EPFL

Thèse n° 10 048

Why do we need so many proteins? A physical insight into the collaboration of Hsp70 and DnaJ

Présentée le 9 juin 2023

Faculté des sciences de base
Laboratoire de biophysique statistique
Programme doctoral en physique

pour l'obtention du grade de Docteur ès Sciences

par

Adélaïde Alice MOHR

Acceptée sur proposition du jury

Prof. F. Mila, président du jury
Prof. P. De Los Rios, directeur de thèse
Dr P. Sartori, rapporteur
Prof. S. Goldt, rapporteur
Prof. A.-F. Bitbol, rapporteuse

École polytechnique fédérale de Lausanne

2023

Une petite impatience peut conduire
à de grandes imprudences.
— Tanguy, 2001

Acknowledgements

At the risk of being unoriginal, I would say that a PhD is, in the first place, a great human adventure. And like many adventures, the essence of this one is to be found in the people who have surrounded me throughout these four and a half years. First of all, I must thank Paolo De Los Rios for having, with infinite patience, supervised my work during all this time. In a few years, he managed to transmit his curiosity for biology, encourage each of my results and guide my work with kindness and support while leaving me the time to do and learn things in my own way.

I would also like to thank Stefano Zamuner who was always available to help me understand anything and everything at the beginning of this work. His generosity and knowledge were irreplaceable, whether in science or in learning how to make pasta. Similarly, Daniel Busiello has been of essential help in carrying out this PhD, his availability being exceeded only by his creativity in the kitchen, which I luckily or unluckily never tested.

During all those years, I shared my office with Solange Flatt who started her own PhD a few months before me. After Dire Strait's "Brothers in Arms", we were "Sisters in PhD", sharing the ups and downs of the PhD student life, our moments of discouragement as well as our best achievements. We supported each other, traveled in the name of food more than science, shared recipes... Without her, it would clearly not have been the same. The rest of the group, Shilling, Bruno, Davide, Alexandra, Hugo, Mathieu and Satyam have all been of great support and have widened my vision of both theoretical physics and biology.

My first contact with the world of biology was at the CIBM with Joao Duarte and continued with him in Lund for my Master thesis. He trusted me to carry out fascinating experimental work (on mice!), helped me learn to communicate with biologists, and allowed me to finally link my studies to the real world and health issues. I will never thank him enough for his trust, his support, and his words of wisdom: "Everything has an impact on the brain!"

Peu de personnes m'ont autant soutenu qu'Athos, qui, à sa manière bien particulière, a toujours été des plus encourageants. Sans en avoir l'air, tu as réussi calmer mes inquiétudes, à m'apprendre à relativiser mes problèmes et à gagner en flexibilité. Ton soutien et ta présence ont été inébranlables, dans cette épreuve comme dans toutes les autres : ton appartement à Bâle en plein covid, le ronco sans chauffage ni climatisation, les allers et retours interminables entre Lausanne et Bellinzona et bien sûr mon permis de conduire. Pour tout cela, je ne peux que te dire merci, *gattone*.

Je voudrais également adresser un énorme remerciement à ma famille, à ma sœur aînée Charlotte en particulier, qui m'a montré que les liens familiaux sont indestructibles. Malgré tes propres préoccupations, tu as toujours réussi à prendre le temps de m'écouter et de me conseiller. En plus de cela, ton soutien lors de la "dernière ligne droite" de l'écriture de cette thèse a été déterminant dans l'aboutissement de ce travail.

Et finalement, je voudrais remercier chaleureusement mes parents. Tout ce que j'ai pu faire, c'est grâce à vous. Des explications devant les locomotives à vapeur du Science Museum au cours d'anglais dans la campagne vaudoise, tout ce que vous avez fait pour moi, dès mon plus jeune âge, m'a permis de développer la curiosité suffisante pour mener à bien mes études et maintenant ce doctorat. Alors merci beaucoup. Merci d'avoir toujours laissé votre porte ouverte, merci d'avoir écouté, conseillé, consolé inlassablement. Merci pour votre présence et votre soutien indéfectible.

Ah et vous voyez, j'ai presque fait ma Tanguette finalement!

Lausanne, April 27, 2023 A. M.

Abstract

The Hsp70 cycle is a key element of protein homeostasis, which is essential to avoid protein aggregation and protein-related diseases. Despite many experimental observations of the interaction between Hsp70, its co-chaperone DnaJ and various substrates, little focus has been given so far to the development of general and predictive models of such interactions.

In this context, this thesis introduces different models to highlight the key features of the Hsp70/DnaJ system and provides a comprehensive description of its mechanism.

We first describe different selection mechanisms and their impact in out-of-equilibrium dynamics, emphasizing the importance of the catalytic discrimination as a necessary, though not sufficient, step to ensure a fast and accurate selection.

We then focus on the Hsp70 system, first by presenting an innovative model to explain the regulation mechanism of its ATPase activity and the synergistic effect of the substrate and DnaJ on this activity. We highlight how the synergistic effect can be seen as a selection step allowing Hsp70 to discriminate substrates.

Finally, we examine the complete Hsp70 cycle with an emphasis on the impact of 3-body interactions and propose the existence of a regulatory mechanism for DnaJ. Our resulting model allows us to highlight the mechanisms used by Hsp70 to act on misfolded proteins.

Key words: Hsp70 cycle | DnaJ | Proteins folding | Out-of-equilibrium | Kinetic model | Selection mechanisms | Regulatory mechanism | ATPase activity

Résumé

Le cycle de Hsp70 est un élément clé de l'homéostasie des protéines qui est essentielle pour éviter l'agrégation des protéines et les maladies qui en résultent. Malgré de nombreuses observations expérimentales sur l'intéraction entre Hsp70, sa co-chaperone DnaJ et différents substrats, le développement de modèles généraux et prédictifs de ces interactions n'a reçu que peu d'attention à ce jour.

Dans ce contexte, cette thèse propose différents modèles pour mettre en évidence les principales caractéristiques du système Hsp70/DnaJ et fournir une description complète de son mécanisme.

Nous décrivons en premier lieu différents mécanismes de sélection et leur impact lors de dynamiques hors équilibre, en soulignant l'importance de la discrimination catalytique comme étape nécessaire mais non suffisante pour garantir une sélection rapide et correcte.

Nous nous concentrerons ensuite sur le système Hsp70, en proposant tout d'abord un modèle innovant pour expliquer le mécanisme de régulation de son activité d'ATPase et l'effet synergique du substrat et de DnaJ sur cette activité. Nous soulignons comment cet effet synergique peut être considéré comme une étape de sélection permettant à Hsp70 de discriminer les substrats.

Finalement, nous examinons le cycle complet de Hsp70 en mettant l'accent sur l'impact des interactions à trois corps et proposons l'existence d'un mécanisme de régulation de DnaJ. Le modèle qui en résulte nous permet de mettre en évidence les mécanismes utilisés par Hsp70 pour agir sur les protéines mal repliées.

Contents

List of Figures

1 Introduction

From my biology courses in high school, I remember a very simple description of proteins. In my memory, they were simply the result of an alignment of building blocks, called amino acids, folded in a particular and organized structure. As I dabbled in biophysics, I discovered that proteins were significantly more complex than I initially thought.

The seemingly infinite possible combinations of amino acids allow proteins to take on a vast array of shapes, ranging from fibrous structures like collagen in connective tissues and bones to globular proteins like hemoglobin that transport oxygen to cells, and there are even proteins that form channels allowing elements to enter or leave cells. As their shape is so closely related to their function, proteins perform a vast variety of functions, being involved in various mechanisms from gene expression regulation to the immune system.

The same protein can take a multitude of different 3-dimensional structures, which in some cases is crucial to its function, and in other cases, can lead to severe and even fatal diseases. Proteins possess complex regulatory mechanisms, are able to sense and recognize elements [1], and can even treat information [2; 3]! Furthermore, a significant proportion of proteins in the human body (approximately 10% of the total protein mass [4]) exist with the sole purpose of assisting other proteins in their functions. In summary, proteins are far from being simple and far from being well understood.

Developing a better understanding of their structure and their interaction could eventually lead to life-changing discoveries for many people affected by protein-related diseases.

It is precisely for those reasons that I decided to dive into the world of proteins a few years ago. With this thesis, I hope to trigger some interest in the fascinating and complex particularities of proteins.

But let us first begin with a general description of what constitutes a protein.

1.1 Life and death of a protein

Proteins are synthesized by the ribosome during a process called translation. The ribosome translates a messenger RNA (mRNA) sequence into a sequence of amino acids that form a protein. The unique combination of those amino acids in a sequence and the intramolecular

forces resulting from that alignment, due to the amino acids chemical properties, lead the protein to fold in a specific 3-dimensional structure, which ultimately determines its function.

This functional state is called the native state and corresponds to a local free energy minimum, as illustrated in Figure 1.1, such that the protein should, in theory, naturally fall into this state given the correct initial conditions.

This native state, however, is a meta-stable state that can easily be perturbed by various forms of stress such as chemical stress, heat shocks, and oxidative stress [5]. In such cases, proteins may lose their native state and adopt a partially folded state or misfolded state, which is characterized by the exposure of hydrophobic amino acids to the environment. This exposure increases the likelihood of misfolded proteins binding with other also partially folded proteins [6; 7], resulting in the formation of aggregates.

Those aggregates can take different forms, depending on their organization and their content. While amorphous aggregates are disordered accumulations of misfolded proteins, oligomers and amyloid fibrils have a well-organized structure. Those aggregates can be highly stable due to their low free energy state, making them difficult to break apart.

To give the reader a better sense of the different possible structures for proteins and their associated energy level, a schematic representation of the energy landscape of protein is provided in Figure 1.1.

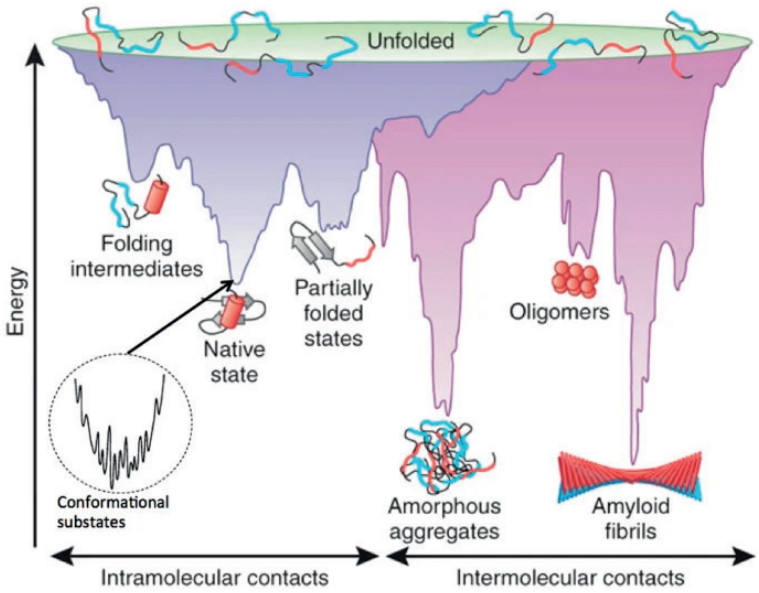

Figure 1.1: Energy landscape of protein structures. Taken from [8]

The structural transitions are encouraged or discouraged depending on their benefit to the cell, by a specific group of proteins, called chaperone proteins [9; 10] as illustrated in Figure 1.2.

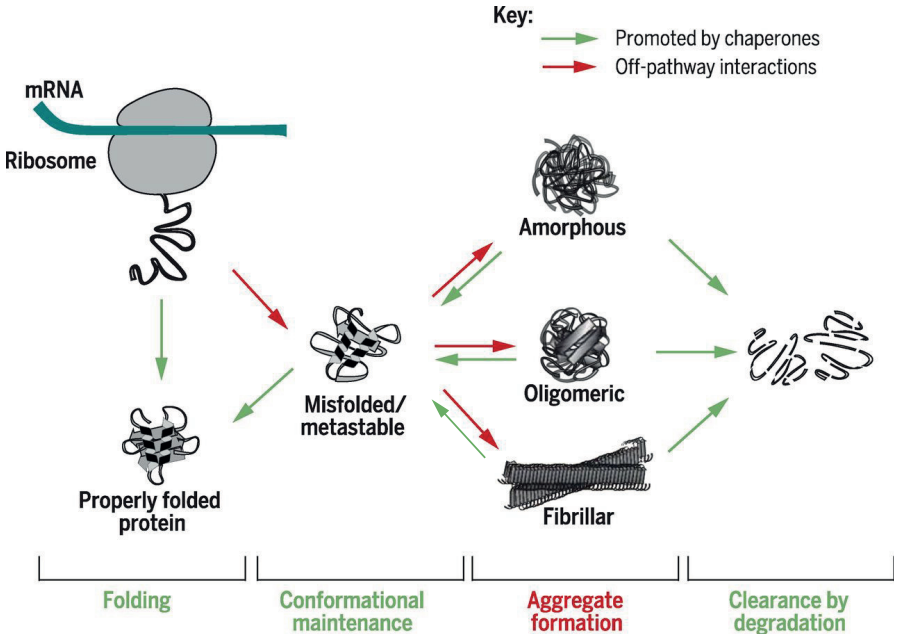

Figure 1.2: Life of proteins, influence by chaperones. Adapted from [11]

Chaperone proteins constitute a very diverse group of proteins that are responsible for ensuring the proper folding and function of other proteins and avoiding their aggregation. Originally, the role of chaperones was understood as a prevention mechanism [12]. By targeting misfolded proteins and holding them, chaperones could avoid aggregate formations by occupying the hydrophobic area of the protein.

However, it is now known that their primary function is to break down protein aggregates and malfunctioning oligomers via a mechanism called *entropic pulling* [10; 13]. Recent studies [14; 15] have also demonstrated that chaperones have the ability to break apart amyloid fibrils through this same mechanism, despite their increased stability.

Proteins aggregates that cannot be recovered are transported by chaperones to the lysosomes where they are degraded [16] and their amino acids are "recycled" for the synthesis of new proteins.

All these processes are part of the *proteostasis*, also called protein homeostasis, defined as the collection of biological pathways assuring the regulation of proteins.

However, this idealistic cycle does not always function properly and misfolded proteins and aggregates can accumulate in living systems with negative consequences.

1.2 When trouble arises

Given the critical role of proteins in so many different biological processes, any protein malfunction or loss of structure can has devastating consequences if it is not recovered. Indeed, the deregulation of protein structure has been linked to a number of diseases, including

neurodegenerative disorders and type 2 diabetes [17; 18].

For example, Alzheimer's disease is closely linked to the degradation of neuronal function and density, as well as the formation of extracellular amyloid fibrils that aggregate into amyloid plaques in the brain. [19]. These plaques have been identified as a possible cause of the disease [20] and therefore, targeting them may act on the progression of the disease. The recently FDA-approved drug called Leqembi intended to do just that. This drug based on Lecanemab-irmb, an antibody "directed against aggregated soluble and insoluble forms of amyloid beta" [21], has been shown to effectively reduce the amount of amyloid beta plaques in Alzheimer patients in an 18 months period [22] in pre-clinical studies.

While the recent FDA approval of Leqembi has brought new hope in the search for a cure for Alzheimer's disease, much remains to be understood about the root cause of protein aggregation and its role in diseases.

The decline of the proteostasis network, the network to which the chaperones belong, responsible for maintaining protein regulation and function, is believed to contribute to the formation of protein aggregates [23]. While this decline is associated with aging, the accumulation of stress and mutations, the underlying mechanism behind it remains unknown.

More generally, a complete understanding of the proteostasis network and its function in maintaining proteins in healthy conditions is lacking and further research is necessary to gain a deeper understanding of this critical network.

1.3 A physical approach

In order to gain a better understanding of the proteostasis network, we decided to concentrate on the 70-kDa heat shock protein (Hsp70) chaperone family, a crucial component of the network. Hsp70 is a highly conserved among living systems family of molecular chaperones carrying out a large variety of functions. It is widely recognized for its role in facilitating protein disaggregation and refolding [24]. It carries out its function in collaboration with two other proteins, called its co-chaperones, DnaJ and a nucleotide exchange factor (NEF).

Despite the large number of experimental observations that have been made on the impact of Hsp70 and DnaJ on misfolded proteins as well as the particularity of their interactions together, there is still a need for a better overall understanding. We hope that through presenting different models, we can emphasize the key features of the Hsp70/DnaJ system and provide a comprehensive description of their mechanism.

Because these proteins work together to target misfolded proteins, we first investigate the principles behind accurate substrate selection. By proposing a minimal model of substrate selection, inspired by the kinetic proofreading model proposed by Hopfield and Nino [25; 26], we aim to identify the essential elements of accurate and fast substrate selection. In this process, we demonstrate the importance of energy consumption as a vital component of the procedure, enabling the system to be more efficient and accurate.

We then turn our attention to the Hsp70 system, starting with a focus on Hsp70 ATPase activity. Hsp70s are ATP-dependent chaperones that use the energy from ATP hydrolysis to actively desegregate proteins. However, their ability to hydrolyze ATP, known as their ATPase activity, is strongly regulated by the presence of either a substrate protein or their co-chaperone, DnaJ [27]. The combination of substrate and DnaJ has been shown to significantly enhance Hsp70 ATPase activity in a synergistic manner [28; 29]. We therefore present an innovative model that can explain both the regulation mechanism of Hsp70 ATPase activity and the synergistic effect of substrate and DnaJ.

Finally, we focus on the role of DnaJ in collaboration with Hsp70. We illustrate the existence of a regulatory mechanism in DnaJ activity as well as study the impact of 3-body interactions on the dynamics between DnaJ, Hsp70 and the substrate. We examine how these elements affect substrate binding and its processing by Hsp70.

Before diving into the details, the following chapter will provide an overview of the specificity of out-of-equilibrium dynamics, which is the foundation of any biological system.

2 Equilibrium and out-of-equilibrium physics

In physics, equilibrium refers to a state in which a system is no longer undergoing any net changes. In thermodynamics, equilibrium means that the system has reached a stable macroscopic state from which no work can be extracted. Microscopically, all changes are compensating one another to keep the state visibly stable. Although this concept is very useful in the study of physical processes, it is less appropriate when considering biological systems. Life itself is intrinsically in a non-equilibrium state, and all living organisms rely on constant energy input and consumption to survive. It is therefore crucial to have an appropriate description of living systems phenomena in out-of-equilibrium conditions.

Before going into the details of non-equilibrium dynamics, let's first describe the properties of biochemical systems at equilibrium.

2.1 Detailed balance

Biological systems are usually represented by kinetic diagrams, where each node corresponds to a possible state of the system, and transitions between nodes represent the biochemical reactions of the system.

For a kinetic model to be at equilibrium, the following condition has to hold:

$$p_i^{eq} k_{ij} = p_j^{eq} k_{ji} \; \forall i, j \tag{2.1}$$

where p_i^{eq} is the equilibrium probability to be in state i and k_{ij} is the first order transition rate from state i to state j. This condition is called **detailed balance** and states that there is no net flux between any connected states i and j.

