of the β -site. Red: Channeling instantaneously changes the states of both sites. Numbers represent numerical notations of the respective chemical states of a subunit. Aex_2 is the state of subunit β with tryptophan inside it.

as shown below, we can use additional experimental data and thermodynamic consistency conditions to determine the missing rate of reverse indole channeling. Additionally, the rate for the reverse reaction $Q_3 \rightarrow \text{indole} + \text{A-A}$ in the β -subunit has been taken from⁵⁶. Numerical values of the rate constants of each transition are shown next to the respective arrows in Fig. 3. The rates of the binding of substrate and product molecules are proportional to substrate and product concentrations. The substrate and product concentrations under physiological conditions were taken from⁵⁷. The respective binding rate constants are given in Table II.

For the combined states (a, b) , time dependent probabilities $p(a, b; t)$ can be introduced. They satisfy the master equation

$$\frac{d}{dt}p(a, b; t) = \sum_{a'=1}^4 \sum_{b'=1}^6 [w_{a',a'}^{b,b'}p(a', b'; t) - w_{a',a}^{b',b}p(a, b; t)] \quad (1)$$

where $w_{a',a}^{b,b'}$ denotes the transition rate from a state (a', b') to the state (a, b) . The numerical values of the transition rates are shown in Fig. 3

As we have already noticed (see, e.g., Fig. 1), only the transition representing indole channeling involves simultaneous changes of the states of both α - and β -sites. All other transitions change the state of only one subunit although the rates of such transitions can be controlled by the state of the other subunit. Therefore, the Markov network of tryptophan synthase has a special structure. It is almost bipartite (see^{19–21}) and the transition matrix elements can be written as

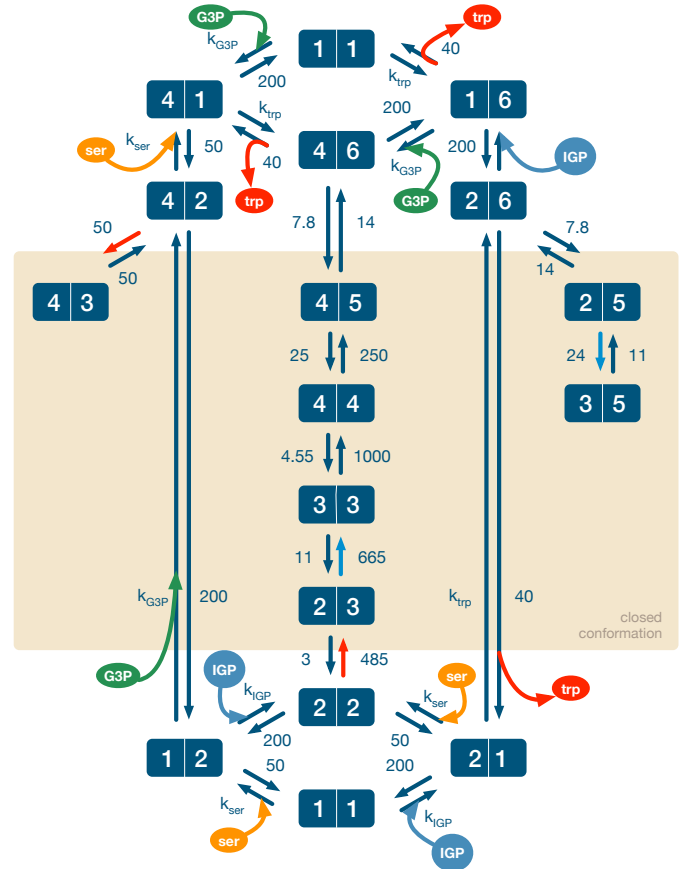


FIG. 3. The kinetic Markov network of tryptophan synthase with numerical values of all transition rates in units of s^{-1} .

TABLE II. Ligand binding rate constants k , ligand concentrations c and the respective transition rates $w = kc$ under physiological conditions.

Reaction	Binding rate constant k	Concentration c	Transition rate w	Reference
β -empty + Ser \rightarrow Q ₁	$7.5 \cdot 10^{-2} \mu\text{M}^{-1}\text{s}^{-1}$	$c(\text{Ser})=68 \mu\text{M}$	5.1 s^{-1}	45
β -empty + Trp \rightarrow Aex ₂	$0.15 \mu\text{M}^{-1}\text{s}^{-1}$	$c(\text{Trp})=12 \mu\text{M}$	1.8 s^{-1}	45
α -empty + IGP \rightarrow α -IGP	$10 \mu\text{M}^{-1}\text{s}^{-1}$	$c(\text{IGP})=3.5 \mu\text{M}$	35 s^{-1}	46,58
α -empty + G3P \rightarrow α -G3P	$0.2 \mu\text{M}^{-1}\text{s}^{-1}$	$c(\text{G3P})=49 \mu\text{M}$	9.8 s^{-1}	46

$$w_{a,a'}^{b,b'} = \begin{cases} w_a^{b,b'} & \text{if } a = a' \\ w_{a,a'}^b & \text{if } b = b' \\ w_{4,3}^{4,3} & \text{if } (a', b') = (3, 3) \text{ and } (a, b) = (4, 4) \\ w_{3,4}^{3,4} & \text{if } (a', b') = (4, 4) \text{ and } (a, b) = (3, 3) \\ 0 & \text{else.} \end{cases} \quad (2)$$

The indole channeling couples the two subunits and perturbs the complete bipartite structure of the Markov network.

It is convenient to define the probability fluxes

$$J_{a,a'}^{b,b'} = w_{a,a'}^{b,b'} p(a', b'; t) - w_{a',a}^{b',b} p(a, b; t) \quad (3)$$

and write the master equation as a continuity equation

$$\frac{d}{dt} p(a, b; t) = \sum_{a', b'} J_{a', b'}^{a, b}. \quad (4)$$

Taking into account the special form (2) of the transition matrix, the master equation can also be written as

$$\frac{d}{dt} p(a, b; t) = \sum_{b'} J_a^{b,b'} + \sum_{a'} J_{a,a'}^b + [\delta_{(a,b)}^{(3,3)} - \delta_{(a,b)}^{(4,4)}] J^{channel} \quad (5)$$

where $\delta_i^j = 1$, if $i = j$ and $\delta_i^j = 0$ otherwise. The fluxes corresponding to transitions inside the β -subunit are $J_a^{b,b'} = w_a^{b,b'} p(a, b'; t) - w_{a',a}^{b',b} p(a, b; t)$ and the fluxes $J_{a,a'}^b$ for the transitions within the α -subunit are defined similarly. The flux corresponding to channeling is $J^{channel} = w_{4,3}^{4,3} p(3, 3; t) - w_{3,4}^{3,4} p(4, 4; t)$.

IV. ENERGIES AND DETAILED BALANCE

Thermodynamics implies that all transitions between the states should satisfy the condition of detailed balance. This condition is that at thermal equilibrium the net probability flux between any two states is absent. For the considered network it implies that the ratio of the rates $w_{a',a}^{b',b}$ and $w_{a,a'}^{b,b'}$ for forward and backward transitions between any two states (a, b) and (a', b') satisfies the equation

$$\frac{w_{a,a'}^{b,b'}}{w_{a',a}^{b',b}} = \exp\left(\frac{G(a, b) - G(a', b')}{k_B T}\right) \quad (6)$$

where $G(a, b)$ and $G(a', b')$ are Gibbs energies of the respective states in the network at equilibrium, T is the temperature, and k_B is the Boltzmann constant.