For states connected in a closed loop, called a cycle, this condition translates to

$$\frac{\prod_i k_{i,i+1}}{\prod_i k_{i+1,i}} = 1 \tag{2.2}$$

where the product of the rates going in one direction of the cycle is equal to the product of the rates in the opposite direction, assuming that the states are labeled such that the state i is

followed by state $i+1$.

The variation of Gibbs free energy of any chemical system is given by:

$$dG = -SdT + VdP + \sum_i \mu_i dn_i. \tag{2.3}$$

where S is the entropy of the system, dT the temperature variation, V the volume, dP the pressure variation, μ_i the chemical potential of the state i and dn_i the variation in population of state i. For biochemical systems, we usually assume that $dP = dT = 0$, such that:

$$G = \sum_i \mu_i n_i. \tag{2.4}$$

At equilibrium, each state must have the same chemical potential, therefore $\mu_i = \mu_j \quad \forall i, j$.

Knowing that the chemical potential of a given state i is expressed as $\mu_i = \mu_i^0 + k_B T \ln(c_i)$ [30], where μ_i^0 is the standard chemical potential and c_i the concentration of element in state i, we have that:

$$\mu_i - \mu_j = \mu_i^0 - \mu_j^0 + k_B T \ln\left(\frac{c_i}{c_j}\right) \underset{\text{at equ.}}{=} 0 \tag{2.5}$$

Implying that at equilibrium,

$$\frac{c_i^{eq}}{c_j^{eq}} = \exp^{-\beta(\mu_i^0 - \mu_j^0)}. \tag{2.6}$$

Furthermore, considering (2.1), we have $\frac{c_i^{eq}}{c_j^{eq}} = \frac{k_{ji}}{k_{ij}} = K_{ij}$,

$$K_{ij} = \exp^{-\beta \Delta G_{ij}^0} \tag{2.7}$$

The difference in chemical potential, $\mu_i^0 - \mu_j^0$ is usually considered equal to the difference in standard free energy ΔG_{ij}^0. This equality is formally wrong as the chemical potential is expressed in joules/mol or joule/molecule while the Gibbs free energy is expressed in joule. The term "partial free energy" is sometimes used instead of free energy to solve that discrepancy although this is sometimes considered to add to the confusion [31].

In reality, the equality is actually valid in the case of unitary concentration such that $c_i \mu_i^0 - c_j \mu_j^0 = \Delta G_{ij}^0$ with $c_i = c_j = 1$.

Substrate-to-product cycle

Let us illustrate the characteristics of a system at equilibrium by considering the most general reaction in biology: the transformation of a substrate S into a product P assisted by an enzyme as represented in Figure 2.1A. Although the usual description of this process is uni-directional, from substrate to product, all possible reactions should be considered in a thermodynamically consistent description.

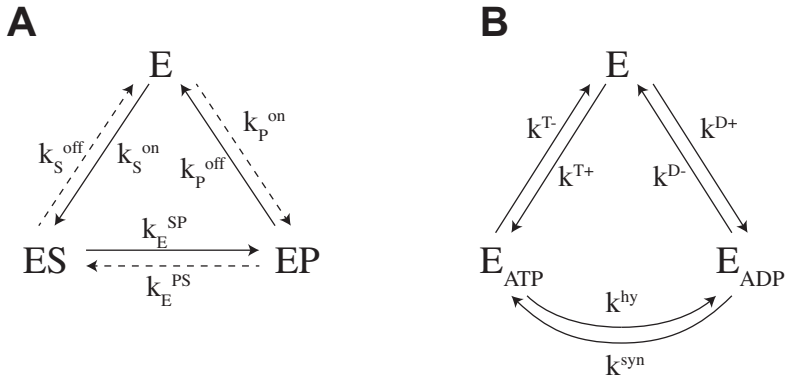

Figure 2.1: **A**: Schematic representation of a substrate to product cycle. Solid arrows represent the cycle going from substrate to product and dashed arrows the reverse cycle from product to substrate. **B**: Schematic representation of ATP hydrolysis induced by the enzyme.

In the case of the substrate-to-product cycle, the detailed balance condition as presented in (2.2) becomes:

$$\frac{[S]_{eq} k_S^{\text{on}} k_E^{SP} k_P^{\text{off}}}{[P]_{eq} k_P^{\text{on}} k_E^{PS} k_S^{\text{off}}} = 1 \tag{2.8}$$

where $k_S^{\text{on}}, k_S^{\text{off}}$ and $k_P^{\text{on}}, k_P^{\text{off}}$ are the binding and unbinding rates of the substrate, respectively product, to the enzyme and k_E^{SP}, k_E^{PS} the rate of product formation and inverse product formation when bound to the enzyme E.

Assuming that the transition between S and P can also happen without the enzyme, via rates k^{SP}, k^{PS}. The equilibrium concentrations of S and P are restrained by:

$$[S]_{eq} k^{SP} - [P]_{eq} k^{PS} = 0 \Longrightarrow \frac{[S]_{eq}}{[P]_{eq}} = \frac{k^{PS}}{k^{SP}} \tag{2.9}$$

Therefore, the equilibrium condition over the cycle becomes:

$$\frac{k^{PS} k_S^{\text{on}} k_E^{SP} k_P^{\text{off}}}{k^{SP} k_P^{\text{on}} k_E^{PS} k_S^{\text{off}}} = 1 \tag{2.10}$$

Considering (2.1), we also have

$$[E]_{eq}[S]_{eq} k_S^{\text{on}} - [ES]_{eq} k_S^{\text{off}} = 0 \tag{2.11}$$

Defining $K_S^d := \dfrac{k_S^{\text{off}}}{k_S^{\text{on}}}$ as the dissociation constants between E and S, we can easily express the equilibrium concentration of ES such that

$$[ES]_{eq} = \frac{[E]_{eq}[S]_{eq}}{K_S^d} \tag{2.12}$$

In addition, we have that

$$\Delta G^0_{E,ES} = k_B T \ln(K^d_S) \tag{2.13}$$

where $\Delta G^0_{E,ES} = \mu^0_{ES} - \mu^0_E - \mu^0_S$ is the molecular free energy difference between states E and ES, corresponding to the energy needed to bind the substrate.

Therefore, at the equilibrium, the concentration of each state is determined by the molecular free energy level, or chemical potential, of that state.

2.2 Out of equilibrium and NESS

Let us now push the system out of equilibrium, by assuming the system to have an inflow of substrate keeping the concentration of substrate constant such that $[S] > [S]_{eq}$ and a constant outflow of product such that $[P] \to 0$.

As only the concentration of S and P have been modified, (2.10) is still valid because the rates of the system remain identical. However, the detailed balance condition (2.8) is not holding anymore:

$$\frac{k^{PS} k^{\text{on}}_S k^{SP}_E k^{\text{off}}_P}{k^{SP} k^{\text{on}}_P k^{PS}_E k^{\text{off}}_S} = 1 \qquad \frac{[S] k^{\text{on}}_S k^{SP}_E k^{\text{off}}_P}{[P] k^{\text{on}}_P k^{PS}_E k^{\text{off}}_S} \neq 1 \tag{2.14}$$

Our system is therefore not able to reach equilibrium. It can however reach a steady-state where all time derivatives are equal to zero, called a non-equilibrium steady-state (NESS) [30].

The system can be described by a set of master equations:

$$\begin{aligned}
\frac{d[E]}{dt} &= [ES] k^{\text{off}}_S + [EP] k^{\text{off}}_P - [E]([S] k^{\text{on}}_S + [P] k^{\text{on}}_P) \\
\frac{d[ES]}{dt} &= [E][S] k^{\text{off}}_S + [EP] k^P_E S - [ES](k^{\text{off}}_S + k^S_E P) \\
\frac{d[EP]}{dt} &= [E][P] k^{\text{off}}_P + [ES] k^S_E P - [EP](k^{\text{off}}_P + k^P_E S)
\end{aligned} \tag{2.15}$$

such that the steady-state concentrations of the systems $([E]_{ss}, [ES]_{ss}, [EP]_{ss})$ can be found by solving:

$$\frac{d[E]}{dt} = 0, \quad \frac{d[ES]}{dt} = 0, \quad \frac{d[EP]}{dt} = 0. \tag{2.16}$$

The chemical potential difference between two states, as given in (2.5), is also not equal to zero anymore.

$$\mu_S - \mu_P = \mu^0_S - \mu^0_P + k_B T \ln\left(\frac{[S]}{[P]}\right) = k_B T \ln\left(\frac{[S]}{[P]} \frac{[P]_{eq}}{[S]_{eq}}\right) \tag{2.17}$$

This difference in chemical potential is referred to as the chemical force X due to its ability to push the system in a given direction.

Michaelis-Menten

The Michaelis-Menten scheme [32] is one of the first models describing the enzymatic reaction of product formation and is a special case of out of equilibrium system.

In this model, the enzyme E bind to the substrate S and directly release the product P such that:

$$\mathrm{E} + \mathrm{S} \underset{k_S^{\mathrm{off}}}{\overset{k_S^{\mathrm{on}}}{\rightleftharpoons}} [\mathrm{ES}] \xrightarrow{k^{cat}} \mathrm{E{+}P}$$

The absence of reverse reactions imposes a direction on the system making it by essence out of equilibrium.

To solve the system, one needs to solve the following set of master equations describing the dynamics of the system:

$$\begin{aligned} \frac{d[E]}{dt} &= [ES](k_S^{\mathrm{off}} + k^{cat}) - [E][S]k_S^{\mathrm{on}} \\ \frac{d[ES]}{dt} &= [E][S]k_S^{\mathrm{off}} - [ES](k_S^{\mathrm{off}} + k^{cat}) \end{aligned} \tag{2.18}$$

Two key hypotheses of the model allow to simplify it even further: the system is assumed to be at steady-state and the concentration of substrate is considered large enough so that $[S]_{tot}] \simeq [S]$. With this assumption, one can compute the rate of the reaction

$$v = [ES]k^{cat} = [E]_{tot}\frac{[S]k^{cat}}{[S] + K_M} \tag{2.19}$$

where $K_M = \dfrac{k_S^{\mathrm{off}} + k^{cat}}{k_S^{\mathrm{on}}}$ is the Michaelis constant.

This model is a simplification of the substrate-product cycle proposed in Figure 2.1A, where we assume an infinite source of substrate S and a sink of P such that $[P] = 0$. Additionally, the Michaelis Menten scheme treats the formation and release of products as a single coarse-grained step: $k^{cat} = \dfrac{k_E^{SP}k_P^{\mathrm{off}}}{k_E^{SP} + k_P^{\mathrm{off}}}$.

Similarly, the inverse reaction of product formation is assumed to be nonexistent: $k_E^{PS} \to 0$.

Under the assumption of $[P] \to 0$ and $k_E^{PS} \to 0$, the rate of product released from the cycle in Figure 2.1A is given by:

$$v = [EP]k_P^{\mathrm{off}} = [E]_{tot}\frac{k_P^{\mathrm{off}}k_S^{\mathrm{on}}k_E^{SP}[S]}{k_S^{\mathrm{on}}k_E^{SP}[S] + k_P^{\mathrm{off}}(k_S^{\mathrm{off}} + k_E^{SP} + k_S^{\mathrm{on}}[S])} \tag{2.20}$$

where $[EP]$ was found by solving the master equations in (2.15) and assuring that $[E]_{tot} = [E] + [ES] + [EP]$.

Arranging the equation to look more like the Michaelis-Menten equation in (2.19), we found

$$v = [E]_{tot} \frac{[S] \frac{k_E^{SP} k_P^{\text{off}}}{(k_E^{SP} + k_P^{\text{off}})}}{[S] + \frac{k_P^{\text{off}}}{k_E^{SP} + k_P^{\text{off}}} \frac{(k_S^{\text{off}} + k_E^{SP})}{k_S^{\text{on}}}} \tag{2.21}$$

where $\frac{k_E^{SP} k_P^{\text{off}}}{k_E^{SP} + k_P^{\text{off}}} = k^{cat}$ and $\frac{k_P^{\text{off}}}{(k_E^{SP} + k_P^{\text{off}})} \frac{(k_S^{\text{off}} + k_E^{SP})}{k_S^{\text{on}}}$ corresponds to the Michaelis-Menten constant for a more complex mechanism consisting of two first-order steps instead of one.

2.2.1 The case of ATP hydrolysis and ATP driven cycles

In biological systems, the primary source of energy comes from ATP hydrolysis. The transition from ATP to ADP is accompanied by the release of approximately 30 kJ/mol $\sim 12 k_B T$ [33] of energy through the unbinding of one phosphate group.

$$\text{ATP} \xrightarrow{k^{hy}} \text{ADP} + \text{Pi} \quad \Delta\mu = 12 k_B T$$

The role of ATP hydrolysis in biological systems is extremely diverse and can usually be associated with a gain of information about the system in question [2]. This is for example the case when considering protein transport [3] or accurate selection [34].

To demonstrate the role of ATP as the source of non-equilibrium, we will consider the biochemical reaction of the cycle of an ATPase E, represented in Figure 2.1B, favoring the transition from ATP to ADP. Initially, the ATP binds to the enzyme E and is then hydrolyzed. Finally, the ADP, resulting from the hydrolysis, is released.

At equilibrium, the detailed balance condition holds such that:

$$\frac{[ATP]_{eq} k^{T+} k^{\text{hy}} k^{D-}}{[ADP]_{eq} k^{D+} k^{\text{syn}} k^{T-}} = 1 \tag{2.22}$$

where k^{T-}, k^{T+} and k^{D-}, k^{D+}, are the unbinding and binding rates of ATP, respectively ADP, and k^{hy} and k^{syn} are the ATP hydrolysis and synthesis rates.

As for the substrate-to-product cycle, the system is pushed out of equilibrium when the ratio of $\frac{[ATP]}{[ADP]} \neq \frac{[ATP]_{eq}}{[ADP]_{eq}}$. More specially we can assume $\frac{[ATP]}{[ADP]} > \frac{[ATP]_{eq}}{[ADP]_{eq}}$.

In the cell, the actual ratio $\frac{[ATP]}{[ADP]} \sim 10$ [33], while the equilibrium ratio $\frac{[ATP]_{eq}}{[ADP]_{eq}} \sim \exp^{-\beta\Delta\mu} = \exp^{-12} \sim 10^{-6}$.

Out of equilibrium, the chemical force of a cycle containing an ATP hydrolysis transition is given by

$$X_T = \mu_T - \mu_D = k_B T \ln\left(\frac{[ATP]}{[ADP]} \frac{[ADP]_{eq}}{[ATP]_{eq}}\right) = k_B T \ln\left(\frac{\alpha}{\alpha_{eq}}\right) \tag{2.23}$$

where we defined $\alpha = \frac{[ATP]}{[ADP]}$ and similarly $\alpha_{eq} = \frac{[ATP]_{eq}}{[ADP]_{eq}}$.

To easily describe the transition from ATP to ADP via exchange, namely the unbinding of one nucleotide and the binding of the other one, we assume that the enzyme in apo-state, with no nucleotide attached, is very short-lived. This means that when an enzyme bound to ATP unbinds its nucleotide, it will quickly either bind to ADP or come back to its ATP state by binding ATP. This allows us to depict the transitions from one nucleotide to the other as a direct process, with rates given by:

$$\begin{aligned} k^{TD}([ATP],[ADP]) &= k^{T-}\frac{k^{D+}[ADP]}{k^{D+}[ADP]+k^{T+}[ATP]} \\ k^{DT}([ATP],[ADP]) &= k^{D-}\frac{k^{T+}[ATP]}{k^{D+}[ADP]+k^{T+}[ATP]} \end{aligned} \tag{2.24}$$

where k^{TD} and k^{DT} are the overall exchange rate from ATP to ADP, respectively ADP to ATP, k^{T-} and k^{T+} are the unbinding and binding rates of ATP, respectively (and similarly for the binding/unbinding rates of ADP). As we assume the concentration of ATP and ADP to be large enough not to limit the exchange, we can further simplify (2.24) to

$$\begin{aligned} k^{TD}(\alpha) &= k^{T-}\frac{k^{D+}}{k^{D+}+\alpha k^{T+}} \\ k^{DT}(\alpha) &= k^{D-}\frac{\alpha k^{T+}}{k^{D+}+\alpha k^{T+}} \end{aligned} \tag{2.25}$$

Free energy transduction

We mentioned above how ATP hydrolysis is the most common source of energy in biological systems. However, the reaction only releases energy in the form of heat, therefore biological systems must have found a way to take advantage of that release of energy to perform work. This is done via a process called **free energy transduction** [35].

To demonstrate this concept, let's consider a closed system with a substrate S and a product P where the substrate is in a more favorable configuration than the product. This translates to $\mu_S^0 < \mu_P^0$. At equilibrium, regardless of the enzyme binding ability, the free concentration of substrate will always be greater than the free concentration of product: $[S]_{eq} > [P]_{eq}$.

However, if the enzymatic conversion of substrate to product is coupled to the hydrolysis of ATP then the system can behave differently.

The joint reaction of ATP hydrolysis and product formation is, in out of equilibrium condition ($\alpha > \alpha_{eq}$), imposing the following direction to the cycle: $E \rightarrow ES \rightarrow EP$. Therefore it is favoring the transition from substrate to product against the chemical gradient.

Out of equilibrium, the chemical force of the hydrolysis reaction is given by $X_T = k_B T \ln\left(\frac{\alpha}{\alpha_{eq}}\right)$, therefore the rate of energy dissipated by the hydrolysis is:

$$\phi_T X_T \tag{2.26}$$

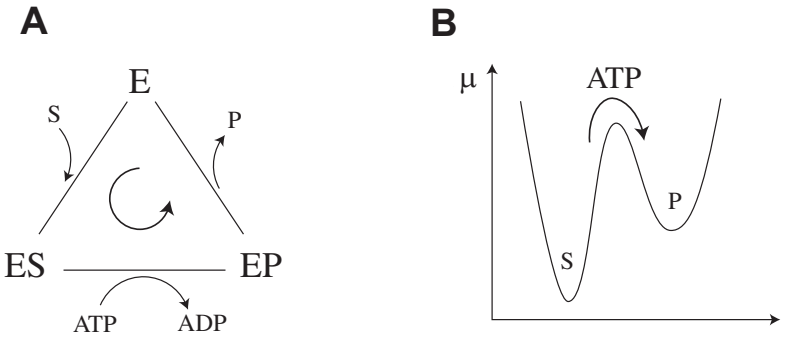

Figure 2.2: **A** Enzyme-assisted substrate to product cycle, power by ATP hydrolysis. **B** Schematic representation of the standard chemical difference between S and P and the impact of the ATP hydrolysis.

where ϕ_T corresponds to the flux of ATP hydrolysis.

Similarly, the chemical force needed to maintain a higher concentration of product over the substrate is $X_P = k_B T \ln\left(\frac{[P]}{[S]}\frac{[S]_{eq}}{[P]_{eq}}\right)$. The work necessary to maintain this situation is:

$$\phi_{SP} X_P \tag{2.27}$$

As long as $\phi_{SP} X_P$ is lower than $\phi_T X_T$, the energy released from the hydrolysis rate, then the system is able to maintain the non-equilibrium concentrations of S and P with an efficiency defined by:

$$\eta = \frac{\phi_{SP} X_P}{\phi_T X_T} = \frac{X_P}{X_T} \tag{2.28}$$

due to the fact that, in our example, $\phi_T = \phi_{SP}$.

This system is described by a similar set of master equations as the one proposed in (2.15), however with the addition of a chemical drive $\Delta\mu$.

Practically, this chemical drive due to the ATP hydrolysis is added to the transition from ES to EP such that:

$$\text{ES} \underset{k_E^{PS}}{\overset{k_E^{SP} e^{\beta\Delta\mu}}{\rightleftharpoons}} \text{EP with } \Delta\mu \geq 0$$

For $\Delta\mu = 0$, the detailed balance condition is respected and the system reaches equilibrium. Any positive value for $\Delta\mu$ breaks the detailed balance conditions and the system can only be in a NESS.

As expected, an increase in $\Delta\mu$ allows the enzyme to maintain a higher concentration of free product in solution compared to the equilibrium one, properly shifting the balance between substrate and product. This phenomenon can be observed in Figure 2.3A.

At equilibrium, the ratio of substrate to product $\frac{[S]_{eq}}{[P]_{eq}} = 5$ while for $\Delta\mu \sim 4.6 k_B T$ ($e^{\beta\Delta\mu} = 100$),

the ratio becomes $\frac{[S]}{[P]} = 0.5$.

In addition to the impact of $\Delta\mu$, we observe that an increase in the binding rate of P, k_P^{on}, and consequently in the unbinding rate k_P^{off} in order to keep a fixed value for $\mathrm{d}K_P^d$ the dissociation constant between E and P, also allows the product concentration to increase (Figure 2.3B).

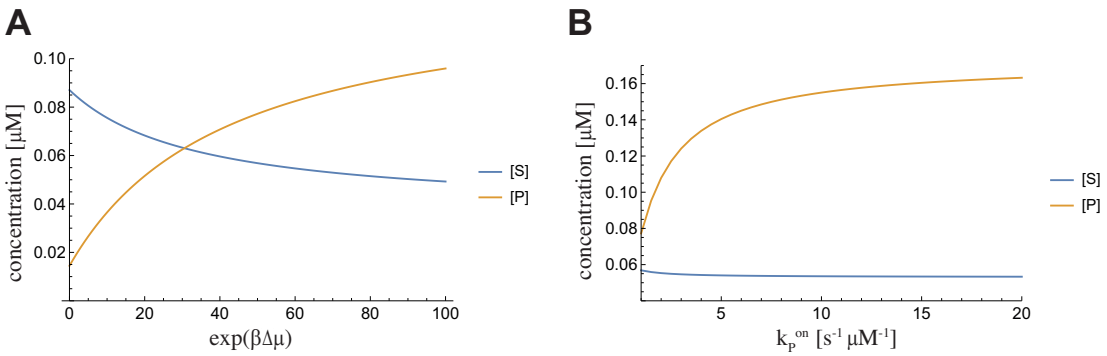

Figure 2.3: **A**: Effect of a chemical potential $\Delta\mu$ on the concentration of free substrate and product. **B**: Free concentration of S and P in function of the binding rate of P to the enzyme, while keeping K_P^d fixed in out-of-equilibrium dynamics

Indeed, the presence of net fluxes in out-of-equilibrium dynamics makes the systems sensitive to the rate kinetics, even when reaching a steady-state. Unlike at equilibrium, where the distribution of the system's population is solely determined by the free energy difference between states and therefore the equilibrium or dissociation constant, in a NESS, the kinetics play a significant role in determining the dynamics of the system.

2.2.2 Energy consumption for protein unfolding

We introduced above the importance of energy consumption in living systems, their ability to use that energy to perform work against a chemical gradient, and specific characteristics of out-of-equilibrium systems.

In the context of protein folding, these concepts play a crucial role in determining the stability and functionality of proteins. Proteins are only functional when they are in their native state. However, this state lacks stability; chemical and physical perturbations can cause proteins to lose their native structure and form aggregates characterized by a low energy level.

For proteins to fold back into their native state, they must be taken out of the aggregate and reach an unfolded state, associated with a high free energy level (Figure 1.1). This transition from a low energy level to a high energy level can only happen through free energy transduction as described above.

This is where chaperones come into play. By consuming ATP, chaperones provide the energy necessary for the transition from aggregates to unfolded proteins [36] and allow for the proper folding of proteins back into their native state.

Energy consumption also plays an essential role in the accuracy of selection processes, fundamental to many biological systems. In the following chapter, we will describe on how energy consumption enables accurate discrimination in biological systems.

3 Selection mechanisms

The work presented in this section comes from an article under review during the writing of this thesis. A pre-print version is available on bioRxiv [37]

3.1 Introduction

In biology, many essential processes are facilitated by specialized entities, called enzymes. These enzymes play a crucial role in catalyzing biological processes by enabling reactions to occur without undergoing any changes themselves. For instance, transporters allow the passage of proteins across membranes, pumps facilitate the movement of ions in and out of cells against chemical gradients, and the ribosome decodes the genetic information in the mRNA to synthesize proteins.

At the basis of all these processes, the enzyme needs to accurately choose its substrate, among many, usually similar, other proteins. A certain accuracy is, therefore, desirable in order to be efficient.

This is especially true for processes handling genetic information, which are known to be extremely accurate. The most common example is the synthesis of proteins by the ribosome where observations show 1 error is made every 1000 codons [38] and DNA replication with an astonishing 1 error every 10^{10} base pairs [39]. Such low error rates are clearly understandable as mistakes during gene replication or gene expression can lead to various issues from non-functioning proteins to the accumulation of misfolded proteins leading to protein-related diseases [40]. But the importance of accurate selection holds for any other biological process.

In a very general approach, the classic model of enzyme-substrate interaction is given by the Michaelis-Menten model, as discussed in Section 2.1 and represented in Figure 3.1. In this minimal model, the enzyme E binds to a substrate S to form the complex ES before the formation and the release of a product P. For simplicity, these two last steps are coarse-grained in one unidirectional transition.

Following this approach, the flux of product released by the enzyme is given by:

$$J = \frac{[E][S]}{K_S^d} k_{\text{cat}} \tag{3.1}$$

where K_S^d is the dissociation constant between E and S and k_{cat} the rate of product formation. We assume that the enzyme has the choice between binding a correct substrate, define as the right substrate, or an incorrect substrate, define as the wrong substrate. It can therefore form either a correct product or a incorrect product, called right or wrong product. The error rate is defined as the ratio between the flux of wrong product J_w and right J_r produce such that:

$$f = \frac{J_w}{J_r} = \frac{[W]}{[R]} \frac{K_R^d}{K_W^d} = \frac{[W]}{[R]} e^{\Delta G / k_B T}. \tag{3.2}$$

Therefore, in an environment with both right and wrong substrates in competition, an enzyme following the Michaelis-Menten scheme will operate with an error rate f directly related to the difference of free energy between choosing a right or a wrong substrate $\Delta G = G_R - G_W$.

This simple description of discrimination between right and wrong substrates is, however, not able to reproduce the low error rate observed in very precise biological systems where ΔG is usually around only a few $k_B T$.

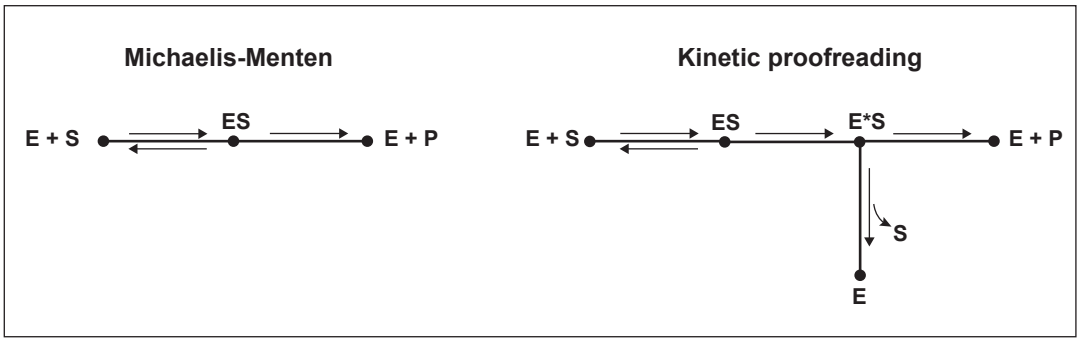

Figure 3.1: Left: Michaelis-Menten scheme where a substrate S is targeted by an enzyme E and released into a product P. Right: Kinetic proofreading model as proposed by Hopfield [25]. Similar to the Michaelis-Menten scheme but with the addition of an intermediate state $E^* S$ acting as a proofreading step.

To solve this paradox, Hopfield [25] and Nino [26] both proposed a model of **kinetic proofreading** with an additional intermediate state before the catalysis and product release (right side of Figure 3.1). In that case, the additional state $E^* S$ acts as a proofreading step and can discard wrong substrates and reset the system without any release of wrong products. This system, when optimized, drastically improves the error rate such that $f \sim \left(e^{\Delta G / k_B T}\right)^2$.

$$f_{Hop} := \frac{J_w}{J_r} \geq \frac{[W]}{[R]} \left(\frac{K_R^d}{K_W^d} \right)^2 \tag{3.3}$$

3.2 Model

Inspired by the kinetic proofreading model, we have developed a model of substrate selection thermodynamically consistent and which explicitly incorporates energy consumption. Indeed,

the directionality of the transition between ES and E^*S, as proposed by Hopfield and Nino, is due to energy consumption, e.i. ATP hydrolysis in the case of biological systems. With the idea of chaperones in mind, we assumed the enzyme to either be attached to ATP, in a state characterized by E, or to ADP, namely E^*. These two states are assumed to be present when the enzyme is bound to a substrate, as in Hopfield's model, but also when it is bound to a product P or on its own.

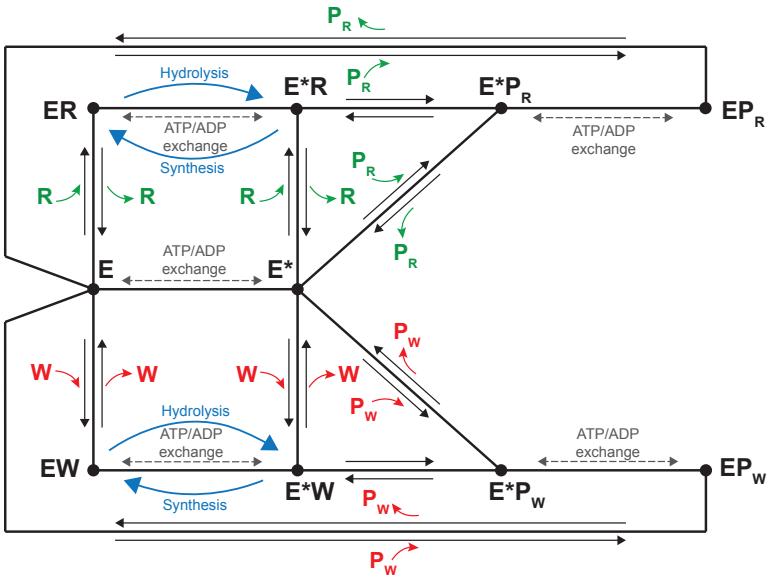

Figure 3.2: Model of substrate selection with a right and wrong path. The enzyme is represented by two states, E, representing the enzyme in its ATP-state, and E^*, corresponding to the enzyme in its ADP-state. The right and wrong substrates are denoted as R and W, respectively, with the corresponding products represented by P_R and P_W. The transitions between NTP and NDP-states of the enzyme can occur through nucleotide exchange (gray dashed arrows) or by NTP hydrolysis or NDP synthesis (blue arrows) when the enzyme is bound to either R or W. The latter occurs as a result of substrate stimulation. Other transitions in the diagram simply correspond to binding or unbinding to substrates or products.

In summary, the selection cycle of a substrate in our model, illustrated in Figure 3.2, is as follow: the enzyme in its ATP state, E, binds to a right or wrong substrate, R or W, to form the dimer ER, EW respectively. The presence of the substrate stimulates ATP hydrolysis to form E^*R or E^*W. After hydrolysis, the substrate can either be discarded via the proofreading pathway or transformed into a product, right or wrong, and then released. The enzyme goes back to its ATP state, either before or after releasing the product. In addition to the hydrolysis and synthesis, the transition from ATP to ADP state can occur through nucleotide exchange, namely unbinding of the current nucleotide and binding of the other one. These transitions are detailed in (2.25). All reactions in the model must be reversible and, at equilibrium, the detailed balance condition should be respected as introduced in Chapter 2.

The resolution of our system depends greatly on the transition rates between each state. An explicit description of each transition accounting for energy consumption and respecting detailed balance at equilibrium is therefore necessary.

The transitions between the ATP bound states and the ADP bound states have to consider a parameter $\alpha = \frac{[ATP]}{[ADP]}$, as shown in (2.25), allowing to tune the system out of equilibrium. Indeed, rate constraints are applied in order to respect detailed balance when $\alpha = \alpha_{eq} = 10^{-6}$ while increasing α brings the system out of equilibrium.

For all energy state α, we consider a fixed non-zero concentration for the substrates $[R]$ and $[W]$ and null concentrations for the products, $[P_R] = [P_W] = 0$, as they are assumed to be immediately removed. This choice of concentrations induces a net flux towards the release of products, even at equilibrium.

Once all transition rates are explicit, a set of master equations can be used to describe the temporal evolution of the system. For clarity, let us first consider a unique state N_i connected to different states to N_j, $j = 1,2,3....$ The evolution of the concentration of N_i is given by its master equation:

$$\frac{d[N_i]}{dt} = \sum_{j \neq i} k_{ji}[N_j] - k_{ij}[N_i] \tag{3.4}$$

with k_{ij} is the transition rate from N_i to any state N_j. As an example, the master equation for the state E is given by:

$$\frac{d[E]}{dt} = \sum_{S=R,W} [ES]k_S^{\text{off}} - [E][S]k_S^{\text{on}} + \sum_{P=P_r,P_w} [EP]k_P^{\text{off}} - [E][P]k_P^{\text{on}} + [E^*]k^{DT}(\alpha) - [E]k^{TD}(\alpha) \tag{3.5}$$

Extending this principle to all states N_i allows us to write the whole set of equations as

$$\frac{d\vec{N}}{dt} = M\vec{N} \tag{3.6}$$

where $\vec{N}$ is a column vector containing the concentration of each state and $M = M(\alpha, \{k_{ij}\})$ is the rate matrix which varies in function of α.

For an easier description of our system, we look for the non-equilibrium steady-state (NESS) solution of Equation (3.6), which is given by:

$$M\vec{N}_{NESS} = 0 \tag{3.7}$$

where N_{NESS} contains the steady-state concentrations of all the states present in our system. From those steady-state concentrations, we can compute the production fluxes

$$J_r = [EP_r]k_{P_r}^{\text{off}} + [E^*P_r]k_{P_r^*}^{\text{off}} - [E][P_r]k_{P_r}^{\text{on}} - [E^*][P_r]k_{P_r^*}^{\text{on}} \tag{3.8}$$

$$J_w = [EP_w]k_{P_w}^{\text{off}} + [E^*P_w]k_{P_w^*}^{\text{off}} - [E][P_w]k_{P_w}^{\text{on}} - [E^*][P_w]k_{P_w^*}^{\text{on}} \tag{3.9}$$

where k_X^{on} and k_X^{off} correspond to the binding and unbinding rates of the products, $X = P_r$ or P_w, as well as the error rate

$$f = \frac{J_w}{J_r} \tag{3.10}$$

These two quantities allow us to assess the accuracy and the velocity of different selection methods.

3.3 Different strategies

3.3.1 Kinetic proofreading

As our model is a thermodynamically consistent extension of Hopfield's kinetic proofreading model presented in Figure 3.1, we should be able to observe a similar behavior when using the same hypothesis.

To perform any selection, Hopfield's model is relying only on differences in dissociation constants and assumes the unbinding rates to be large compared to the other rates in the model.

To reproduce Hopfield's model, we assume that our system is symmetric between the right and wrong path, except for the unbinding rates. To achieve this, we set $k_W^{\text{off}} > k_R^{\text{off}}$ and $k_{W^*}^{\text{off}} > k_{R^*}^{\text{off}}$. Moreover, we assume that these unbinding rates are sufficiently large, such that $k_{W^*}^{\text{off}}, k_{R^*}^{\text{off}} \gg k_{SP}$, where k_{SP} is the production formation rate for both correct and wrong substrates. To minimize the error rate, Hopfield's assumption is that the system must reach its minimum when the transition between the two selection steps, corresponding to the hydrolysis rate, is near zero.

We, therefore, study the impact of the hydrolysis rate on the error and production rates in our system, as illustrated in Figure 3.3.
When the system is at equilibrium (red line), both the error rate and the production rate are constant. In the energy landscape description, increasing the hydrolysis rate is equivalent to lowering the energy barrier of the transition from ES to E^*S, without changing those states' energy levels. Therefore, the dynamics of the system are not impacted at equilibrium.

Furthermore, as no directionality is imposed between the two selection steps, the system cannot use the proofreading mechanism and only the last selection step has an effect on the error rate. This leads to the same selection mechanism as in the Michaelis-Menten scheme (left side in Figure 3.1) and the same error rate

$$f_{eq} = \frac{[W]}{[R]} \frac{K_R^d}{K_W^d}. \tag{3.11}$$

On the contrary, when the system is pushed out of equilibrium, the change in the hydrolysis rate modifies the system dynamics and favors a directionality of the transition between the two selection steps as in Hopfield's scheme.