For transitions between the states $(a, b) \rightarrow (a', b')$ that do not involve binding or release of ligands, the rates $w_{a,a'}^{b,b'}$ coincide with the respective rate constants $k_{a,a'}^{b,b'}$ and the Gibbs energies $G(a, b)$ are the internal Gibbs energies $g(a, b)$ of the molecular states. In this case, equation (6) takes the form

$$\frac{k_{a,a'}^{b,b'}}{k_{a',a}^{b',b}} = \exp\left(\frac{g(a, b) - g(a', b')}{k_B T}\right). \quad (7)$$

Note that, for macromolecules the Gibbs energies $g(a, b)$ of internal states are different from the internal energies $\epsilon(a, b)$ of such states because they additionally include entropic contributions and solvent effects.

The transitions that involve binding or release of a ligand should be treated separately. Suppose that a transition from (a, b) to (a', b') is accompanied by binding of a ligand and the ligand is released in the backward transition. Then the forward transition rate is proportional to the ligand concentration c , i.e. $w_{a',a}^{b',b} = k_{a',a}^{b',b} c$, whereas for the backward transition we have $w_{a,a'}^{b,b'} = k_{a,a'}^{b,b'}$. Moreover, the Gibbs energies in (6) include now contributions from ligand particles, i.e. $G(a, b) = g(a, b) + \mu$, where μ is the chemical potential of the ligand. For the considered weak solutions, we have $\mu = \mu_0 + k_B T \ln c$. Substitution of these expressions into equation (6) yields

$$\frac{k_{a,a'}^{b,b'}}{k_{a',a}^{b',b}} = \exp\left(\frac{g(a, b) + \mu_0 - g(a', b')}{k_B T}\right). \quad (8)$$

In this equation, the ligand can be either a substrate or a product if reverse binding of a product molecule takes place.

As shown by Schnakenberg¹, one can derive further identities by considering different pathways in a Markov network. Suppose that the chosen pathway represents a closed cycle Γ that involves only the internal states of the molecule without the events of ligand binding or release. Then, by using equation (7), one can show that the identity

$$\prod_{\Gamma} \frac{k_{a,a'}^{b,b'}}{k_{a',a}^{b',b}} = \exp \left(\sum_{\Gamma} \frac{g(a,b) - g(a',b')}{k_B T} \right) = 1 \quad (9)$$

holds, with the multiplication on the left side performed over all transitions that belong to the chosen cycle.

If the pathway Γ involves a conversion of substrate to a product or back, application of condition (8) leads to a modified identity. For tryptophan synthase, it has the form

$$\prod_{\Gamma} \frac{w_{a,a'}^{b,b'}}{w_{a',a}^{b',b}} = \exp \left(\frac{\mu(\text{trp}) + \mu(\text{G3P}) - \mu(\text{ser}) - \mu(\text{IGP})}{k_B T} \right) \quad (10)$$

if the pathway Γ leads from the bottom to the top empty states (1,1) in the Markov network in Fig. 3, i.e. if it corresponds to conversion of the two substrate molecules IGP and serine to the two product molecules G3P and tryptophan.

The detailed balance condition (6) and the Schnakenberg identities (9) and (10) can be used to check the thermodynamic consistency of a Markov network, to find missing rate constants of some transitions, and to determine Gibbs energies of different states. Particularly, in the Markov network of tryptophan synthase, there is a transition from the state (4,4) to (3,3) that corresponds to the channeling of indole from the β - to the α -site. This transition has never been observed experimentally and its rate constant could not be measured. This rate constant can however be determined, as explained below, by using the identity (10) and additional experimental data.

In each turnover cycle of tryptophan synthase two substrate molecules (IGP and serine) are converted into two product molecules (G3P and tryptophan) and some energy Δq is dissipated during this conversion process. In microcalorimetric measurements under isothermal conditions, it is possible to determine the heat Q released by the reaction in a given volume per unit time. This heat can be estimated as $Q = N\Delta q/\tau$ where τ is the mean turnover time and N is the number of enzyme molecules in the reaction volume. The last two quantities can be determined independently and thus the amount of heat Δq in one turnover cycle can be obtained. Such calorimetric measurements have been performed for tryptophan synthase, under standard concentration conditions $c_0(\text{IGP}) = c_0(\text{ser}) = c_0(\text{G3P}) = c_0(\text{trp}) = 1 \text{ M}$,

and they yield $\Delta q = 20.46 k_B T^{59}$.

On the other hand, the released heat Δq corresponds to the difference of chemical potentials of substrates and products

$$\Delta q = \mu(\text{IGP}) + \mu(\text{ser}) - \mu(\text{G3P}) - \mu(\text{trp}) \quad (11)$$

where, for example, $\mu(\text{IGP}) = \mu_0(\text{IGP}) + k_B T \ln c(\text{IGP})$. Thus, the difference of the chemical potentials $\mu_0(\text{IGP}) + \mu_0(\text{ser}) - \mu_0(\text{G3P}) - \mu_0(\text{trp})$ in tryptophan synthase under standard conditions is equal to $20.46 k_B T$.

By using the identity (10) and the known value of Δq for tryptophan synthase, reverse channeling transition rate can be determined as $k_{3,4}^{3,4} = 4.55 \text{ s}^{-1}$. This is indeed much smaller than the measured rate $k_{4,3}^{4,3} = 1000 \text{ s}^{-1}$ of the forward channeling transition. Therefore, the reverse channeling transitions should be very rare and this is why they have not been experimentally observed.

Furthermore, the detailed balance conditions (6) and (10) can be used to determine, by repeated application, Gibbs energies $G(a,b)$ with respect to the Gibbs energy of a certain reference state.

Our reference state corresponds to the free enzyme with two products (tryptophan and G3P) and its Gibbs energy is chosen as $G_{\text{final}} = 0$. In the initial state, the enzyme is free, there are two additional substrate molecules (serine and IGP) and the two product molecules (tryptophan and G3P) are missing. The Gibbs energy of the initial state is therefore $G_{\text{initial}} = \mu(\text{IGP}) + \mu(\text{ser}) - \mu(\text{G3P}) - \mu(\text{trp})$. It should be noted that it depends on the involved ligand concentrations c because $\mu = \mu_0 + k_B T \ln c$. It coincides with the amount of heat Δq released in one turnover cycle. The value above given $\Delta q = 20.46 k_B T$ corresponds to the standard conditions $c_0(\text{IGP}) = c_0(\text{ser}) = c_0(\text{G3P}) = c_0(\text{trp}) = 1 \text{ M}$. Recalculating this under the physiological concentrations (Table II), we obtain $G_{\text{initial}} = \Delta q = 19.56 k_B T$.

There are also several states where one of the subunits is empty and the other subunit has a ligand bound to it. For example, the state (IGP, empty) has IGP bound to the α -subunit and no ligand in the β -subunit. The Gibbs energy of this state is $G(\text{IGP,empty}) = g(\text{IGP,empty}) - g_0 + \mu(\text{ser}) - \mu(\text{G3P}) - \mu(\text{trp})$. It includes both the difference of the chemical potentials, depending on the concentrations, and the internal Gibbs energies $g(\text{IGP,empty})$ and $g_0 = g(\text{empty,empty})$ of the state (IGP, empty) and the free state of the enzyme.

Finally, there are states where both subunits are occupied. For example, for the state (IGP, Q_1), we have $G(\text{IGP}, Q_1) = g(\text{IGP}, Q_1) - g_0 - \mu(\text{G3P}) - \mu(\text{trp})$. For

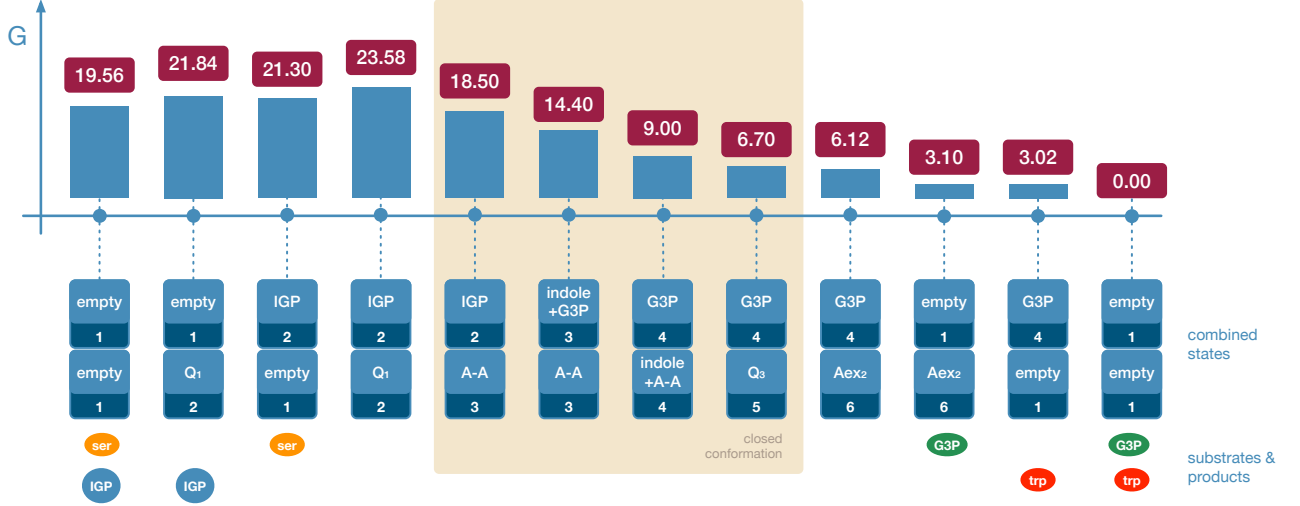


FIG. 4. The Gibbs energy landscape along the main pathway of tryptophan synthase under physiological ligand concentrations. The Gibbs energies are given in units of $k_B T$. Physiological concentration values are chosen. In the states within the beige box, the molecular gates are closed and the enzyme is disconnected from the chemostats. Tryptophan is present inside the β -subunit in the state Aex₂.

the state (IGP, A-A), we have $G(\text{IGP}, \text{A-A}) = g(\text{IGP}, \text{A-A}) - g_0 - \mu(\text{G3P}) - \mu(\text{trp})$. Note that the difference $G(\text{IGP}, \text{Q}_1) - G(\text{IGP}, \text{A-A}) = g(\text{IGP}, \text{Q}_1) - g(\text{IGP}, \text{A-A})$ is determined only by the internal Gibbs energies of the states and is independent of ligand concentrations. This difference gives the amount of heat dissipated in the respective transition.

Figure 4 shows the Gibbs energy landscape of tryptophan synthase along its main pathway. After the binding of substrates requiring activation energies of $1.74 k_B T$ for IGP binding and $2.28 k_B T$ for serine binding, all transitions towards product formation are exergonic. The four catalytically important transitions $(\text{IGP}, \text{Q}_1) \leftrightarrow (\text{IGP}, \text{A-A}) \leftrightarrow (\text{indole} + \text{G3P}, \text{A-A}) \leftrightarrow (\text{G3P}, \text{indole} + \text{A-A}) \leftrightarrow (\text{G3P}, \text{Q}_3)$ in the closed conformation of the enzyme are highly exergonic and accompanied by heat release in the range between 5.40 and $2.30 k_B T$. The release of the products G3P and tryptophan is accompanied by the heat release of 3.10 and $3.02 k_B T$, respectively.

V. ENTROPY PRODUCTION AND FLOW

The Shannon entropy at time t is given by

$$S(t) = - \sum_{a,b} p(a, b; t) \ln p(a, b; t) \quad (12)$$

where $p(a, b; t)$ is the probability to find the enzyme in the state (a, b) at time t . Its rate of change is

$$d_t S = \frac{1}{2} \sum_{a,a',b,b'} J_{a,a'}^{b,b'} \ln \frac{p(a', b'; t)}{p(a, b; t)} \quad (13)$$

where $J_{a,a'}^{b,b'}$ is the probability flux for the transition from the state (a', b') to (a, b) defined in equation (3). This rate of change can be decomposed as

$$d_t S = \sigma - h \quad (14)$$

into the difference of the entropy production σ inside the enzyme and of the net flow h of entropy *from* the enzyme, i.e. of the rate of entropy export by it.

According to equation (13), the rate of change of the total entropy of the enzyme can be also written as a sum of the rates of entropy change $i_{a,a'}^{b,b'}$ in each individual transition, i.e.

$$d_t S = \frac{1}{2} \sum_{a,a',b,b'} s_{a,a'}^{b,b'}, \text{ with } s_{a,a'}^{b,b'} = J_{a,a'}^{b,b'} \ln \frac{p(a', b')}{p(a, b)}. \quad (15)$$

The same holds for the total entropy production σ and the rate of entropy export h . Thus the quantities $\sigma_{a,a'}^{b,b'}$ and $h_{a,a'}^{b,b'}$ corresponding to individual transitions can be introduced,

$$h = \frac{1}{2} \sum_{a,a',b,b'} h_{a,a'}^{b,b'}; \quad \sigma = \frac{1}{2} \sum_{a,a',b,b'} \sigma_{a,a'}^{b,b'} \quad (16)$$

where (cf.7)

$$h_{a,a'}^{b,b'} = J_{a,a'}^{b,b'} \ln \frac{w_{a,a'}^{b,b'}}{w_{a',a}^{b',b}}, \quad (17)$$

$$\sigma_{a,a'}^{b,b'} = J_{a,a'}^{b,b'} \ln \frac{w_{a,a'}^{b,b'} p(a',b')}{w_{a',a}^{b',b} p(a,b)}. \quad (18)$$

Note that $\sigma_{a,a'}^{b,b'} = \sigma_{a',a}^{b',b}$, $h_{a,a'}^{b,b'} = h_{a',a}^{b',b}$ and $s_{a,a'}^{b,b'} = s_{a',a}^{b',b}$ and therefore these quantities do not depend on the choice of the transition direction, i.e. they are the same both for $(a,b) \rightarrow (a',b')$ and $(a',b') \rightarrow (a,b)$. The entropy production $\sigma_{a,a'}^{b,b'}$ characterizes dissipation or irreversibility whereas the entropy flow $h_{a,a'}^{b,b'}$ represents the average heat exchanged per unit time with the environment in a transition⁷. We have $s_{a,a'}^{b,b'} = \sigma_{a,a'}^{b,b'} - h_{a,a'}^{b,b'}$.

In the state of thermal equilibrium, all fluxes $J_{a,a'}^{b,b'}$ vanish and therefore according to equations (17) and (18) there are no transitions where entropy is produced or exported. Under physiological conditions, however, the enzyme tryptophan synthase operates far from thermal equilibrium, with the difference of Gibbs energies of $19.56 k_B T$ for one cycle. Thus, its operation is characterized by nonequilibrium steady-state. In the respective nonequilibrium steady-state with the stationary probability distribution $\bar{p}(a,b)$, the fluxes $\bar{J}_{a,a'}^{b,b'}$ do not vanish and therefore the transitions are accompanied by entropy production and entropy export. Because the entropy S is conserved in this state, $d_t S = \sigma - h = 0$. Hence the total entropy production σ is counterbalanced by the entropy export h . Note that, although $d_t S = 0$, the rates of entropy change $s_{a,a'}^{b,b'}$ for individual transitions are not zero even in the nonequilibrium steady-state.