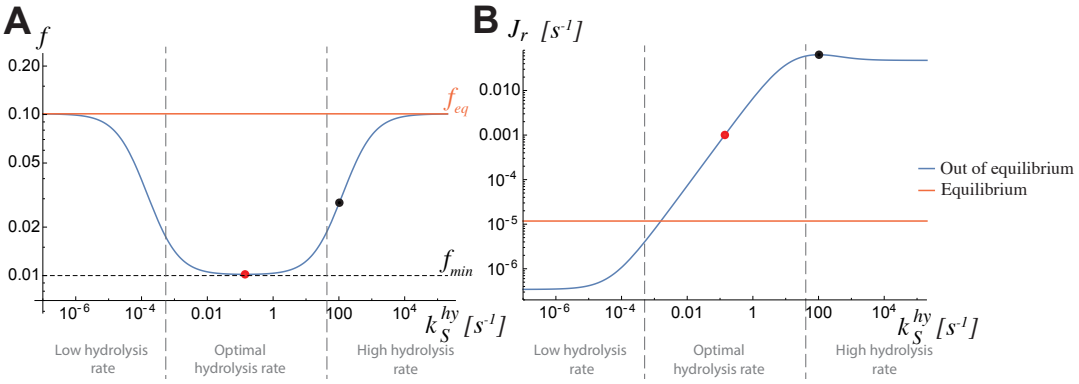

Figure 3.3: Hopfield-like selection system highly depends on available energy, α, and on the transition rate between ES and E^*S, k_S^{hy}. The figure shows **A** the error rate and **B** the production rate in function of the hydrolysis rate out of equilibrium in physiological conditions *i.e.* $\alpha = 10$ (blue) and at equilibrium $\alpha = \alpha_{eq}$ (orange). Red dots: minimum error rate under physiological conditions ($\alpha = 10$). Black dots: maximum production rate under physiological conditions ($\alpha = 10$). Dashed horizontal line: theoretical limit for the error rate under Hopfield's assumption (3.3).

This results in a decrease in the error rate, at least for an appropriate value of the hydrolysis rate. Indeed, as depicted in Fig. 3.3A, the relation between the hydrolysis rate and the error rate (blue line), can be separated into three different categories. For very low and very high hydrolysis rates, the error rate approaches its equilibrium values, meaning that the system is not able to take advantage of the additional selection step. However, there is an optimal hydrolysis rate between these two extremes, marked by a red dot, where the system reaches Hopfield's minimal error rate ($f_{min} = f_{Hop}$) given by Equation (3.3).

In Hopfield's model, the minimal error rate is obtained for close to zero transition rate between ES and E^*S, which is equivalent to the hydrolysis rate k_S^{hy} in our model. In our case, slowing down the hydrolysis rate first improves the error rate up to $f_{min} = f_{Hop}$. From that, any further decrease in the hydrolysis rate only leads to an increase in the error rate which eventually reaches its equilibrium value. There is therefore an optimal hydrolysis rate that allows to fully utilize the two selection steps, as illustrated in Figure 3.4.

If the hydrolysis rate is too slow, the enzyme will mostly bind the substrate from its ADP state E^* (instead of its ATP state E), thus removing the first selection step from the cycle. On the contrary, a large hydrolysis rate, as stated by Hopfield [25], pushes the substrate directly towards the E^*S state, thus bypassing the ES state. In both cases the system does not benefit from the error correction that would be provided by the first selection step. This explains why very high and very low values of k_S^{hy} reach the same error rate as the equilibrium one.

On top of lowering the system error rate, energy consumption can also speed up the system (Fig.3.3B). At equilibrium, $\alpha = \alpha_{eq}$, the production rate is constrained by the imbalance between the substrate and product concentrations. Out of equilibrium, the large concentration of ATP is pushing the enzymes in their ATP state, therefore decreasing the concentration of E^*.

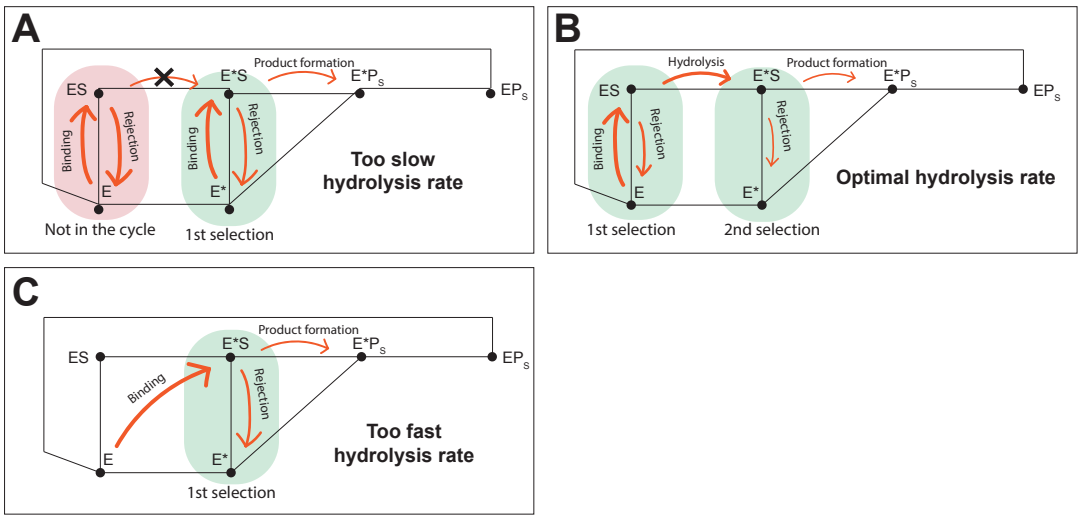

Figure 3.4: Scheme of the selection pathways following Hopfield kinetics for different values of the hydrolysis rate. **A** the slow hydrolysis rate prevents the substrate from entering the selection pathway from the E state. Therefore the system is using only one selection step (between E^* and E^*S, implying a high error rate. **B** the optimal hydrolysis rate allows the substrate to enter the selection pathway from the E state and make the most of the two selection steps leading to a low error rate. **C** with a fast hydrolysis rate, the substrate entering the selection is directly brought to the state E^*S where it can be rejected or directly proceed to product formation. This prevents the system from benefiting from the two selection steps, leading to a higher error rate.

Because the substrate enters the system from the E^* at low hydrolysis rate (Figure 3.4A), its low concentration limits the flux of substrate processed by the enzyme. This leads to a very low production rate, even lower than in equilibrium conditions.

As the hydrolysis rate increases, the substrate starts to enter the system from the E state (Figure 3.4A) and they are processed at a higher pace due to the relatively fast transition from ES to E^*S. As described above, the system also starts to utilize both selection steps, therefore also lowering the error rate and favoring the release of the right product. These three considerations together explain the increase in the production rate until it reaches a maximum value marked as a black dot in Figure 3.3B.

Finally, when k_S^{hy} is too high, the accuracy of the system starts to be low enough to lead to a small decrease in the production rate, due to the system releasing an increasing amount of wrong products.

In this Hopfield-like scheme, the system is not able to achieve both minimum error rate and maximum production rate at the same time. Decreasing the hydrolysis rate allows the system to fully utilize Hopfield's two selection steps, but also slows down the substrate's progress in the proofreading pathway, resulting in high accuracy but a production rate that is not

optimal (represented by the red dots in Figure 3.3A and B) and far from the maximum value (represented by the black dot on Figure 3.3B) by orders of magnitude, in agreement with [34]. This trade-off between speed and accuracy, often accompanied by energy dissipation, is commonly seen in models of information processing systems such as proofreading or error correction [41; 42; 43; 44; 45; 46], sensory adaptation [47], and polymer synthesis [48; 49; 50].

3.3.2 Use of the induced-fit mechanism - recognition of the substrate

The kinetic proofreading model allows to reach error rates as low as the ones observed in biological systems. However, such low error rates are reached at the cost of slowing down extensively the selection system. Biologically, most systems require both a low error rate and a rapid release of products. This observation leads us to explore another mean of discrimination, based on the induced-fit model [51; 52]. This model proposes that the interaction between the enzyme active site and the substrate is very flexible with the enzyme undergoing configurational changes to accommodate the substrate. Those modifications are not restricted to the binding site alone and can lead to a completely new configuration of the enzyme with a different activity. In that case, it is said that the enzyme is allosterically regulated by the substrate [53; 54].

With this idea, the enzyme can "recognize" the nature of the substrate: binding to the right substrate leads to the correct configurational changes, usually promoting the hydrolysis of the ATP molecule attached to the enzyme, while the wrong substrate is not able to promotes those changes. This allows the right substrate to proceed faster in the system than the wrong one. From a physical point of view, the enzyme is acting as a Maxwell demon [2; 3], where its action depends on the result of the measurement.

We model this phenomenon by simply increasing the hydrolysis rate when the enzyme is bound to the right substrate compared to the hydrolysis rate induced by binding to the wrong one, such that $k_R^{\text{hy}} \gg k_W^{\text{hy}}$. This difference between the hydrolysis rates can be explained by the enzyme possessing two different allosteric states, one inactive, with a low ATPase activity, and one active with a high ATPase activity. The presence of the right substrate tilts the enzyme allosteric balance toward the active state while the impact of the wrong substrate is limited. This idea of hydrolysis acceleration will be further developed in the following chapter.

To isolate this selection method, called **catalytic discrimination** due to the substrate's ability to catalyze the hydrolysis rate, we assume that all transitions, except the hydrolysis rate, are identical for the right and wrong substrate.

As the dissociation constants of binding the right and the wrong substrate are equal: $K_R^d = K_W^d$ and $K_{R^*}^d = K_{W^*}^d$, the states ER and EW and the states E^*R and E^*W have the same energy level. The only difference is the height of the energy barrier to go from ES to E^*S as the presence of the right substrate significantly lowers that energy barrier, as represented on Figure 3.5.

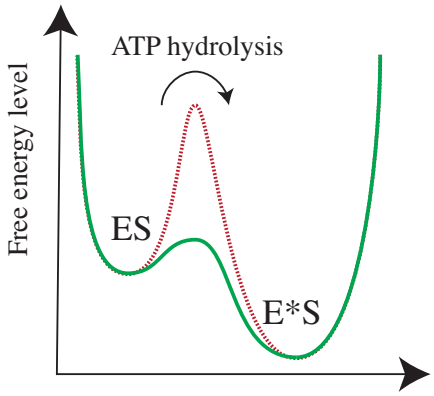

Figure 3.5: Scheme of the effect of the catalytic discrimination of the energy landscape between ES and E^*S. The red dashed line corresponds to the case with the wrong substrate while the green one, to the right substrate

As we already established, the discrimination at equilibrium can only depend on the difference in free energy level. With this discrimination method, the system, therefore, needs to be pushed out of equilibrium to perform any kind of discrimination between the right and the wrong substrate as seen in Figure 3.6. Indeed at equilibrium, the error rate is only determined by the ratio of right and wrong substrates available to the system. Pushing the system out of equilibrium by increasing the parameter α, allows to take advantage of the energy barrier difference between the transition from ER to E^*R and EW to E^*W and favors the right substrate.

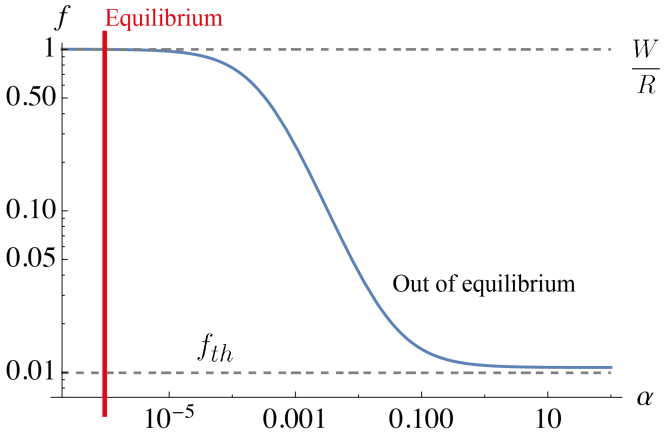

Figure 3.6: Error rate f in a catalytic discrimination scheme as a function of the available energy, α. Catalytic discrimination is only effective away from equilibrium. The upper dashed line corresponds to the error rate at equilibrium where no selection is effective, while the lower dashed line is the error rate limit (3.12) when the available energy is approaching ∞. The red vertical line represents the equilibrium condition with $\alpha = \alpha_{eq}$.

In the limit of an infinite amount of energy available to the system ($\alpha \to \infty$ and $\alpha_{eq} \to 0$), the

error rate based on the catalytic discrimination is given by:

$$f_{th} = \frac{W}{R} \frac{k_W^{\text{hy}}}{k_R^{\text{hy}}} \frac{(k_R^{\text{hy}} + k_S^{\text{off}})}{(k_W^{\text{hy}} + k_S^{\text{off}})} \tag{3.12}$$

where k_S^{off} is the unbinding rate for both W and R, and k_W^{hy}, k_R^{hy}, are the hydrolysis rate for EW, respectively ER. This formula corresponds to the asymptotic limit reached in Fig.3.6 and is very similar to the theoretical error rate found in [50] when considering similar discrimination mechanisms.

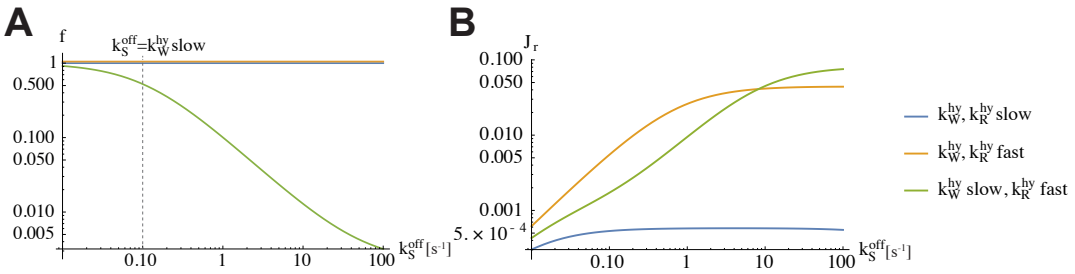

Figure 3.7: Error rate (**A**) and production rate (**B**) for slow hydrolysis rate for W and R (blue), fast hydrolysis rates for both W and R and slow hydrolysis rate for W and fast hydrolysis rate for R. Slow hydrolysis rate = 0.1 s^{-1}, fast hydrolysis rate = 100 s^{-1}. For a selection mechanism based on the catalytic discrimination, the fidelity is dictated by a slow hydrolysis rate for incorrect substrate, k_W^{hy}, a fast hydrolysis rate for correct substrate k_R^{hy}, and a fast enough unbinding rate k_S^{off}.

From (3.12), we know that the error rate mainly depends on three rates (k_S^{off}, k_R^{hy} an k_W^{hy}) and their relative values to each other. While the difference in the hydrolysis rate allows the system to be selective, this condition is not enough, as displayed in Figure 3.7A. Indeed, a very slow unbinding rate k_S^{off} prevents the discrimination step to be efficient and still maintains a high error rate. Increasing the unbinding rate such that $k_S^{\text{off}} > k_W^{\text{hy}}$, activates the selection step and lowers the error rate. This result is completely intuitive, a wrong substrate should have the time to unbind the enzyme before the hydrolysis, while the right substrate should continue in the process, hence $k_W^{\text{hy}} < k_S^{\text{off}} \leq k_R^{\text{hy}}$.

Interestingly and counter-intuitively, the production rate also benefits from a high value of k_S^{off} as seen in Figure 3.7B. Indeed, the detailed balance condition imposes the ATP synthesis rate to be inversely proportional to the unbinding rate such that $k_S^{syn} \propto \frac{1}{k_S^{\text{off}}}$. Therefore, a low value of k_S^{off} imposes a fast synthesis rate k_S^{syn} trapping the substrate between the ES and E^*S states and reducing the production rate. As k_S^{off} increases, k_S^{syn} decreases, and substrates begin to leave the E^*S state to be correctly processed by the enzyme.

This phenomenon is however limited as pushing the unbinding rate k_S^{off} to a really high value will eventually impair the enzyme's ability to bind any substrate and will decrease both the error rate and the production rate.

Despite this, there is a whole range of value for k_S^{off} where both the error rate and the production rate are optimized, leading to an absence of trade-off between accuracy and velocity with this selection scheme.

While our description of catalytic discrimination is insightful, it is not yet sufficient to realistically describe real biological systems. First, the configurational changes needed to trigger the hydrolysis rate most likely also modify the affinity between the substrate and the enzyme. Secondly, our observation of the production rate where the system is able to produce ATP for low unbinding rates is clearly biologically unrealistic. To accurately depict real biological processes, we must therefore make some additions to the selection mechanism.

3.3.3 Combining mechanisms for optimal performance

We previously introduced two means of discrimination, organized in different manners. The first one is based on the repetition of energetically discriminating steps, separated by a dissipating transition. In the second one, the affinities between the substrates and the enzyme are all equal, and only the energy-dissipating transition act as a unique discriminating step.

There is no reason to separate these two selection mechanisms as the presence of the first shouldn't impair the second. It is therefore interesting to allow these two mechanisms to work simultaneously. For that matter, let's consider a unique system where both selection schemes can be turned on or off using two parameters, λ and ω. λ is the parameter controlling the selection based on the kinetic proofreading, and ω the catalytic discrimination, such that

$$\lambda := \frac{k_W^{\text{off}}}{k_R^{\text{off}}} = \frac{k_{W^*}^{\text{off}}}{k_{R^*}^{\text{off}}} \qquad \omega := \frac{k_R^{\text{hy}}}{k_W^{\text{hy}}} \tag{3.13}$$

$\lambda = 1$ implies no selection based on the free energy difference and $\lambda > 1$ means that the system is using a kinetic proofreading scheme, similarly when $\omega = 1$ no selection based on the difference in the energy barrier of the hydrolysis transition is possible and $\omega > 1$ turns on the catalytic discrimination.

By modifying these two parameters, we can transition from one selection strategy to the other but also allow both to act on the system simultaneously.

In Figure 3.8, we investigate the interplay between the production rate and the error rate in the phase space defined by the parameters $\lambda \in [1, 100]$ and $\omega \in [1, 1000]$ under three different energy conditions, $\alpha = \alpha_{eq}, \alpha = 0.005$ and $\alpha = 10$.

At equilibrium, the phase space is represented by a vertical line, with both systems having the same production rate driven by the chemical force due to the unbalance of the substrate-to-product ratio. The system relying on catalytic discrimination is not able to perform any selection ($f = \frac{[W]}{[R]} = 1$), as previously observed. On the other hand, the kinetic proofreading is able to utilize one of its two selection steps such that $f \sim 0.1$.

As α increases, both systems can move away from their equilibrium state and differences

between them arise. While the kinetic proofreading scheme (KP) is mostly able to improve its accuracy, the catalytic discrimination (C) mostly increases the system velocity.

However, the combination of the two (KP+C) leads to an improvement of both the accuracy and the velocity of the selection without any threshold. In that case, the system benefits from both strategies by first being able to discard the wrong substrate on multiple occasions due to the proofreading, and secondly by accelerating the correct substrate into the product formation thanks to the catalytic discrimination.