The stationary probability distribution $\bar{p}(a,b)$ can be found by solving the master equation (1) in the nonequilibrium steady-state. Numerical values of the probabilities $\bar{p}(a,b)$ corresponding to all possible states are given Table S1. Then, by using equation (3), the fluxes $\bar{J}_{a,a'}^{b,b'}$ can be determined. According to equations (15), (17) and (18), the values of $\sigma_{a,a'}^{b,b'}$, $h_{a,a'}^{b,b'}$ and $s_{a,a'}^{b,b'}$ can be calculated afterwards.

The results are displayed in Figs. 5 and 6. Here, we show the same network as in Fig. 3, but, for simplicity, retain only numerical notations of the states. Only the links between the states are shown because the transition directions are not important (see the comment after equation (18)) For each link, the value of the quantities $\sigma_{a,a'}^{b,b'}$ or $h_{a,a'}^{b,b'}$ is indicated. Additionally, color coding is used.

Here and below, all numerical values for entropy and information are given in units of bits. We have 1 bit =

$\ln 2 = 0.693$, because natural logarithms are used in our definition of the Shannon entropy.

The rates of entropy or information change are given in bits per seconds. Alternatively, they can also be expressed by the respective amounts per a catalytic cycle. Note that the substrate conversion rate of the enzyme is equal to the probability flux $J^{channel}$ because each productive cycle includes this transition. The mean catalytic cycle time is the inverse of the substrate conversion rate. Under physiological concentrations, we find that the mean cycle time is 0.75 s. Tryptophan synthase is a slow molecular machine.

Figure 5 shows numerical values of entropy production for all individual transitions within the enzyme. The entropy is mostly produced along the main catalytic pathway. The highest entropy production (10.22 bit s⁻¹) is found for the allosterically activated transition $Q_1 \leftrightarrow A-A$ in the β -site. In contrast to this, all transitions involving futile states (side branches of the network) have values of entropy production below 0.01 bit s⁻¹ per second. Ligand binding and release is characterized by entropy production below 1.78 bit s⁻¹ per second.

The values for entropy export are given in Fig. 6. The entropy export takes is maximal (between 7.77 bit s⁻¹ and 10.22 bit s⁻¹) for the transitions $(IGP,Q_1) \leftrightarrow (IGP,A-A) \leftrightarrow (\text{indole}+G3P,A-A) \leftrightarrow (G3P,\text{indole}+A-A)$ where most of the heat exchange with the environment takes place. All other transitions have absolute values smaller than 4.61 bit s⁻¹. Note that transition $(G3P,Q_3) \leftrightarrow (G3P,Aex_2)$ has a small entropy export, but a high entropy production.

Because the rate of entropy change in a transition is given by the difference of entropy production and export, this rate can be found by subtracting the respective values in Figs. 5 and 6. Thus the transition $(G3P,Q_3) \leftrightarrow (G3P,Aex_2)$ in the main catalytic pathway has the largest rate of entropy increase $s_{4,4}^{6,5} = 4.82$ bit s⁻¹. In contrast to this, channeling and the subsequent transition $(G3P,\text{indole}+A-A) \leftrightarrow (G3P,Q_3)$ are accompanied by the net export of entropy at the rates $s_{3,3}^{4,4} = -4.53$ bit s⁻¹ and $s_{4,4}^{5,4} = -3.55$ bit s⁻¹.

Using the computed rates of entropy production and export for individual transitions, total amounts for the whole enzyme per a turnover cycle can be obtained. We find that, within a single catalytic cycle of tryptophan synthase, 27.79 bits of entropy are produced. The same amount of entropy is on the average exported by the enzyme per one cycle.

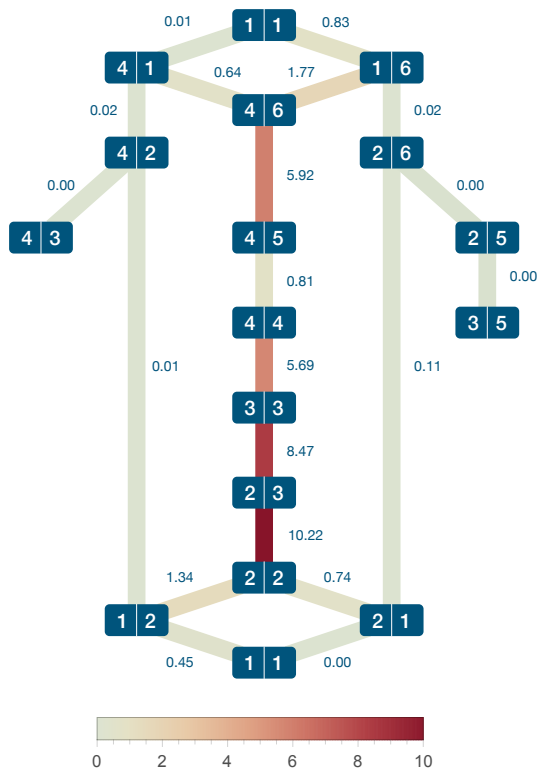


FIG. 5. Entropy production in different transitions in the nonequilibrium steady-state. The values of entropy production are given in units of bit s^{-1} next to the links between the states. Additionally, color coding of the links according to the corresponding entropy production is used.

VI. INFORMATION EXCHANGE BETWEEN THE SUBUNITS

There is a complex pattern of allosteric interactions between the two subunits of tryptophan synthase. Additionally, one transition that corresponds to indole channeling and affects simultaneously both subunits takes place. The allosteric cross-regulations and channeling lead to the development of correlations between the internal states of the subunits. Previously, the presence of correlations has been demonstrated by computing the Pearson correlation coefficients for all possible pairs of states⁵². In this section, the concept of mutual information will be employed to further quantify the effects of allosteric cross-regulation and channeling. Our method is based on the theory of information exchange between the subsystems^{19–21} which had, however, to be extended to the situation where transitions involving simultaneously both subunits can also occur. The formulation of the extended theory is provided in the Appendix.

The correlation $i(a, b)$ between the states a and b of the two subunits is defined as

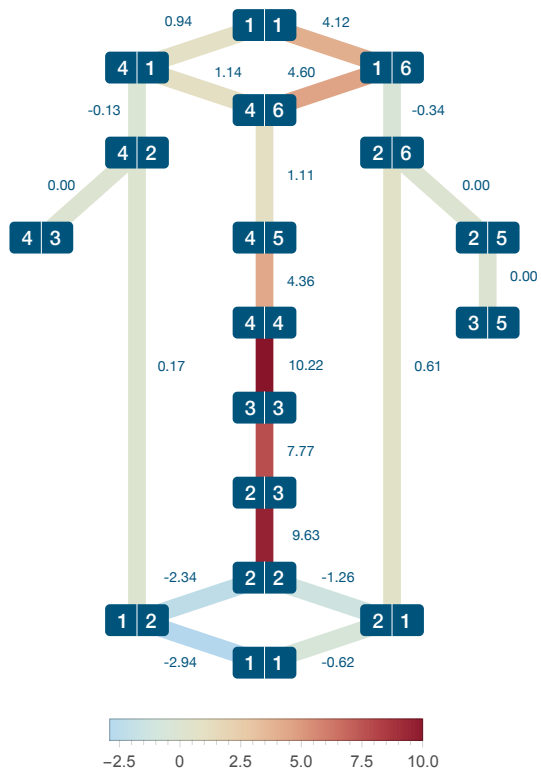


FIG. 6. Rates of entropy export in individual transitions in tryptophan synthase. The same notations as in Fig. 5.

$$i(a, b) = \ln \frac{p(a, b)}{p_\alpha(a)p_\beta(b)} \quad (19)$$

where $p_\alpha(a) = \sum_{b=1}^6 p(a, b)$ and $p_\beta(b) = \sum_{a=1}^4 p(a, b)$ are the probability distributions for the states of α - and β -subunits. This property quantifies correlations between the states a of the α -subunit and b of the β -subunit, it vanishes if these states are statistically independent, i.e. if $p(a, b) = p_\alpha(a)p_\beta(b)$. If it is negative, anti-correlations between the states are present.

The values $i(a, b)$ under physiological conditions are shown in Fig. 7 for all states (a, b) . We find high correlations (2.39 and 2.20 bits) between the states (G3P, indole+A-A) and (G3P, Q₃) after indole channeling and after the indole reaction at the β -site in the main pathway. This agrees with the previous analysis using the Pearson correlation coefficients⁵². As a result of channeling, both subunits simultaneously arrive at the state (G3P, indole+A-A) and high positive correlations are characteristic for it. On the other hand, we find anticorrelation (-1.04 bits) in the state (IGP, A-A) before channeling. This is an effect of allosteric interactions: when the β -subunit is in the state A-A, the cleavage of IGP into G3P and indole is blocked when the β -subunit is in the state Q₁, but it is possible in the state A-A.

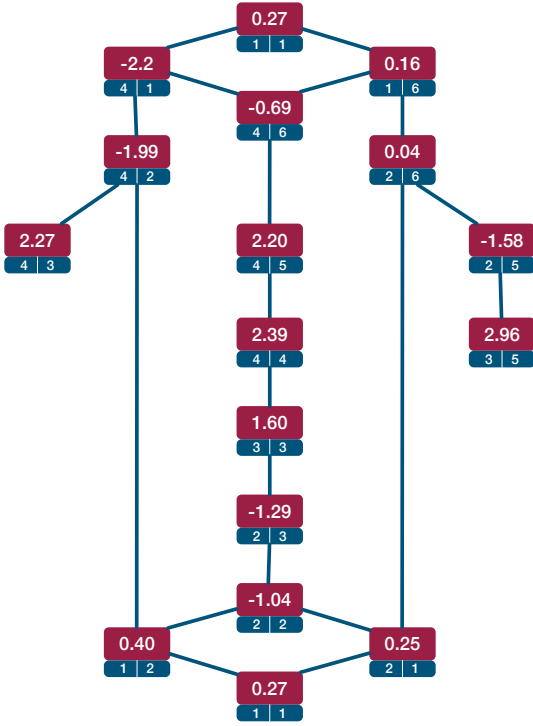


FIG. 7. Correlations $i(a, b)$ in units of bits for different states a and b .

The statistical average of $i(a, b)$ over all pair states (a, b) yields the mutual information I of the whole system

$$I = \sum_{a=1}^4 \sum_{b=1}^6 p(a, b) \ln \frac{p(a, b)}{p_\alpha(a)p_\beta(b)} = \sum_{a=1}^4 \sum_{b=1}^6 p(a, b) i(a, b). \quad (20)$$

This property is positive and it characterizes the strength of statistical correlations between the α - and β -subunits. For tryptophan synthase under physiological conditions we have $I = 0.49$ bit.

As shown in the Appendix, the rate of mutual information change for the entire system is

$$d_t I = \frac{1}{2} \sum'_{a, a'=1}^4 f_{a, a'}^\alpha + \frac{1}{2} \sum'_{b, b'=1}^6 f_{b, b'}^\beta + f^{channel}. \quad (21)$$

Here, the sums exclude the forward and backward channeling transitions and we have

$$f_{a, a'}^\alpha = \sum_b J_{a, a'}^b [i(a, b) - i(a', b)], \quad (22)$$

$$f_{b, b'}^\beta = \sum_a J_a^{b, b'} [i(a, b) - i(a, b')], \quad (23)$$

$$f^{channel} = J^{channel} [i(4, 4) - i(3, 3)]. \quad (24)$$

Note that in a steady state $d_t I = 0$ and therefore the terms (22) - (24) satisfy one additional constraint. Moreover, the terms $f_{a, a'}^\alpha$ and $f_{b, b'}^\beta$ do not depend on the choice of a direction for the transitions between a and a' or between b and b' . The quantity $f_{a, a'}^\alpha$ gives the contribution by the transition between the states a and a' in the α -subunit to the rate of change of the total mutual information of the system; this contribution is averaged over all possible regulatory states of the subunit β . A similar interpretation holds for the quantity $f_{b, b'}^\beta$.

By solving the master equation under physiological concentrations, we obtain the steady state probabilities $\bar{p}(\alpha, \beta)$. Substituting them into equation (3) and into the equations (22) - (24) yields the values for $f_{a, a'}^\alpha$, $f_{b, b'}^\beta$ and $f^{channel}$.

Figure 8 shows how the generation (or loss) of mutual information is distributed over the network. Mutual information is generated in three transitions in the α -subunit. Its highest generation rate is 3.79 bit s^{-1} in the transition (IGP \leftrightarrow indole+G3P) preceding channeling. The channeling transition itself generates mutual information at a smaller rate (1.04 bit s^{-1}). All transitions in the β -subunit are accompanied by mutual information loss with the highest rate (-3.79 bit s^{-1}) achieved in the transition immediately after channeling ($Q_3 \leftrightarrow Aex_2$).

Furthermore, information interactions between entire subunits can also be discussed. To do this, we write the rate of change of mutual information in the form

$$d_t I = F^\alpha + F^\beta + F^{channel} \quad (25)$$

where

$$F^\alpha = \frac{1}{2} \sum_{a, a', b} f_{a, a', b}^\alpha + f_{channel}^\alpha, \quad (26)$$

$$F^\beta = \frac{1}{2} \sum_{a, b, b'} f_{a, b, b'}^\beta + f_{channel}^\beta. \quad (27)$$

Here, we have divided the rate of generation of mutual information in the channel $f^{channel}$, given by equation (24) into three parts, i.e. $f^{channel} = F^{channel} + f_{channel}^\alpha + f_{channel}^\beta$, where

$$F^{channel} = J^{channel} \ln \frac{p(3, 3)}{p(4, 4)}, \quad (28)$$

$$f_{channel}^\alpha = J^{channel} \ln \frac{p_\alpha(3)p(4, 4)}{p_\alpha(4)p(3, 3)}, \quad (29)$$

$$f_{channel}^\beta = J^{channel} \ln \frac{p_\beta(3)p(4, 4)}{p_\beta(4)p(3, 3)}. \quad (30)$$

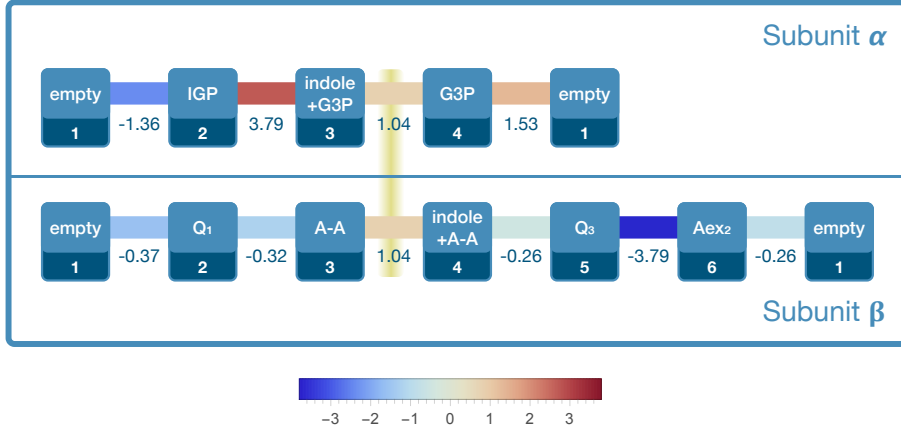


FIG. 8. Rates of change of mutual information in units of bits per second for the transitions within α - and β -subunits and for the channeling transition.