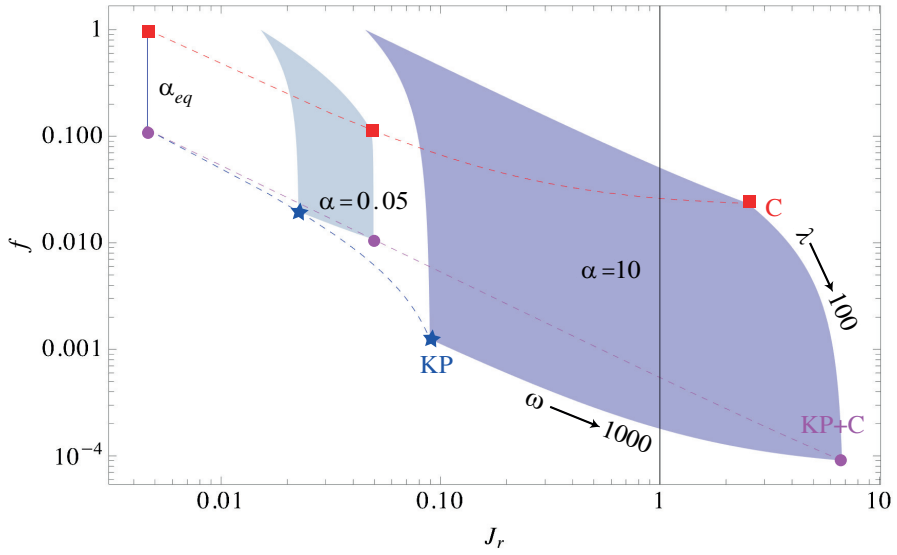

Figure 3.8: Error and production rates phase space for $\omega = \frac{k_R^{hy}}{k_W^{hy}} \in [1,1000]$, $\lambda = \frac{k_W^{off}}{k_R^{off}} = \frac{k_{W^*}^{off}}{k_{W^*}^{off}} \in [1,100]$ and different energy conditions (equilibrium: $\alpha = 10^{-6}$, non-equilibrium: $\alpha = 0.005$ and $\alpha = 10$, physiological conditions). Blue star: KP scheme ($\lambda = 100$ and ω=1), red square: catalytic discrimination C ($\lambda = 1$ and $\omega = 1000$) and purple dot: KP + catalytic discrimination C ($\lambda = 100$, $\omega = 1000$).

In the limit of an infinite amount of energy available, $\alpha \to \infty$, $\alpha_{eq} \to 0$, we can once again compute the theoretical error rate.

$$f \sim \frac{W}{R}\frac{k_W^{hy}}{k_R^{hy}}\frac{(k_R^{hy}+k_R^{off})(k_{SP}+k_{R^*}^{off})}{(k_W^{hy}+k_W^{off})(k_{SP}+k_{W^*}^{off})} = \frac{W}{R}\frac{1}{\omega}\frac{(\omega k_W^{hy}+k_R^{off})(k_{SP}+k_{R^*}^{off})}{(k_W^{hy}+\lambda k_R^{off})(k_{SP}+\lambda k_{R^*}^{off})} \tag{3.14}$$

where k_R^{hy}, k_W^{off} and $k_{W^*}^{off}$ were replaced to introduce the parameters λ and ω. Assuming a pure kinetic proofreading model, $\omega = 1$ and $\lambda \gg 1$, the error rate can be simplified as

$$f \sim \frac{W}{R}\frac{(k_W^{hy}+k_R^{off})(k_{SP}+k_{R^*}^{off})}{(k_W^{hy}+\lambda k_R^{off})(k_{SP}+\lambda k_{R^*}^{off})} \to \frac{W}{R}\frac{1}{\lambda^2} \tag{3.15}$$

in the limit of slow hydrolysis and production rates (k_W^{hy}, $k_{SP} \ll k_R^{off}$, $k_{R^*}^{off}$). This result is equiva-

lent to the minimal error rate obtained in a pure kinetic proofreading state, (3.3).

Similarly, with a purely catalytic system ($\lambda = 1$ and $\omega \gg 1$) we recover (3.12):

$$f \sim \frac{W}{R}\frac{1}{\omega}\frac{(\omega k_W^{\text{hy}} + k_R^{\text{off}})\cancel{(k_{\text{SP}} + k_{R^*}^{\text{off}})}}{(k_W^{\text{hy}} + k_R^{\text{off}})\cancel{(k_{\text{SP}} + k_{R^*}^{\text{off}})}} \tag{3.16}$$

Therefore, the general equation for the error rate, given by (3.14) contains the important elements of the two past formulas for the error rate. Interestingly, both selection mechanisms are working independently without interfering with each other. While the impact of catalytic discrimination is limited by

$$f \underset{\omega\to\infty}{\longrightarrow} \frac{W}{R}\frac{(k_{\text{SP}} + k_{R^*}^{\text{off}})}{(k_W^{\text{hy}} + \lambda k_R^{\text{off}})(k_{\text{SP}} + \lambda k_{R^*}^{\text{off}})}, \tag{3.17}$$

the kinetic proofreading can, in theory, always improve the accuracy by increasing λ:

$$f \underset{\lambda\to\infty}{\longrightarrow} 0. \tag{3.18}$$

As the increase in the energy available to the system improves the accuracy and the velocity, it is interesting to observe how this additional energy is used in each selection strategy. If we consider the chemical drive of the system $X_T = k_B T \ln(\alpha/\alpha_{eq})$, we can determine how the energy is used, favoring accuracy or velocity. To achieve this, we define the selection and production efficiency, γ_f and η_{J_r} respectively, as

$$\gamma_f = \frac{k_B T \ln(f^{(eq)}/f)}{k_B T \ln(\alpha/\alpha_{eq})} \tag{3.19}$$

$$\eta_{J_r} = \frac{k_B T \ln(J_r/J_r^{(eq)})}{k_B T \ln(\alpha/\alpha_{eq})} \tag{3.20}$$

where $1/f$ is considered the fidelity of the system.

Figure 3.9 provides insights into how KP, C and KP+C systems use energy to improve their production rate J_r or the fidelity $1/f$. Across all energy conditions, the KP system is favoring fidelity over speed, as demonstrate by a selection efficiency always higher than the production one ($\gamma_f > \eta_{J_r}$).

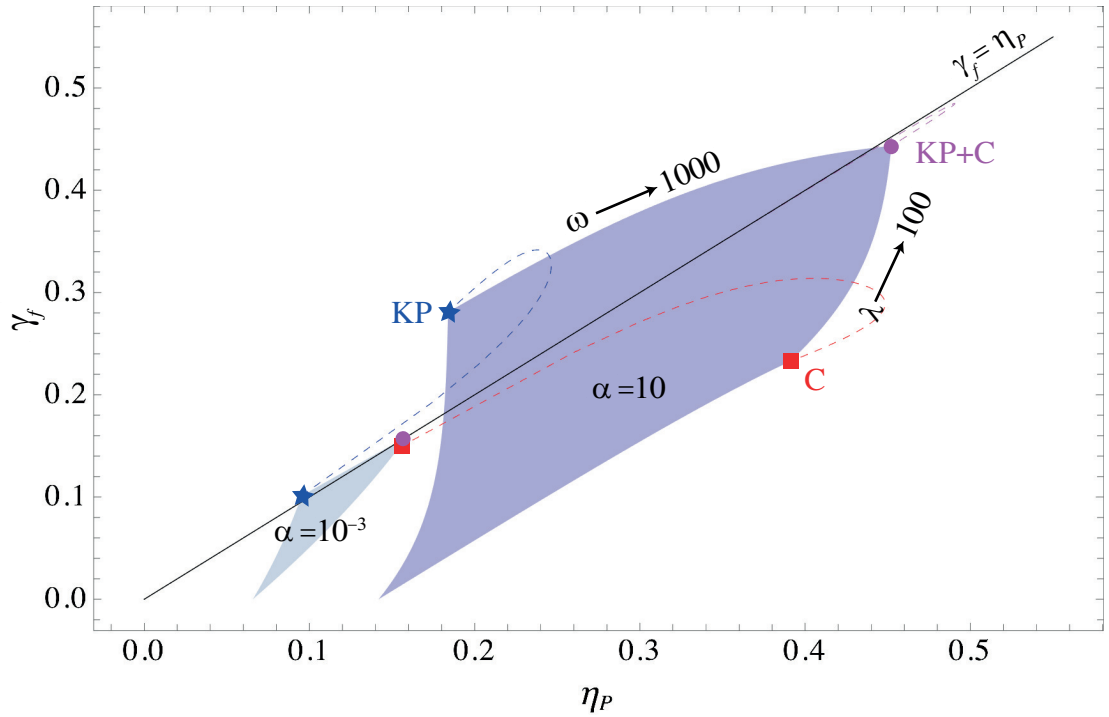

Figure 3.9: Production and selection efficiencies phase space for $\omega \in [1, 1000]$, $\lambda \in [1, 100]$ and different energy conditions ($\alpha = 10^{-3}, 10$). Black line: $\gamma_f = \eta_P$, blue star: KP scheme ($\lambda = 100$ and ω=1), red square: catalytic discrimination C ($\lambda = 1$ and $\omega = 1000$) and purple dot: KP + catalytic discrimination C ($\lambda = 100$, $\omega = 1000$).

On the other hand, the C system always consistently exhibits a higher production efficiency compared to its selection efficiency, $\eta_{J_r} > \gamma$.

The combined KP+C system offers a unique advantage, with both selection and production efficiencies improving with the energy input. Additionally, the KP+C system is more energy efficient than both the KP and C systems, as it exhibits a higher selection efficiency than the C system and a higher production efficiency than the KP system across all energy conditions.

Interestingly, all these three systems reach their optimal efficiencies (max γ and max η_{J_r}) in energy condition lower than $\alpha = 10$ and the additional energy input is not able to improve the systems further. Therefore, these systems have a certain tolerance to variation in energy conditions, as minor changes in α do not have a significant effect on their performance.

3.3.4 Example of selection in biological systems

As real life examples, is worth noting that protein synthesis and DNA replication, biological systems known for their high accuracy, actually rely on a combination of the kinetic proofreading mechanism and catalytic discrimination.

The accuracy of protein synthesis is dependent on the ribosome's ability to select the correct aminoacyl-tRNA complex matching the targeted mRNA anti-codon. This process involves an

initial selection step where only cognate and near-cognate codons are accepted, followed by a proofreading step [55; 56; 57], as proposed by Hopfield. After the initial selection, the interactions between the cognate codon and the correct anti-codon induce changes in the complex configuration, leading to a fast GTPase activation [55], needed to reach the proofreading step, while the near-cognate codon is not able to trigger such changes.

Hopfield initially considered DNA replication as an example of a proofreading mechanism in nature [25]. However, it is now widely accepted that the transitions from the polymerase activity, which assembles nucleotides, and the exonuclease activity, which removes incorrectly incorporated nucleotides, rely on an induced-fit mechanism [58]. Therefore, like protein synthesis, DNA replication relies on a combination of kinetic proofreading and catalytic discrimination [59].

3.4 Conclusion

In this chapter, we proposed a general model for the enzymatic selection and processing of substrates. Our model incorporates explicit energy-dependent transitions, which enable us to drive the system out of equilibrium by tuning the available energy. In this framework, we have evaluated several selection strategies, based on the kinetic proofreading, as proposed by Hopfield and Nino [25; 26], and on catalytic discrimination [34; 44] relying on the induced-fit mechanism and allosteric regulations [53].

The kinetic proofreading mechanism was found to be highly effective in reaching high levels of fidelity, but at the cost of slowing down the process significantly. On the contrary, catalytic discrimination has the advantage to induce a fast release of the product out of equilibrium, but cannot reach the same low error rate as the kinetic proofreading.

By incorporating both selection mechanisms in our model, we were able to reach a high fidelity and a fast production rate at the same time. The combined effect of both selection strategies resulted in a lower error rate than that observed in the kinetic proofreading scenario and a higher production rate than with reliance solely on catalytic discrimination.

We also could highlight the importance of energy consumption in biological processes. By explicitly considering the ATP consumption, we could observe how our system behaves in different energy conditions, particularly when far from equilibrium. We were also able to highlight how the different selection strategies use the available energy and how effective they are in this respect.

Finally, the prevalence of similar selection mechanisms in real biological systems highlights the significance of considering them in our model. This also suggests that a combination of these selection mechanisms could also be prevalent in systems that are not necessarily known for their high accuracy, making our proposed model applicable to a wide range of situations.

4 Hsp70 chaperone and its co-chaperone DnaJ

4.1 Introduction to Hsp70

We previously mentioned the importance of maintaining proteins in their native structure. Indeed, it is only when a protein is in its native structure that it can assure its particular function. However, the protein's native state is not the most stable state accessible to the protein, and external stresses, such as heat shock, chemical stress, or oxidative stress, are enough for a protein to lose its native structure.

The loss of the native structure is characterized by a misfolded state where the protein is not able to carry out its function anymore. Apart from the loss of function which clearly has a negative impact on the cells, the presence of multiple misfolded proteins can lead to the formation of protein aggregates which have a toxic impact on the cells and can lead to diseases.

Protein homeostasis, which includes all the biological processes allowing proteins to function correctly, is maintained, among others, by a specific class of proteins known as chaperones.

Among all classes of chaperones, one of the most essential and abundant in cells is the Hsp70 family. Hsp70s are 70k Dalton (70 kDa) heat shock proteins first known for their role in preventing protein aggregation during heat shock, hence their name. Hsp70s are able to unfold misfolded proteins as well as disassemble protein aggregates via the entropic pulling mechanism [10; 13].

These proteins are composed of two distinct domains, a nucleotide-binding domain (NBD) and a substrate-binding domain (SBD) attached by a linker (Figure 4.1). The SBD comprises two elements: a β-basket where substrates bind and an α-helical subdomain that acts as a lid on top of the β-basket. The NBD can either bind to ATP or ADP, influencing the overall configuration of the chaperone. When the NBD is bound to ATP, the SBD docks onto the NBD in an open configuration allowing rapid binding and unbinding of substrates [63; 64]. Binding to ADP leads to important configurational changes allowing the SBD to undock [61] and the closure of the α-helical lid onto the β-basket. In this configuration, the closed configuration of the SBD highly impairs the binding and unbinding of substrates with an overall higher affinity compared to the ATP state [65].

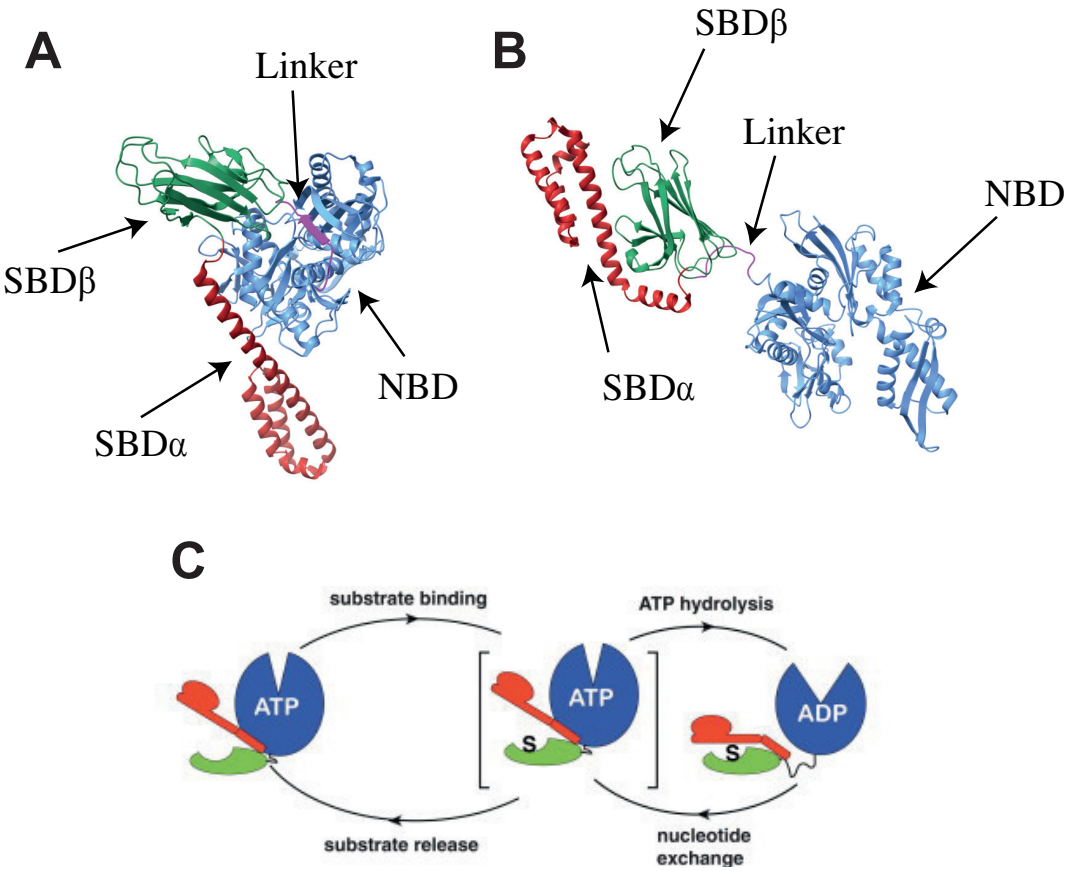

Figure 4.1: Hsp70 structures in ATP and ADP states, with, in blue, the nucleotide-binding domain (NBD), in the green, the β-basket of the substrate binding domain (SBD), and in red the α-helical lid of the SBD. **A**: Crystal structure of Hsp70 in ATP state with an open SBD (PDB ID code 4B9Q) from [60] **B**: Crystal structure of Hsp70 in ADP state with a close SBD (PDB ID code 2KHO) from [61]. **C**: Schematic representation of the closure of the SBD on a substrate, figure taken from [62]

In the process of targeting misfolded proteins, Hsp70 is usually assisted by two other proteins called its co-chaperones: DnaJ and the nucleotide exchange factors (NEF). The role of the NEF is to promote the ADP to ATP exchange after the hydrolysis whilst DnaJ is responsible for bringing the substrate to Hsp70 and further increasing Hsp70 ATPase activity. More details about DnaJ-like protein will be given in Section 4.3. For now, we will only state that DnaJ contains a highly conserved domain called the J-domain which interacts with Hsp70 and a substrate binding domain targeting misfolded proteins. Due to the presence of the J-domain, DnaJ is part of the J-domain protein (JDP) family and is sometimes referred to as a JDP.

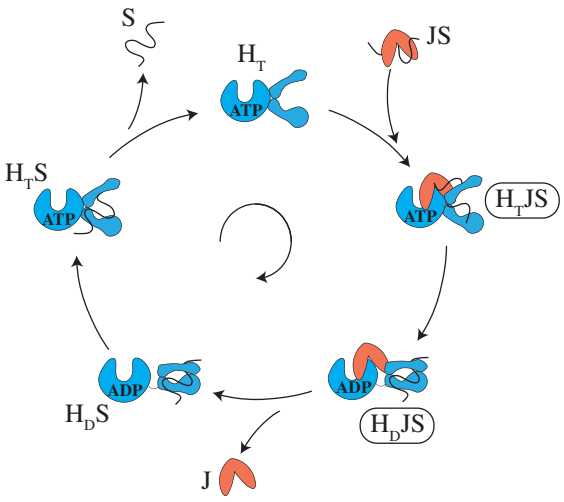

Figure 4.2: Traditional representation of the Hsp70 cycle assisted by DnaJ. Hsp70 is represented in blue and changes its configuration depending on the nucleotide attached to it: H_T for the ATP state and H_D for the ADP state. DnaJ, J, in red, targets the substrate S, represented as a floppy strand in black and brings it to Hsp70, in blue.