Thus the rates of mutual information change in α - and β -subunits include now contributions $f_{channel}^{\alpha}$ and $f_{channel}^{\beta}$ from the channeling transition. The advantage of this definition is that, as shown in the Appendix, important thermodynamic inequalities for entropy production in both subunits become then satisfied.

In a steady state, $d_t I$ vanishes and we have $F^{\alpha} + F^{\beta} + F^{channel} = 0$. If the channeling were absent, we would have had $F^{\alpha} = -F^{\beta}$. In this case, the mutual information generated in one subunit would have been completely consumed in the other subunit, cf.^{19–21}. Because $F^{channel} \neq 0$, this is, however, no longer valid. Some mutual information for the entire enzyme is additionally generated in the channeling transition involving simultaneously both subunits.

We have computed the values for F^{α} , F^{β} and $F^{channel}$ under physiological concentrations. We find that they all have the same order of magnitude. The mutual information $F^{channel} = -4.53 \text{ bit s}^{-1}$ generated per unit time by the transition corresponding to indole channeling flows to both subunits where is consumed at the rates of $F^{\alpha} = 3.09 \text{ bit s}^{-1}$ and $F^{\beta} = 1.42 \text{ bit s}^{-1}$. Note that F^{β} is positive whereas all contributions $f_{b,b'}^{\beta}$ from individual transitions in the β -subunit are negative. This is an effect of the large contributions from the cross term $f_{channel}^{\beta} = 6.43 \text{ bit s}^{-1}$ (whereas $f_{channel}^{\alpha} = -0.86 \text{ bit s}^{-1}$).

VII. DISCUSSION

In this study, methods of stochastic thermodynamics have been applied to characterize the operation of the channeling enzyme tryptophan synthase.

Using thermodynamic identities related to the detailed balance, the Gibbs energy landscape of this enzyme along its main catalytic pathway could be reconstructed from the experimental data. We found that, under *in vivo* conditions, the cycle of this enzyme is driven by the Gibbs energy gradient of approximately $19.56 k_B T$ between its substrates and products. Thus, under physiological substrate and product concentrations, the enzyme operation is far from thermal equilibrium.

Inside the cycle of tryptophan synthase, only the first substrate binding transitions are thermally activated, with activation energies about $1 k_B T$. All other transitions, including the events of product release, correspond to a decrease in the Gibbs energy. In particular, channeling is driven by the energy difference of $5.4 k_B T$ and does therefore not represent a diffusion process.

Because the enzyme operates far from equilibrium, entropy is persistently produced. We found that 27.79 bits of entropy are produced and the same amount of entropy is exported, on the average, to the environment within one catalytic cycle. The distribution of entropy production over the Markov network is largely nonuniform.

Information interactions between the two catalytic subunits of the enzyme have been analyzed. Both the allosteric interactions between the subunits and the channelling of an intermediate product from one of them to another contribute to the change of mutual information. Thus, the previously existing theory^{19–21} had to be generalized to the situations where, in addition to regulatory interactions between the subsystems, the transitions simultaneously changing the states of both of them can also take place. We found that mutual information is generated both in α - and β -subunits at the rates 3.09 and 1.49 bits per second. This mutual information is consumed in the channeling transition so

that the balance is maintained. Moreover, contributions from individual allosterically regulated transitions in each of the subunits to the total mutual information change were determined too.

Thus, we have demonstrated that, through the use of stochastic thermodynamics, a rich quantitative characterization of the nonequilibrium operation of an enzyme can be produced. It would be interesting to perform analogous investigations for other enzymes with several catalytic subunits. Such further investigations can clarify the connections between various thermodynamic properties of such nanomachines and the aspects of the chemical function of the enzymes.

SUPPLEMENTARY MATERIAL

See supplementary material for the full reaction scheme of tryptophan synthase, a figure of its molecular structure and numerical data for the stationary probability distribution on the Markov network under physiological conditions.

ACKNOWLEDGMENT

The authors are grateful to G. Ertl for stimulating discussions and persistent support. We thank T. Brandes, J Horowitz and U. Seifert for valuable discussions.

Appendix

In this Appendix, the previous analysis of information exchange and entropy production for systems with a bipartite structure¹⁹⁻²¹ is extended to systems that also have cross-transitions between the two subsystems. For tryptophan synthase, there is one such transition and it corresponds to indole channeling. Our results are, however, general and hold for systems with any number of cross-transitions.

We consider a system (A, B) composed of two subsystems A and B . The states of the system are labeled as (a, b) . We assume that there are regulatory transitions that occur within one subsystem with their rates, however, affected by the state of the other subsystem, i.e. the transitions of the form $(a, b) \leftrightarrow (a, b')$ and $(a, b) \leftrightarrow (a', b)$. Moreover, there are also cross-transitions where the states of both subsystems become simultaneously changed, i.e. $(a, b) \leftrightarrow (a', b')$ with $a \neq a'$ and $b \neq b'$. For tryptophan synthase, there is one such transition and it corresponds to indole channeling.

The evolution of the joint probability distribution $p(a, b; t)$ obeys the master equation

$$\frac{d}{dt}p(a, b; t) = \sum_{a', b'} [w_{a, a'}^{b, b'} p(a', b'; t) - w_{a', a}^{b', b} p(a, b; t)] \quad (\text{A.1})$$

where $w_{a, a'}^{b, b'}$ denotes the transition rate from a state (a', b') to the state (a, b) . Distinguishing between the regulatory and cross-transitions, we can write

$$w_{a, a'}^{b, b'} = \begin{cases} w_a^{b, b'} & \text{if } a = a' \\ w_{a, a'}^b & \text{if } b = b' \\ w_{a, a'}^{b, b'} & \text{if } a \neq a' \text{ and } b \neq b' \end{cases} \quad (\text{A.2})$$

Moreover, probability fluxes can be introduced as

$$J_{a, a'}^b = w_{a, a'}^{b, b'} p(a', b'; t) - w_{a', a}^{b', b} p(a, b; t) \text{ if } b = b', \quad (\text{A.3})$$

$$J_a^{b, b'} = w_{a, a'}^{b, b'} p(a', b'; t) - w_{a', a}^{b', b} p(a, b; t) \text{ if } a = a', \quad (\text{A.4})$$

$$J_{a, a'}^{b, b'} = w_{a, a'}^{b, b'} p(a', b'; t) - w_{a', a}^{b', b} p(a, b; t) \text{ if } a \neq a' \text{ and } b \neq b'. \quad (\text{A.5})$$

The mutual information $i(a, b)$ for a pair of states (a, b) is defined as

$$i(a, b) = \ln \frac{p(a, b)}{p_A(a)p_B(b)} \quad (\text{A.6})$$

where $p_A(a) = \sum_b p(a, b)$ is the probability to find the subsystem A in the state a and $p_B(b) = \sum_a p(a, b)$. The average of $i(a, b)$ over all states (a, b) yields the mutual information I of the entire system

$$I = \sum_{a, b} p(a, b) \ln \frac{p(a, b)}{p_A(a)p_B(b)} = \sum_{a, b} p(a, b) i(a, b). \quad (\text{A.7})$$

Its time derivative $d_t I$ can be written in the form

$$d_t I = \frac{1}{2} \sum_{a, a'} f_{a, a'}^A + \frac{1}{2} \sum_{b, b'} f_{b, b'}^B + \frac{1}{2} \sum_{a, a', b, b'} f_{a, a'}^{b, b'} \quad (\text{A.8})$$

where the first two sums are taken over regulatory transitions in subsystems A or B and the last sum includes all cross-transitions in the considered system. We have

$$\begin{aligned} f_{a, a'}^A &= \sum_b J_{a, a'}^b [i(a, b) - i(a', b)] \\ &= \sum_b J_{a, a'}^b \ln \frac{p_B(b|a)}{p_B(b|a')}, \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} f_{b, b'}^B &= \sum_a J_a^{b, b'} [i(a, b) - i(a, b')] \\ &= \sum_a J_a^{b, b'} \ln \frac{p_A(a|b)}{p_A(a|b')}, \end{aligned} \quad (\text{A.10})$$

$$f_{a, a'}^{b, b'} = J_{a, a'}^{b, b'} [i(a, b) - i(a', b')]. \quad (\text{A.11})$$

Here, $p_A(a|b) = p(a,b)/p_B(b)$ is the conditional probability to find the A -system in state a if the B -system is in the state b and $p_B(b|a)$ is defined similarly.