The ATPase cycle of Hsp70, depicted in Figure 4.2, typically starts with the binding of DnaJ and a substrate, the target protein that Hsp70 is interacting with, onto the open SBD, with Hsp70 bound to ATP. The presence of DnaJ and the substrate stimulates ATP hydrolysis, shifting the Hsp70 in its ADP state. The ATP hydrolysis switches the configuration of the SBD from an open to a closed one keeping the substrate locked between the β-basket and the α-helical lid.

The configurational changes in Hsp70 due to the hydrolysis also impair its affinity for DnaJ, which is pushed out of the trimer quickly after the hydrolysis.

Finally, the ADP-ATP exchange, release of ADP and its replacement with ATP on Hsp70, allows the SBD to open and the substrate to escape. The transition from an ADP state to an ATP state is usually promoted by the NEF but can also happen by itself as illustrated in Figure 4.2.

4.2 Regulation of the ATPase activity

4.2.1 Introduction

The Hsp70 ATPase activity is an essential part of the Hsp70 cycle. This energy-consuming step pushes the system in out of equilibrium dynamics and imposes a directionality in the cycle. However, Hsp70 by itself has a very low ATPase activity with a basal rate of around 6×10^{-4} s^{-1} [28] and relies on the presence of other proteins to stimulate its ATPase activity.

Experiments have shown that the presence of Hsp70 co-chaperones, DnaJ, [28; 66; 67] or substrates increases the Hsp70 ATPase activity, leading to an acceleration of 5 to 20-fold compared to the basal rate [29; 68]. When both DnaJ-like proteins and substrates are in solution, the effect on the ATP hydrolysis is drastically increased with a rate able to reach nearly 2 s^{-1} [28]. This phenomenon was intensely investigated in [29] where the author shows higher ATPase stimulation when DnaK (bacteria Hsp70) and DnaJ were in solution with a peptide allowing the formation of a ternary complex rather than with a single-motif peptide designed to only bind DnaK. The same phenomenon was observed in [28] where the use of peptide instead of a full-length substrate didn't allow the system to reach the same hydrolysis rate in the presence of DnaJ.

The ATP hydrolysis stimulation is believed to come from structural changes in Hsp70·ATP which are triggered by binding to a substrate [62; 69] or the J-domain of DnaJ [70]. However, there are no explanations for the synergistic effect these two proteins have when they are able to interact with each other. Therefore, we want to propose a comprehensive and general model able to explain the increase in the hydrolysis rate, as well as the synergistic effect DnaJ and the substrate have on the ATPase activity.

4.2.2 Model

Two allosteric states model

In our model (presented in Figure 4.3), Hsp70 bound to ATP can freely fluctuate between two allosteric configurations, H_1 and H_2, corresponding to an inactive and a highly active state regarding ATP hydrolysis ability, respectively. Both configurations are able to bind to a substrate S and/or Hsp70 co-chaperone DnaJ, J, to form dimers represented by H_iS, H_iJ, $i = 1,2$, and trimers. The trimers can either be in a fully connected state where each protein is bound to the other two ($\boxed{H_iJS}$) or in a 2-bond trimer, H_iJS, characterized by the presence of only two bonds connecting the trimer.

The difference between the two trimer states is represented in Figure 4.4 where A represents the fully connected trimer, and B a trimer formed without contact between S and J.

In our model, Hsp70 is able to bind independently to S and J to form H_iS, H_iJ respectively, or directly to the dimer JS reaching the H_iJS state. Furthermore, the dimer or trimer formation doesn't prevent Hsp70 to transition from its inactive (H_1) to its active state (H_2) or vice-versa. There is therefore an equilibrium between the H_1 and H_2 states, which is marked by the

transitions between the red and green planes of Figure 4.3.

Finally, in a way similar to [71], we assume that the chaperone's ability to trigger ATP hydrolysis depends on its allosteric state, with $k_{1,hy}$ the hydrolysis rate for H_1 and $k_{2,hy}$ the hydrolysis rate for H_2.

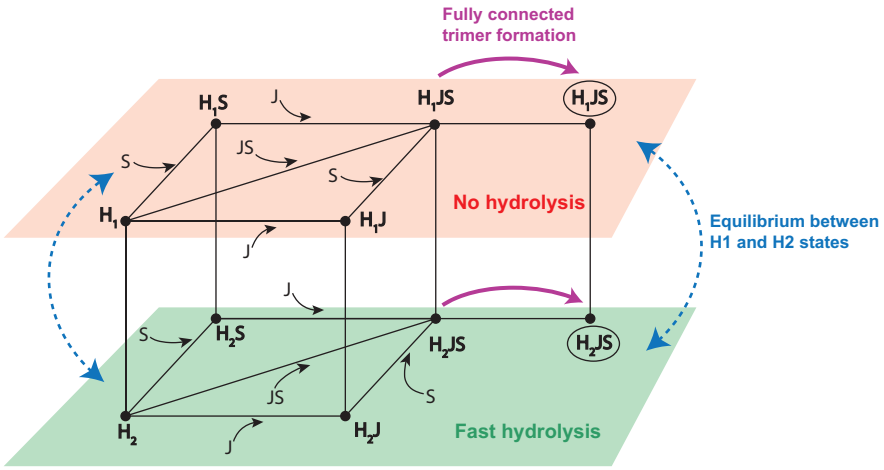

Figure 4.3: Pre-equilibrium model of Hsp70·ATP interaction with substrate S and its co-chaperone Hsp40, J. H_1 and H_2 (in the red and green planes) are the two allosteric configurations of Hsp70·ATP. In configuration H_1, ATP hydrolysis is blocked while in configuration H_2 the hydrolysis is a fast process.

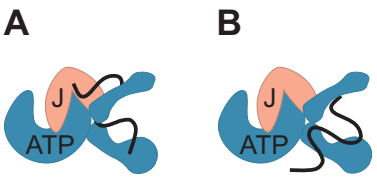

Figure 4.4: Schematic representation of the Hsp70 in ATP configuration, Hsp40, and substrate trimer . **A**: fully connected trimer, **B**: one of the three possible 2-bond trimers.

In order to explain the apparent acceleration of the ATP hydrolysis, we assume that H_2 can hydrolyze ATP extremely fast to complete the ATP cycle, so that the hydrolysis rate $k_{2,hy} = k^{max}$ while H_1 configuration is blocking the hydrolysis, leading to $k_{1,hy} \simeq 0$. The overall hydrolysis rate observed is then simply given by the contribution of the two allosteric states:

$$\begin{aligned} k_{hy}^{obs} &= \frac{[H_1]_{tot} k_{1,hy} + [H_2]_{tot} k_{2,hy}}{[H_1]_{tot} + [H_2]_{tot}} \\ &= \frac{[H_2]_{tot}}{[H_1]_{tot} + [H_2]_{tot}} k^{max} \end{aligned} \tag{4.1}$$

where $[H_1]_{tot}$, $[H_2]_{tot}$, corresponds to the total concentration of Hsp70 in configuration 1, respectively 2, regardless of the elements attached to it. The basal hydrolysis rate, of Hsp70

alone, is simply determined by the natural equilibrium between the active and inactive states. In order for a substrate or a J-domain protein to stimulate the ATPase activity, the equilibrium between active and inactive states needs to be tilted toward the active state. This will be the case if the active state has a higher affinity for the substrate and the J-domain protein than the inactive state.

Intuitive understanding of the synergistic effect

As we already mentioned, J-protein and substrates work together to highly stimulate Hsp70 ATPase activity. Their joint impact on the hydrolysis rate is then higher than the combination of their individual effect. However, this phenomenon seems only to take place when the two proteins are able to bind to each other, as shown in [29].

Before going into detailed mathematics, let us first propose an intuitive explanation of this phenomenon.

We assume that the observed hydrolysis rate is simply due to the proportion of Hsp70 in its active states (H_2) compared to its inactive states (H_1). Binding to a substrate S or to a J-domain protein J shifts the equilibrium between those two states toward the active state provoking an increase in the hydrolysis rate. In the presence of both S and J, the binding of the two proteins to Hsp70 results in the formation of the HJS or $\textcircled{HJS}$ complex. This binding further shifts the ratio towards the H2 states, resulting in an increased rate of hydrolysis.

Without synergy/affinity: If S and J are not able to bind to each other, they both act independently and the probability of them binding to Hsp70 is directly linked to their bulk concentration. They can form the 2-bond trimers H_1JS and H_2JS but for relatively low concentration of J and S, most of the Hsp70 in solution will be found in dimers with either J or S.

With synergy/affinity: If J and S can bind to each other, the formation of the dimer JS modifies the whole system. Indeed, when one protein of this dimer binds to Hsp70, the other one, due to the proximity with Hsp70 can easily bind as well. The same phenomenon is also happening for the HS dimer when binding to free J and for the HJ dimer with free S. In other terms, the localization of the three proteins in a small space creates a local concentration ρ of these proteins which is high enough compared to the dissociation constants to allow for the formation of the fully connected trimer $\textcircled{HJS}$. Thus, the co-localization stabilizes the fully connected trimer and increases the proportion of Hsp70 simultaneously bound S and J.

Mathematical details

Although we are considering an intrinsically out-of-equilibrium system, due to the release of energy during ATP hydrolysis, we decided to limit ourselves to a pre-equilibrium description of the system. This assumes that the transitions described in Figure 4.3 are fast compared to the hydrolysis rate. In addition, the ADP states of Hsp70 are neglected due to the assumption of fast nucleotide exchanges after the hydrolysis.

At equilibrium, all transitions should respect detailed balance. This simple condition imposes, for example,

$$k_{1,2}[H_1] = k_{2,1}[H_2] \Longrightarrow [H_2] = K_e[H_1] \tag{4.2}$$

where $K_e = \frac{k_{1,2}}{k_{2,1}}$ is the equilibrium constant between H_1 and H_2 and $k_{1,2}$, $k_{2,1}$ are the transition rates between H_1 and H_2, respectively H_2 and H_1. In the same way, the equilibrium concentrations of $H_i S$ is given by satisfying the detailed balance condition:

$$[H_i][S]k_{i,S}^{\text{on}} = [H_i S]k_{i,S}^{\text{off}} \Longrightarrow [H_i S] = \frac{[H_i][S]}{K_{i,S}^d} \qquad i = 1,2 \tag{4.3}$$

where $K_{i,S}^d = \dfrac{k_{i,S}^{\text{off}}}{k_{i,S}^{\text{on}}}$ is the dissociation constant between H_i and S. Similarly, the other dimers equilibrium concentrations are

$$[H_i J] = \frac{[H_i][J]}{K_{i,J}^d} \quad [JS] = \frac{[J][S]}{K_{J,S}^d} \tag{4.4}$$

where $K_{i,J}^d$ is the dissociation constant of H_i with J for $i = 1,2$ and $K_{J,S}^d$ the dissociation constant between S and J.

The equilibrium concentrations of the trimer are separated in two cases. Either the trimer is formed with two bonds linking the three proteins one after the other, or with three bonds with the three proteins all attached to each other forming the fully connected trimer. In the first case, three configurations are possible depending on the protein that is forming the two bonds.

$$[H_i JS] = [H_i][J][S]\left(\frac{1}{K_{i,S}^d K_{i,J}^d} + \frac{1}{K_{J,S}^d K_{i,J}^d} + \frac{1}{K_{J,S}^d K_{i,S}^d}\right) \tag{4.5}$$

where the terms on the right-hand side correspond to a trimer with Hsp70 bound to both S and J (Figure 4.4), then with J forming the two bounds and the last with S.

In the 2-bond trimer configuration, the two proteins not interacting with each other are very close in space. This co-localization, pushing them to bind, is associated to an apparent high concentration ρ such that

$$\boxed{[H_i JS]} = \frac{[H_i][J][S]\rho}{K_{i,S}^d K_{i,J}^d K_{J,S}^d} \tag{4.6}$$

Prevention of the synergistic effect of S and J is achieved by impairing their binding ability. In that case, assuming $K_{j,S}^d \to \infty$, the trimers concentrations are modified such that:

$$[H_i JS] = \frac{[H_i][J][S]}{K_{i,S}^d K_{i,J}^d} \tag{4.7}$$

$$\boxed{[H_i JS]} = 0 \tag{4.8}$$

with $i = 1,2$ representing the two allosteric states.

Similarly to [72], we also assume that protein complexes containing more than three proteins can be formed. For example, Hsp70 in the dimer HJ may still bind to a substrate, even if that substrate is already bound to another DnaJ. This will form the $HJJS$ complex, where the proteins are connected as such: $J-H-S-J$. However, as those complexes become significant only for high enough concentration of substrate and/or DnaJ, they are ignored in the following description here but are still considered in the model. More details are provided in the appendices.

To give the full hydrolysis rate equation, let's first compute $[H_1]_{\text{tot}}$ and $[H_2]_{\text{tot}}$. With eqs. (4.2) and (4.4) to (4.6), we have that:

$$\begin{aligned}
[H_1]_{\text{tot}} &= [H_1] + [H_1S] + [H_1J] + [H_1JS] + \boxed{[H_1JS]} \\
&= [H_1]\left(1 + \frac{[S]}{K_{1,S}^d} + \frac{[J]}{K_{1,J}^d} + \frac{[J][S]}{K_{1,S}^d K_{1,J}^d} + \underbrace{\frac{[J][S]}{K_{J,S}^d K_{1,J}^d} + \frac{[J][S]}{K_{J,S}^d K_{1,S}^d} + \frac{[J][S]\rho}{K_{1,S}^d K_{1,J}^d K_{J,S}^d}}_{=0 \text{ if no affinity between } S \text{ and } J}\right) \\
&= [H_1]F_1(J,S)
\end{aligned} \tag{4.9}$$

$$\begin{aligned}
[H_2]_{\text{tot}} &= [H_2] + [H_2S] + [H_2J] + [H_2JS] + \boxed{[H_2JS]} \\
&= [H_1]K_e\left(1 + \frac{[S]}{K_{2,S}^d} + \frac{[J]}{K_{2,J}^d} + \frac{[J][S]}{K_{2,S}^d K_{2,J}^d} + \underbrace{\frac{[J][S]}{K_{J,S}^d K_{2,J}^d} + \frac{[J][S]}{K_{J,S}^d K_{2,S}^d} + \frac{[J][S]\rho}{K_{2,S}^d K_{2,J}^d K_{J,S}^d}}_{=0 \text{ if no affinity between } S \text{ and } J}\right) \\
&= [H_1]K_eF_2(J,S)
\end{aligned} \tag{4.10}$$

where we are considering here F_1 and F_2 as functions of the free concentrations of substrates and JDPs to ease the description [1].

Using the definition of F_1 and F_2, the hydrolysis rate (4.1) can be written as:

$$k_{hy} = K_e\frac{F_2(J,S)}{F_1(J,S) + K_eF_2(J,S)}k^{max} \simeq K_e\frac{F_2(J,S)}{F_1(J,S)}k^{max} + O(K_e^2) \tag{4.11}$$

where we considered $K_e \ll 1$.

Without any substrate or JDP, we find $F_i(J=0,S=0) = 1$ for $i = 1,2$ therefore the basal hydrolysis rate is simply given by:

$$k_{hy}^0 \simeq K_e k^{max} \tag{4.12}$$

Once we have the basal rate, we can compute, in each concentration condition, the acceler-

[1] The actual calculation were however performed while considering the total concentrations of substrates and DnaJ, but those solutions will not be presented here to keep a light description. The conclusions we present here are however identical to the ones resulting from the more complex calculations.

ation of the hydrolysis rate. This approach allows us to eliminate the dependence in k^{max}, reducing the number of parameters of the system.

With only substrates, $F_i(J=0,S) = 1 + \dfrac{[S]}{K^d_{i,S}}$, $i = 1,2$, the acceleration induced by the substrate is given by:

$$a_S \simeq \frac{K_e \frac{1+\frac{[S]}{K^d_{2,S}}}{1+\frac{[S]}{K^d_{1,S}}} k^{max}}{K_e k^{max}} = \frac{1+\frac{[S]}{K^d_{2,S}}}{1+\frac{[S]}{K^d_{1,S}}} \tag{4.13}$$

In order to have an effective acceleration, $a_S > 1$, a condition on $K^d_{2,S}$ imposes that $K^d_{1,S} > K^d_{2,S}$. This implies, as expected, that the H_2 states, characterized by a fast hydrolysis rate, have a higher affinity for the substrate.

Similarly, if only J is present $F_i(J,S=0) = 1 + \dfrac{[J]}{K^d_{i,J}}$ and the acceleration is given by:

$$a_J \simeq \frac{1+\frac{[J]}{K^d_{2,J}}}{1+\frac{[J]}{K^d_{1,J}}}. \tag{4.14}$$

with the same condition $K^d_{1,J} > K^d_{2,J}$ as above.

The maximal acceleration is obtained at *saturating* concentration of S or J ($S \gg K^d_{1,S}$ and $J \gg K^d_{1,J}$) and is determined by the ratio of the dissociation constants between the active and inactive state:

$$a^{max}_S = \lim_{S \gg K^d_{1,S}, K^d_{2,S}} a_S = \frac{K^d_{1,S}}{K^d_{2,S}} =: \lambda_S, \qquad a^{max}_J = \lim_{J \gg K^d_{1,J}, K^d_{2,J}} a_J = \frac{K^d_{1,J}}{K^d_{2,J}} =: \lambda_J. \tag{4.15}$$

In the presence of both J and S, but without affinity between the two (case without synergy), the functions F_1 and F_2 are given by $F_i(J,S) =)(1+\dfrac{[J]}{K^d_{i,J}})(1+\dfrac{[S]}{K^d_{i,S}})$, $i = 1,2$. The acceleration is then given by:

$$a_{JS} \simeq \frac{(1+\frac{[J]}{K^d_{2,J}})(1+\frac{[S]}{K^d_{2,S}})}{(1+\frac{[J]}{K^d_{1,J}})(1+\frac{[S]}{K^d_{1,S}})} = a_J a_S \tag{4.16}$$

where the two components of the acceleration are completely independent of each other.

At large concentration of S and J, both component of the acceleration, a_J and a_S, will reach their maximal value, allowing the overall acceleration a_{JS} to reach:

$$a^{max}_{JS} = a^{max}_J a^{max}_S = \frac{K^d_{1,S} K^d_{1,J}}{K^d_{2,S} K^d_{2,J}} = \lambda_J \lambda_S \tag{4.17}$$

In the case with synergy, i.e. with affinity between S and J, the functions F_1 and F_2 do not simplify, leading to a really complicated expression for the hydrolysis rate. However, we can assume that with the appropriate concentrations of S and J, the formation of the fully connected trimers $\boxed{[H_1JS]}$ and $\boxed{[H_2JS]}$ will exceed any other dimers or trimers formation, due to the high local concentration ρ. Therefore, we can approximate F_1 and F_2 such that: $F_i(J,S) \simeq 1 + \frac{[J][S]\rho}{K^d_{i,S}K^d_{i,J}K^d_{j,S}}$, $i = 1,2$.

Then, the acceleration of the hydrolysis rate becomes

$$a_{\underline{JS}} \simeq \frac{1 + \frac{[J][S]\rho}{K^d_{2,S}K^d_{2,J}K^d_{J,S}}}{1 + \frac{[J][S]\rho}{K^d_{1,S}K^d_{1,J}K^d_{J,S}}} \tag{4.18}$$

To reach its maximal value, the condition $S \gg K^d_{1,S}, K^d_{2,S}$ an $J \gg K^d_{1,J}, K^d_{2,J}$ is no longer necessary. Instead, the condition is now $[J][S]\rho \gg K^d_{1,S}K^d_{1,J}K^d_{J,S}$ where the high value of ρ can compensate for concentrations of J and S lower than saturation concentrations.

$$a^{max}_{\underline{JS}} = \frac{K^d_{1,J}K^d_{1,S}}{K^d_{2,J}K^d_{2,S}} = \lambda_J\lambda_S = a^{max}_{JS} \tag{4.19}$$

In summary, the stimulation of ATP hydrolysis by S and J is attributed to their ability to shift Hsp70 into its active state. The maximal acceleration one can obtain with S or J depends on the dissociation constant ratio between binding the inactive and active state such that $a^{max}_J = \frac{K^d_{1,J}}{K^d_{2,J}}$ and $a^{max}_S = \frac{K^d_{1,S}}{K^d_{2,S}}$.

Two cases arise in the presence of both S and J. The two proteins can either not interact with each other such that the ATPase stimulation simply corresponds to their combined effect, or interact with each other and therefore promote the formation of a fully connected trimer. In both cases, the same maximal acceleration is reached, $a^{max}_{\underline{JS}} = a^{max}_{JS} = \frac{K^d_{1,J}K^d_{1,S}}{K^d_{2,J}K^d_{2,S}}$ but under different concentration conditions.

Theoretical results

The idea that the synergy helps with trimer formation can be verified in Figure 4.5 displaying the percentage of Hsp70 in a trimer in function of the concentration of S and J.

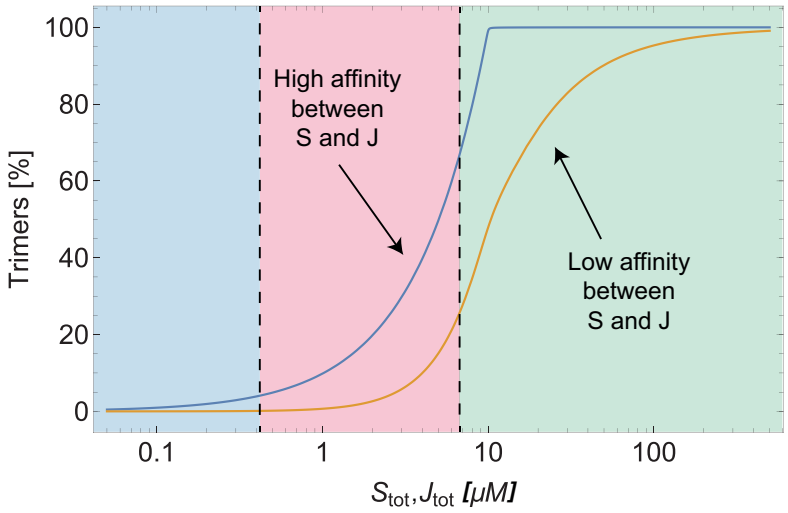

Figure 4.5: Percentage of Hsp70 in a trimer configuration in function of the concentration of substrate and J-domain protein. Blue curve: case with synergy due to the high affinity between J and S, orange curve: case without synergy due to the absence of affinity between J and S. Blue area: low concentration of S and J, nearly no trimers. Red area: intermediate concentration of S and J leading to an increase and saturation of the trimer formation in the case of affinity between S and J due to the effect of the co-localization. Green area: concentration of J and S high enough to create trimers even without affinity. See Table 4.1 for the used parameters

At really low concentrations, no trimers, or very few, are formed as shown in the blue area on Figure 4.5. As the concentrations increase, trimers start to appear in the case **with synergy** thanks to the local high concentration of proteins (red area on Figure 4.5). Finally, for even higher concentrations, trimers in the case **without synergy** are formed and can even saturate Hsp70 but only for extremely high concentrations of S and J.

K_e [−]	$K^d_{1,S}$ [μM]	$K^d_{2,S}$ [μM]	$K^d_{1,J}$ [μM]	$K^d_{2,J}$ [μM]	$K^d_{J,S}$ [μM]	ρ [μM]
10^{-6}	4.4 [64; 73]	0.47 [74]	0.07 [74]	0.0014	0.02 [74]	1000

Table 4.1: Parameters selected for the theoretical plots. The affinity between H_2 and S was assumed to be similar that the affinity between S and Hsp70·ADP. Total concentration of Hsp70 was fixed to 10μM.

As a high proportion of trimers is directly linked to a high hydrolysis rate, we can conclude that the synergistic effect of J and S on the hydrolysis rate is efficient at boosting the ATPase activity for relatively low concentrations.

Indeed, as observed in Figure 4.6A and B, in high concentrations of J and S, both configurations are reaching a maximum hydrolysis rate of ~ 460 times the basal hydrolysis rate, in agreement with [75].

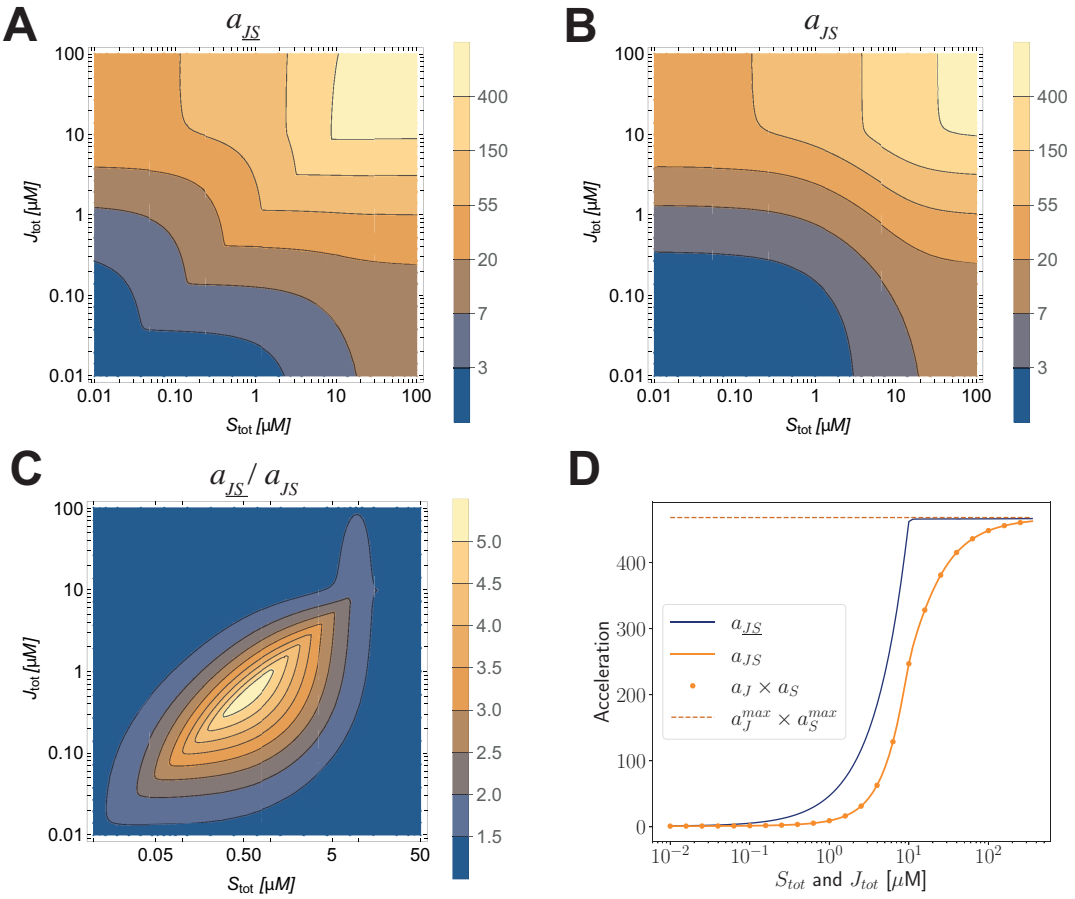

Figure 4.6: **A**: Acceleration with synergy for varying concentrations of substrates and DnaJs. **B**: Acceleration without synergy for varying concentrations of substrates and DnaJs. **C**: Ratio between acceleration with synergy and without synergy in function of the concentration of substrate and DnaJ. **D**: Acceleration of the ATP hydrolysis with and without synergy. Blue curve corresponds to the acceleration with synergy, orange curve is the acceleration of the hydrolysis rate without synergy, orange dot: combination of the individual acceleration due to the substrates and DnaJs, brown horizontal line: theoretical limits of the acceleration. Parameters used given in Table 4.1

While the difference between the acceleration with and without synergy is minimal at low or high concentrations ($a_{\underline{JS}}/a_{JS} \simeq 1$ on panel C), the synergistic effect is clearly visible at intermediate concentrations, between 0.1 and 1 μM, which are in the range of physiological concentration [33; 76]. The synergistic effect is maximal for concentrations around 0.5 μM for both S and J and the ratio $a_{\underline{JS}}/a_{JS} > 5$.

Therefore, we can confirm that the principal effect of the synergy is not to reach a faster hydrolysis rate but to compensate for the low concentration of substrates and/or JDPs in order to reach the same fast hydrolysis rate at lower concentrations.

The results also confirm that the acceleration induced by substrate and DnaJ without synergy is simply the combination of their individual contributions, as demonstrated by the perfect alignment between the a_{JS} and $a_J \times a_S$ curves in Figure 4.6D.

4.2.3 Applications

A special case: the HscA/HscB/IscU system

This work was done in collaboration with Jarosław Marszałek and Aneta Grabińska at the University of Gdansk and Medical University of Gdansk. All the experimental work was performed by Aneta Grabińska.

Steady-state experiments of HscA (from the Hsp70 family) ATPase activity were performed in the presence of its co-chaperon HscB, a DnaJ-like protein, and IscU as substrate.

The HscA/HscB/IscU system was chosen due to HscA fast nucleotide exchange rate after the hydrolysis [77; 78] without the need of a nucleotide exchange factor (NEF). This allowed the use of the pre-equilibrium assumption on which our model is based, even in steady-state experiments.

It is important to know that HscA, HscB and IscU are part of a very specialized Hsp70 system and only interact with each other [79], in opposition to many Hsp70s which are known for their versatility due to their interactions with multiple different DnaJs and substrates.

Two different HscB were considered, the wild type (HscB WT) which can bind to IscU, allowing the formation of a fully connected trimer and a mutant (HscB LMF) modified to impair its binding with IscU. Similarly, IscU MVY, a mutant form of IscU with a reduced affinity for HscB was used in the experiments, in addition to IscU WT.

The experiments were performed multiple times and the measured hydrolysis rates were averaged in-between repetitions. First, to assess the individual impact of IscU and HscB, titrations of both proteins alone were performed separately (Figure 4.7A). The synergistic effect was then observed in titrations of HscB with a fixed initial concentration of IscU (5 μM, 10μM and 20μM), Figure 4.7B. Finally, using mutants impairing the affinity between HscB and IscU, the decrease in the synergistic effect was shown with the titration of HscB LMF with an initial 20 μM concentration of IscU WT and similarly with the titration of HscB WT with 20 μM IscU MVY (Figure 4.7C).

The data presented in Figure 4.7 corresponds, for each experimental condition, to the average acceleration of the hydrolysis rate compared to the first measurement of the titration, at zero concentration of titrant. In this way, we eliminate the dependence on the intrinsic hydrolysis rate of the system. For the titrations of HscB with a fixed concentration of IscU, the initial hydrolysis rate corresponds to the hydrolysis rate induced by IscU, not the basal hydrolysis rate. Therefore the data actually correspond to a_{JS}/a_S and $a_{\underline{JS}}/a_S$.

Both IscU and HscB alone are able to increase the ATPase activity, although the acceleration is limited to about 7-fold for HscB and 3-fold for IscU (Figure 4.7 panel A). The addition of HscB to a solution with a fixed concentration of IscU (5, 10 or 20 M), leads to an acceleration of about 20/25-fold compared to the ATPase activity due to IscU alone (Figure 3B), thus greater than the intrinsic acceleration due to HscB alone. Finally, the presence of a mutant altering the affinity between IscU or HscB (either HscB LMF or IscU MVY) significantly reduces the acceleration compared to the case with the two wild types HscB and IscU. Indeed, an acceleration of about 7-fold was reached with IscU MVY and HscB WT and of about 15-fold for IscU WT and HscB LMF (Figure 4.7C).

Our theoretical model was tested to reproduce those experimental results. Let $f(\{\beta\}, c)$ be the function computing the theoretical acceleration of the hydrolysis rate, a_{th}, in function of a set of parameters intrinsic to the model $\{\beta\}$, and c the substrate or DnaJ concentration, such that

$$f(\{\beta\}, c) = a_{th} \tag{4.20}$$

For each experimental titration $\mathbf{d} = (d_1, d_2, ...)$, we can find the optimal set of parameters $\{\beta_{opt}\}$ that minimizes the error

$$SSE = \sum_i (f(\{\beta\}, c_i) - d_i)^2 \tag{4.21}$$

between the data points d_i and the theoretical acceleration given by the function $f(\{\beta\}, c_i)$ where c_i is the titrant concentration. As many parameters of $\{\beta\}$ are the same in-between different titration experiments, all experiments were fitted simultaneously to get a unique set $\{\beta\}_{opt}$. The actual parameters in $\{\beta\}$ are K_e, the equilibrium constant between the two allosteric states H_1 and H_2, $K^d_{1,S}$, $K^d_{2,S}$, $K^d_{1,J}$ and $K^d_{2,J}$, the substrate and JDP dissociation constants with Hsp70 in states 1 and 2, $K^d_{J,S}$, the dissociation constant of the substrate with JDP and ρ, the apparent concentration of the unbound protein during the trimer formation. Finally, to account for the use of mutants, two other dissociation constants were added $K^d_{J_{LMF},S}$ and $K^d_{J,S_{MVY}}$, which corresponds to the dissociation constants between HscB LMF and IscU WT, and HscB WT and Iscu MVY respectively.

To avoid falling into a local minimum when minimizing SSE, the minimization was repeated multiple times with, each time, a randomly chosen initial guess. Due to the large number of parameters we are trying to estimate, the surface landscape of SSE is difficult to explore thoroughly in one occasion. Hence, the large disparity in the parameters estimated values between each repetition, as represented in the appendices (Figure A.2).

The data and best fit, on Figure 4.7, were obtained with the set of parameters presented in Table 4.2.

K_e [−]	$K^d_{1,S}$ [μM]	$K^d_{2,S}$ [μM]	$K^d_{1,J}$ [μM]	$K^d_{2,J}$ [μM]	$K^d_{J,S}$ [μM]	$K^d_{J_{LMF},S}$ [μM]	$K^d_{J,S_{MVY}}$ [μM]	ρ [μM]
0.014	408.6	65.7	26.6	2.9	10	754.2	12663.0	8100.1

Table 4.2: Parameters obtained from fitting with the least-square methods on the experimental data . The value of K_e was fixed to 0.014, corresponding to the ratio between the basal hydrolysis rate and the highest measure of the hydrolysis rate, and $K^d_{j,S}$ was constrained to [0.1, 10] μM in agreement with experimental observations [80].

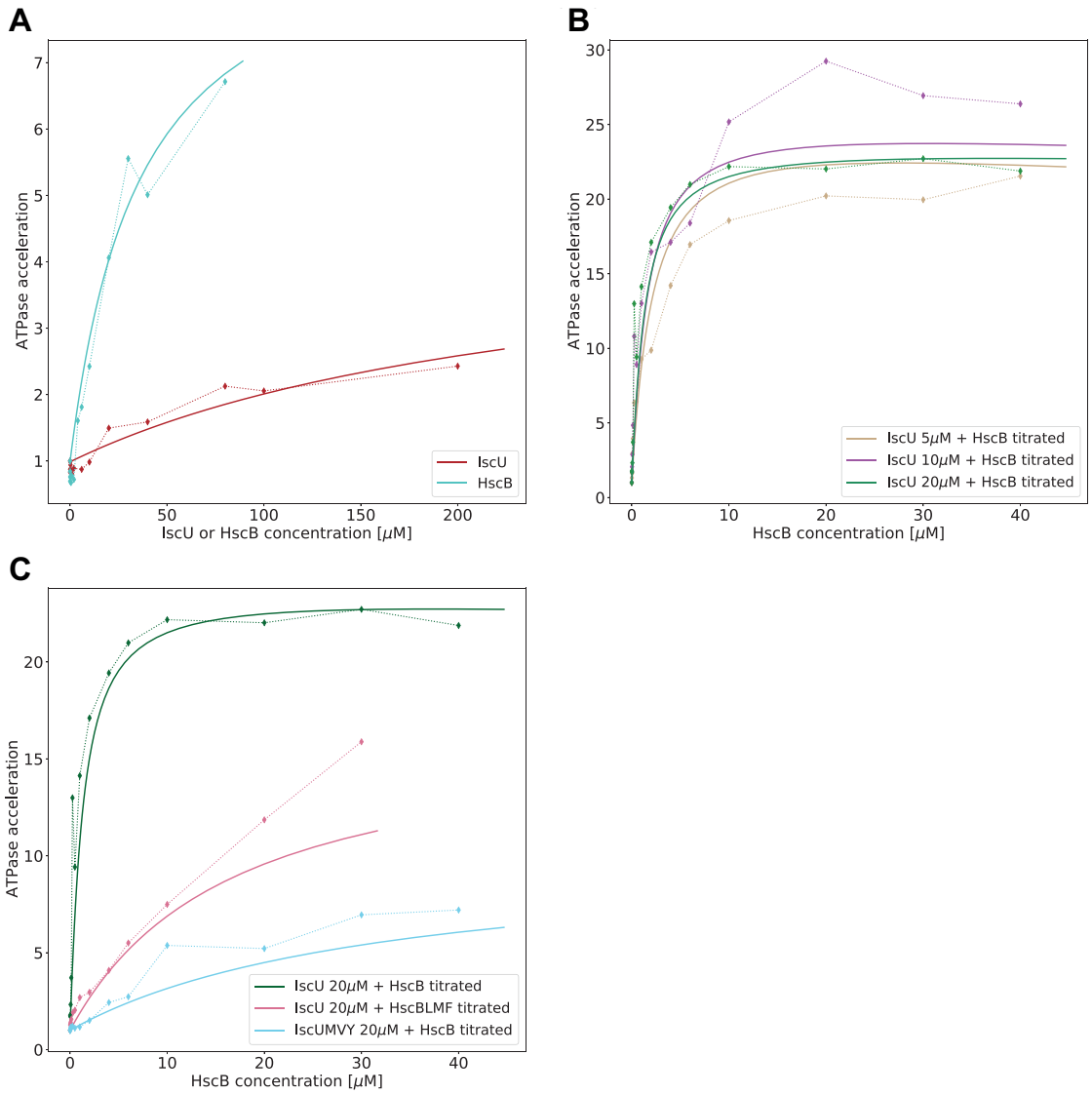

Figure 4.7: Measure of the ATP hydrolysis acceleration induced by the addition of IscU and/or HscB. Data are presented by dots and the model simulation by the solide line. **A**: Titration of IscU (red) or HscB (blue) alone. **B**: Titration of HscB WT with fixed concentrations of IscU WT (5 μM in beige, 10 μM in purple and 20 μM in green). **C**: Titration of HscB with 20 μM IscU with mutants. In green HscB WT with IscU WT, in pink HscB LMF with IscU WT and in blue HscB WT with IscU MYV.