Thus $f_{a,a'}^A$ yields the contribution to the total rate of change of mutual information due to the regulatory transition between a and a' that takes place in the subsystem A and is regulated by the subsystem B . A similar interpretation holds for $f_{b,b'}^B$. The term $f_{a,a'}^{b,b'}$ represents the contribution to the total rate of change of mutual information due to the cross-transition between (a,b) and (a',b') , with $a \neq a'$ and $b \neq b'$, that directly connects the two subsystems A and B .

Now we derive the influence of the coupling through regulatory and cross-transitions on each of the entire subsystems A and B . Therefore, we consider the amount of entropy Σ^A produced per unit time in the transitions that change the state of the A subsystem. It is given by equation (18),

$$\begin{aligned} \Sigma^A &= \frac{1}{2} \sum_{a,a',b} J_{a,a'}^b \ln \frac{w_{a,a'}^b p(a',b)}{w_{a',a}^b p(a,b)} + \\ &+ \frac{1}{2} \sum_{a \neq a', b \neq b'} J_{a,a'}^{b,b'} \ln \frac{w_{a,a'}^{b,b'} p(a',b')}{w_{a',a}^{b,b'} p(a,b)}. \end{aligned} \quad (\text{A.12})$$

In a similar way, the amount of entropy Σ^B produced in the B subsystem can be found

$$\begin{aligned} \Sigma^B &= \frac{1}{2} \sum_{a,b,b'} J_{a,b'}^{b,b'} \ln \frac{w_{a,b'}^{b,b'} p(a,b')}{w_{a',b}^{b,b'} p(a,b)} + \\ &+ \frac{1}{2} \sum_{a \neq a', b \neq b'} J_{a,a'}^{b,b'} \ln \frac{w_{a,a'}^{b,b'} p(a',b')}{w_{a',a}^{b,b'} p(a,b)}. \end{aligned} \quad (\text{A.13})$$

Suppose now that we observe the subsystem A without the knowledge of the states of the subsystem B , i.e. we have no access to the joint probability distribution $p(a,b)$ and use instead of it the probability distribution $p_A(a)$ in equation (A.12). Proceeding in this way, the *apparent* entropy production σ^A assigned to the subsystem A is obtained

$$\begin{aligned} \sigma^A &= \frac{1}{2} \sum_{a,a',b} J_{a,a'}^b \ln \frac{w_{a,a'}^b p_A(a')}{w_{a',a}^b p_A(a)} + \\ &+ \frac{1}{2} \sum_{a \neq a', b \neq b'} J_{a,a'}^{b,b'} \ln \frac{w_{a,a'}^{b,b'} p_A(a')}{w_{a',a}^{b,b'} p_A(a)}. \end{aligned} \quad (\text{A.14})$$

Similarly, we have

$$\begin{aligned} \sigma^B &= \frac{1}{2} \sum_{a,b,b'} J_{a,b'}^{b,b'} \ln \frac{w_{a,b'}^{b,b'} p_B(b')}{w_{a',b}^{b,b'} p_B(b)} + \\ &+ \frac{1}{2} \sum_{a \neq a', b \neq b'} J_{a,a'}^{b,b'} \ln \frac{w_{a,a'}^{b,b'} p_B(b')}{w_{a',a}^{b,b'} p_B(b)}. \end{aligned} \quad (\text{A.15})$$

The real entropy production rates Σ^A and Σ^B are always non-negative, whereas the apparent entropy production rates σ^A and σ^B can also be negative^{19–21}. The influence on the entropy production of system A (respectively, B) through coupling to the whole system is then given by the difference between the apparent and total entropy production. Thus we define

$$F^A = \sigma^A - \Sigma^A \quad (\text{A.16})$$

$$F^B = \sigma^B - \Sigma^B. \quad (\text{A.17})$$

Substituting equations (A.12) to (A.24), we obtain

$$\begin{aligned} F^A &= \frac{1}{2} \sum_{a,a',b} J_{a,a'}^b \ln \frac{p_B(b|a)}{p_B(b|a')} + \\ &\frac{1}{2} \sum_{a \neq a', b \neq b'} J_{a,a'}^{b,b'} \ln \frac{p_B(b|a)}{p_B(b'|a')}, \end{aligned} \quad (\text{A.18})$$

$$\begin{aligned} F^B &= \frac{1}{2} \sum_{b,b',a} J_{a,b'}^{b,b'} \ln \frac{p_A(a|b)}{p_A(a|b')} + \\ &+ \frac{1}{2} \sum_{a \neq a', b \neq b'} J_{a,a'}^{b,b'} \ln \frac{p_A(a|b)}{p_A(a'|b')}. \end{aligned} \quad (\text{A.19})$$

Note that F^A and F^B have contributions from terms $f_{a,a'}^A$ and $f_{b,b'}^B$, defined in equations (A.9) and (A.10) and used in the splitting of $d_t I$ in equation (A.8). In addition, they also include cross-terms originating from non-bipartite transitions.

Using F^A and F^B , equation (A.8) for the rate of change of mutual information can be written as

$$d_t I = F^A + F^B + F^{\text{cross}} \quad (\text{A.20})$$

where the quantity

$$F^{\text{cross}} = \frac{1}{2} \sum_{a \neq a', b \neq b'} J_{a,a'}^{b,b'} \ln \frac{p(a',b')}{p(a,b)}. \quad (\text{A.21})$$

is introduced. Note that the expression (A.21) for F^{cross} can be also formulated as

$$F^{\text{cross}} = \sum_{a \neq a', b \neq b'} s_{a,a'}^{b,b'} \quad (\text{A.22})$$

where $s_{a,a'}^{b,b'}$ is the Shannon entropy produced in the cross-transition from (a, b) to (a', b') . Using the non-negativity of Σ^A and Σ^B , we arrive at the second law-like inequalities

$$\Sigma^A = \sigma^A - F^A \geq 0, \quad (\text{A.23})$$

$$\Sigma^B = \sigma^B - F^B \geq 0. \quad (\text{A.24})$$

where F^A and F^B are related by the change of mutual information and the rate of Shannon entropy in the cross-transitions according to equation (A.20).

The equations (A.23) and (A.24) are the same as previously derived for completely bipartite systems where two subsystems were coupled by regulatory transitions, but no cross-transitions were allowed^{19–21}. In the absence of cross-transitions, the original framework^{19–21} is recovered. Now, these inequalities have been generally derived for the systems where both regulatory and cross-transitions directly connecting the subsystems can take place. Such generalization is only possible if the definitions (A.18) and (A.19) are employed. Once the inequalities have been established, the same interpretation as in refs.^{19–21} can be used.

¹J. Schnakenberg, *Rev. Mod. Phys.* **48**, 571 (1976).

²K. Sekimoto, *J. Phys. Soc. Japan* (1997).

³C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).

⁴G. Crooks, *Phys. Rev. E* **60**, 2721 (1999).

⁵C. Van den Broeck, *J. Stat. Mech.* **2010**, P10009 (2010).

⁶M. Esposito and C. Van den Broeck, *Phys. Rev. E* **82**, 011143 (2010).

⁷U. Seifert, *Rep. Prog. Phys.* **75**, 126001 (2012).

⁸C. Van den Broeck and M. Esposito, *Phys. A: Stat. Mech. Appl.* **418**, 6 (2015).

⁹T. Tomé and M. J. de Oliveira, *Phys. Rev. E* **91**, 042140 (2015).

¹⁰D. Andrieux and P. Gaspard, *Proc. Natl. Acad. Sci.* **105**, 9516 (2008).

¹¹A. E. Allahverdyan, D. Janzing, and G. Mahler, *J. Stat. Mech.* **2009**, P09011 (2009).

¹²T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **109**, 180602 (2012).

¹³S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, *Phys. Rev. Lett.* **109**, 120604 (2012).

¹⁴S. Ito and T. Sagawa, *Phys. Rev. Lett.* **111**, 180603 (2013).

¹⁵A. C. Barato, D. Hartich, and U. Seifert, *J. Stat. Phys.* **153**, 460 (2013).

¹⁶J. Hoppenau and A. Engel, *EPL* **105**, 6 (2014).

¹⁷A. C. Barato and U. Seifert, *Phys. Rev. E* **90**, 042150 (2014).

¹⁸A. Barato and U. Seifert, *Phys. Rev. Lett.* **112**, 090601 (2014).

¹⁹D. Hartich, A. C. Barato, and U. Seifert, *J. Stat. Mech.* **2014**, P02016 (2014).

²⁰J. M. Horowitz and M. Esposito, *Phys. Rev. X* **4**, 031015 (2014).

²¹G. Diana and M. Esposito, *J. Stat. Mech.* **2014**, P04010 (2014).

²²J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, *Nat. Phys.* **11**, 131 (2015).

²³N. Shiraishi and T. Sagawa, *Phys. Rev. E* **91**, 012130 (2015).

²⁴C. Cafaro, S. A. Ali, and A. Giffin, *Phys. Rev. E* **93**, 022114 (2016).

²⁵T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **102**, 250602 (2009).

²⁶U. Seifert, *Eur. Phys. J. E* **34**, 26 (2011).

²⁷C. Van den Broeck, N. Kumar, and K. Lindenberg, *Phys. Rev. Lett.* **108**, 1 (2012).

²⁸F. Jülicher, A. Ajdari, and J. Prost, *Rev. Mod. Phys.* **69**, 1269 (1997).

²⁹C. Bustamante, D. Keller, and G. Oster, *Acc. Chem. Res.* **34**, 412 (2001).

³⁰A. W. C. Lau, D. Lacoste, and K. Mallick, *Phys. Rev. Lett.* **99**, 158102 (2007).

³¹J. M. Horowitz, T. Sagawa, and J. M. Parrondo, *Phys. Rev. Lett.* **111**, 010602 (2013).

³²P. Sartori, L. Granger, C. F. Lee, and J. M. Horowitz, *PLoS Comput. Biol.* **10**, e1003974 (2014).

³³Y. Demirel, *Entropy* **16**, 1931–1948 (2014).

³⁴P. Gaspard and E. Gerritsma, *J. Theor. Biol.* **247**, 672 (2007).

³⁵K. Kawaguchi, S. I. Sasa, and T. Sagawa, *Biophys. J.* **106**, 2450 (2014).

³⁶E. Gerritsma and P. Gaspard, *Biophys. Rev. Lett.* **05**, 163 (2010).

³⁷M. E. Fisher and A. B. Kolomeisky, *Proc. Natl. Acad. Sci.* **98**, 7748 (2001).

³⁸J. E. Baker, *J. Theor. Biol.* **228**, 467 (2004).

³⁹S. Liepelt and R. Lipowsky, *Phys. Rev. Lett.* **98**, 258102 (2007).

⁴⁰S. Liepelt and R. Lipowsky, *Phys. Rev. E* **79**, 011917 (2009).

⁴¹T. R. Barends, M. F. Dunn, and I. Schlichting, *Curr. Op. Chem. Biol.* **12**, 593 (2008).

⁴²M. F. Dunn, D. Niks, H. Ngo, T. R. M. Barends, and I. Schlichting, *Trends Biochem. Sci.* **33**, 254 (2008).

⁴³S. Raboni, S. Bettati, and A. Mozzarelli, *Cell. Mol. Life Sci.* **66**, 2391 (2009).

⁴⁴M. F. Dunn, *Arch. Biochem. Biophys.* **519**, 154 (2012).

⁴⁵A. N. Lane and K. Kirschner, *Eur. J. Biochem.* **129**, 571 (2005).

⁴⁶K. S. Anderson, E. W. Miles, and K. A. Johnson, *J. Biol. Chem.* **266**, 8020 (1991).

⁴⁷E. U. Woehl and M. F. Dunn, *Biochemistry* **Vol. 38**, 7131 (1999).

⁴⁸M. F. Dunn, V. Aguilar, P. Brzović, W. F. Drewe, K. F. Houben, C. A. Leja, and M. Roy, *Biochemistry* **29**, 8598 (1990).

⁴⁹P. S. Brzović, Y. Sawa, C. C. Hyde, E. W. Miles, and M. F. Dunn, *J. Biol. Chem.* **267**, 13028 (1992).

⁵⁰H. Ngo, N. Kimmich, R. Harris, D. Niks, L. Blumenstein, V. Kulik, T. R. Barends, I. Schlichting, and M. F. Dunn, *Biochemistry* **46**, 7740 (2007).

⁵¹H. Ngo, R. Harris, N. Kimmich, P. Casino, D. Niks, L. Blumenstein, T. R. Barends, V. Kulik, M. Weyand, I. Schlichting, and M. F. Dunn, *Biochemistry* **46**, 7713 (2007).

⁵²D. Loutchko, D. Gonze, and A. S. Mikhailov, *J. Phys. Chem. B* **120**, 2179 (2016).

⁵³T. R. Schneider, E. Gerhardt, M. Lee, P. H. Liang, K. S. Anderson, and I. Schlichting, *Biochemistry* **37**, 5394 (1998).

⁵⁴L. Blumenstein, T. Domratheva, D. Niks, H. Ngo, R. Seidel, M. F. Dunn, and I. Schlichting, *Biochemistry* **46**, 14100 (2007).

⁵⁵S. Raboni, S. Bettati, and A. Mozzarelli, *J. Biol. Chem.* **280**, 13450 (2005).

⁵⁶C. A. Leja, E. U. Woehl, and M. F. Dunn, *Biochemistry* **34**, 6552 (1995).

⁵⁷B. D. Bennett, E. H. Kimball, M. Gao, R. Osterhout, S. J. Van Dien, and J. D. Rabinowitz, *Nat. Chem. Biol.* **5**, 593 (2009).

⁵⁸A. N. Lane and K. Kirschner, *Biochemistry* **30**, 479 (1991).

⁵⁹N. Kishore, Y. B. Tewari, D. L. Akers, R. N. Goldberg, and E. W. Miles, *Biophys. Chem.* **73**, 265 (1998).

⁶⁰T. Hatano, *Phys. Rev. E* **60**, R5017 (1999).