The individual effect of HscB (J) and IscU (S) on the ATPase activity is well reproduced by our model, as seen in Figure 4.7A. The synergistic effect of IscU (fixed 20 μM) and HscB is also fully replicated (green line in Figure 4.7B and C). However, for lower concentration of IscU, the fit is not as able to follow the data. Indeed, while the three predictions are extremely close to each other, the three curves differ highly when we look at the data.

Interestingly, the presence of both HscB and IscU does not result in a significant increase in HscA ATPase activity. Upon evaluating the acceleration of the hydrolysis rate in comparison to the basal rate, even at saturation, the highest value obtained is only about 30-fold, which is considerably lower compared to other Hsp70 systems where an acceleration of up to 300-fold can be easily achieved [29]. Finally, our model successfully reproduces the impact of the low affinity between the pairs HscB LMF and IscU WT and HscB WT and IscU MYV. Indeed, the only difference between the three fitted curves of Figure 4.7C is the value of dissociation constant between HscB and IscU.

In Table 4.2, the fitted parameters corresponded to those that provided the best fit to the model. It is worth noting, however, that numerous other parameter sets also resulted in a good fit characterized by $SSE < 1$. Among those different sets of parameters, all parameters exhibit a high variability as shown in Figure 4.8B. This is especially true for $K_{J,S_{MVY}}$ whose distribution spans over an order of magnitude. The value of $K^d_{J,S}$ was artificially limited to 10 μM, due to biological constraints, explaining the difference in its distribution compared to the other parameters.

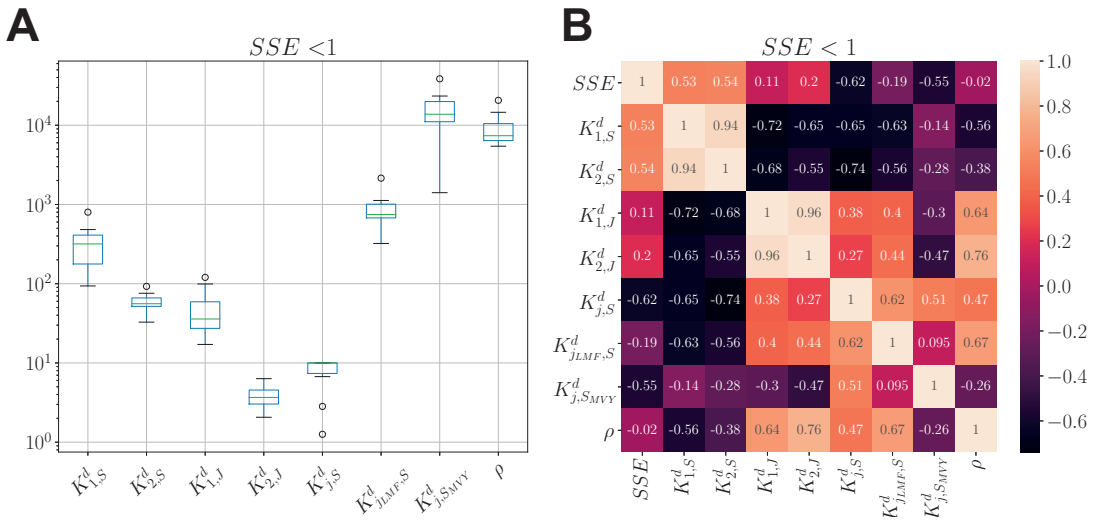

Figure 4.8: Information on the value of our model fitted parameters. Only parameters inducing a good enough fit were considered ($SSE < 1$) corresponding to 12 sets of parameter values. **A**: Distribution of the parameters **B**: Correlation matrix between each parameter. K_e was dropped as its value was fixed.

Although we are not able to define a precise value for each parameter, the very strong correlation between some of them highlights the importance of their relative values to one another

rather than their absolute values. This is, for example, the case with the dissociation constants of Hsp70 in its active and inactive state for a given protein ($K^d_{1,S}$ with $K^d_{2,S}$ and $K^d_{1,J}$ with $K^d_{2,J}$). In both cases, the correlation is nearly perfect with a correlation coefficient of 0.94 for $K^d_{1,S}$ and $K^d_{2,S}$ and of 0.96 for $K^d_{1,J}$ and $K^d_{2,J}$. This strong correlation is due to a pretty stable estimation of the ratio between the constants. Indeed, $\lambda_S := \frac{K^d_{1,S}}{K^d_{2,S}}$ and $\lambda_J = \frac{K^d_{1,J}}{K^d_{2,J}}$ characterize the maximal acceleration the system can reach during the titration of substrates, JDPs respectively and have a relatively low coefficient of variation (CV) compared to other parameters (Table 4.3).

The most consistent parameters along the 12 fits, $\lambda_S\lambda_J$ with CV=0.12, is directly linked to the maximal acceleration the system can reach. This consistency is due to the relatively strong anti-correlation between $K^d_{1,S}, K^d_{2,S}$ pair and the $K^d_{1,J}, K^d_{2,J}$ pair, between -0.55 and -0.72 in Figure 4.8B.

	$K^d_{1,S}$ [μM]	$K^d_{2,S}$ [μM]	$K^d_{1,J}$ [μM]	$K^d_{2,J}$ [μM]	$K^d_{J,S}$ [μM]	$K^d_{J_{LMF},S}$ [μM]	$K^d_{J,S_{MVY}}$ [μM]	ρ [μM]	λ_S [-]	λ_J [-]	$\lambda_J\lambda_S$ [-]
Mean	328.2	58.8	49.5	3.9	8.0	875.9	15720.9	9433.6	5.2	11.6	55.5
SD	194.2	15.7	31.8	1.3	3.1	471.0	9763.9	4614.83	1.8	3.4	6.4
CV	0.59	0.26	0.64	0.34	0.38	0.54	0.62	0.49	0.35	0.29	0.12

Table 4.3: Mean, standard deviation (SD) and coefficient of variation (CV) for the parameters of the model on the 12 best iterations with $\lambda_S = \frac{K^d_{1,S}}{K^d_{2,S}}$ and $\lambda_J = \frac{K^d_{1,J}}{K^d_{2,J}}$.

4.2.4 4 allosteric configurations model

Recent publications by Hendrickson *et al.* [69; 71; 81] have investigated the mechanism of Hsp70 ATPase activity stimulated by substrate binding. The authors have proposed a theoretical model [71] based on the presence of two different allosteric configurations of Hsp70·ATP characterized by their substrate affinity and their ATPase activity.

A *restraining state* with low substrate affinity and low ATPase activity is in equilibrium with a *stimulating state* with high substrate affinity and high ATPase activity. This equilibrium defines the observed hydrolysis rate and the binding of a substrate effectively shifts the equilibrium toward the *stimulating states* due to its higher affinity.

Furthermore, the authors claimed to have identified those two states structures [81]. While the *restraining state* corresponds to the usual Hsp70·ATP structure with an open SBD and the linker docked into the NBD as in Figure 4.1B, the *stimulating state* still display a docked linker but with a closed SBD, represented in Figure 4.9B. This state can be associated to an intermediate structure between the ATP and the ADP configurations.

Finally, the authors compared the effect of the substrate concentration on the ATPase activity (data on Figure 4.10) on WT DnaK and on diverse mutants, including I483D, blocked into the *stimulating state* [69]. While WT DnaK ATPase activity increases as expected with the substrate concentration, I483D is not affected by the addition of substrate and hydrolyses ATP

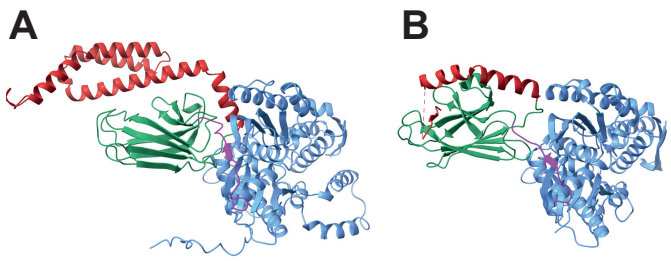

Figure 4.9: **A**: HscA AlphaFold prediction [82; 83], **B**: Hendrickson S-state cristal (PBD ID code: 7KRU) [69]

at a constant higher rate than WT.

The idea of Hsp70·ATP possessing different allosteric structures is consistent with a prior study [62], which proposed the existence of a SBD-undocked configuration with a bound linker as an allosterically active state Hsp70·ATP, in opposition to the classic Hsp70·ATP state with a docked SBD and a bound linker. This finding is similar to the structure proposed by Hendrickson. Additionally, the predicted structure of HscA, a specific Hsp70, as determined by AlphaFold and illustrated in Figure 4.9A, displays similarities, featuring a closed SBD and a bound linker. All those observations can be seen as evidence of the flexibility of the Hsp70·ATP structure and of the possible presence of multiple allosteric structures.

By combining the observed allosteric configurations with the theoretical aspect of the hydrolysis stimulation, Hendrickson *et al.* are pushing forward toward a better understanding of the role of these allosteric structures. Their approach is therefore extremely insightful for our own research.

However, their model is limited to substrate-induced ATPase activity and assumes that, at saturation, all Hsp70 are in the stimulating state. This assumption completely ignores the possible impact the J-domain protein could have on such a system. We now know that DnaJ binds Hsp70 with contact points on the NBD, the linker, and the SBD and is thought to stimulate hydrolysis rate by increasing Hsp70 sensitivity for the substrate's signal [70]. But it also stimulates the hydrolysis rate by itself. If we assume, in agreement with Hendrickson, that the ATPase activity is stimulated by a change in configuration, then DnaJ should also trigger structural changes in Hsp70.

Furthermore, as reported in [69], other DanK mutants with a mutation in the NBD, I160D and N170D, displayed a high hydrolysis rate in their apo-state, characterized by the absence of peptide. Moreover, I160D ATPase activity was further stimulated in the presence of peptides, as depicted in Figure 4.10. This observation is very similar to what would be anticipated in the case of a peptide titration with fixed initial concentrations of DnaK WT and DnaJ.
This leads us to assume the I160D mutant is in a structure close to or equivalent to the one

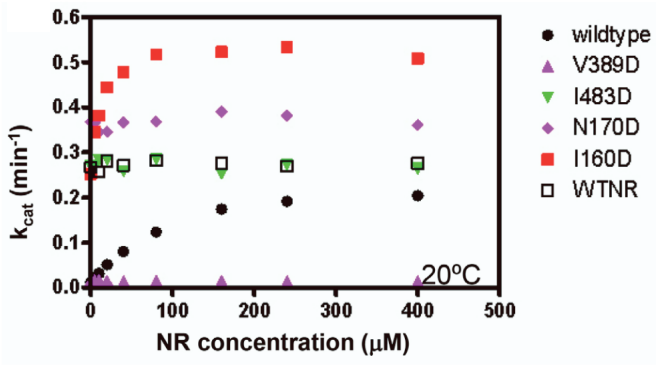

Figure 4.10: ATPase activity of various Hsp70 with different mutations in function of peptide concentration. Figure from [69].

induced by J-domain binding. Therefore, the high hydrolysis rate could be due to the mutation mimicking the action of DnaJ.

If this is true, we can assume that in addition to the new allosteric structure, named here $\text{Hsp70}_{\text{SBD}}$ and proposed in [81], there is an additional allosteric configuration triggered by JDP binding. Due to the location of the mutations, this additional configuration is called $\text{Hsp70}_{\text{NBD}}$, although we have no evidence of the actual location of the configuration change.

Assuming both these changes to be independent of one another, a fourth allosteric structure, $\text{Hsp70}_{\text{SBD·NBD}}$, characterized by configurational changes in both the SBD and the NBD should exist. In that case, we can naturally extend our previous model with one inactive (H_1) and one active state (H_2) to a 4-state model composed of three non-active states (H, H_{SBD}, H_{NBD}) which were coarse-grained in a unique state H_1 and one active state ($H_{\text{SBD·NBD}}$) equivalent to our state H_2.

Mathematical details

Similarly to our previous model, Hsp70 fluctuates between these four different states, as illustrated in Figure 4.11, and the equilibrium between each state is influenced by the presence of substrate and JDP. However, as JDPs are not part of the experiment, mathematical development involving them will be avoided to lighten the description.

The equilibrium between the different allosteric states now depends on two different constants, namely $K_{e,S}$ and $K_{e,J}$, where $K_{e,S}$ is associated with the allosteric change on the substrate binding domain (SBD) induced by the substrate and $K_{e,J}$ is associated to the configurational modification on the NBD and triggered by JDP binding.

Therefore the equilibrium concentrations are given by:

$$[H_{\text{SBD}}] = K_{e,S}[H] \qquad [H_{\text{NBD}}] = K_{e,J}[H]. \tag{4.22}$$

Assuming the two allosteric changes to be independent of each other, the equilibrium concentration for the active state is:

$$[H_{\text{SBD·NBD}}] = K_{e,J}[H_{\text{SBD}}] = K_{e,S}[H_{\text{NBD}}] = K_{e,J}K_{e,S}[H]. \tag{4.23}$$

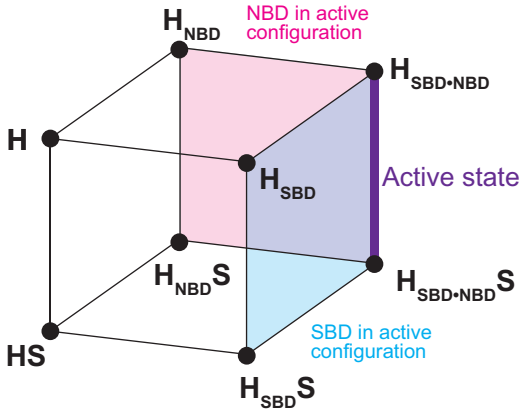

Figure 4.11: Graphical representation of the 4-state Hsp70 model with binding to substrates. Other possible binding states (with JDPs or trimers configurations) are not represented to preserve the reader's sanity.

As previously, we assume that only the active state is able to hydrolyze the ATP at a very fast rate called k^{max}. In that case, the observed hydrolysis rate simply depends on the ratio of Hsp70 in the active state $H_{\text{SBD·NBD}}$ compared to the total concentration of Hsp70. Therefore, the basal hydrolysis rate is simply:

$$k^0_{hy} = \frac{[H_{\text{SBD·NBD}}]}{[H]_{tot}} k^{max} = \frac{K_{e,J}K_{e,S}}{(1+K_{e,J})(1+K_{e,S})} k^{max} \tag{4.24}$$

Adding substrates modifies the Hsp70 equilibrium toward the allosteric states with the configurational change in the SBD. To be able to do that, those states, namely H_{SBD} and $H_{\text{SBD·NBD}}$ should have a higher affinity for the substrate that the states with the SBD in the normal configuration.

$$[HS] = \frac{[H][S]}{K^d_S} \qquad [H_{\text{NBD}}S] = \frac{[H_{\text{NBD}}][S]}{K^d_S} \tag{4.25}$$

$$[H_{\text{SBD}}S] = \frac{[H_{\text{SBD}}][S]}{K'^d_S} \qquad [H_{\text{SBD·NBD}}S] = \frac{[H_{\text{SBD·NBD}}][S]}{K'^d_S} \tag{4.26}$$

where K^d_S is the dissociation constant of S with H and H_{NBD} and K'^d_S is the dissociation constant of S with H_{SBD} and $H_{\text{SBD·NBD}}$. As before, for the substrate to stimulate ATPase activity, $K'^d_S < K^d_S$.

The hydrolysis rate induced by binding to the substrate is therefore given by:

$$k_{hy}^{S} = \frac{[H_{\text{SBD}}] + [H_{\text{SBD}}S]}{[H]_{tot}} = \frac{K_{e,J}K_{e,S}K_S^d(K_S'^d + S)}{(1+K_{e,J})(K_S'^d S + K_S^d(K_S'^d + K_{e,S}K_S'^d + K_{e,S}S))}k^{max} \tag{4.27}$$

Let's once again consider the acceleration of the hydrolysis rate.

$$a_S = \frac{k_{hy}^S}{k_{hy}^0} = \frac{(1+K_{e,S})K_S^d(K_S'^d + S)}{K_S'^d S + K_S^d(K_S'^d + K_{e,S}K_S'^d + K_{e,S}S)} \tag{4.28}$$

At saturation, $S \gg K_S^d$, the acceleration reaches its maximal values:

$$a_S^{max} = \lim_{S \gg K_S^d} a_S = \frac{K_S^d(1+K_{e,S})}{K_S'^d + K_S^d K_{e,S}}. \tag{4.29}$$

The 2-state model as a coarse-grained version of the 4-state model

As previously stated, the 2-state model (with H_1 / H_2) is a coarse-grained version of the more detailed model with four states presented above, where H_1 is characterized by the 3 inactive states H, H_{SBD} and H_{NBD} and H_2 is equivalent to the active state $H_{\text{SBD·NBD}}$. This means that both models should be able to reproduce the same phenomenon regarding the ATPase stimulation with adequate parameters.

The equilibrium constant, K_e, is defined as:

$$H_2 = K_e H_1 \Longrightarrow H_{\text{SBD·NBD}} = K_e(H + H_{\text{NBD}} + H_{\text{SBD}}) \iff K_e = \frac{K_{e,S}K_{e,J}}{(1+K_{e,S}+K_{e,J})} \tag{4.30}$$

where the last equivalence comes from (4.22) and (4.23).

Similarly, we can compute the coarse-grained dissociation constant of Hsp70 with substrates in its inactive state H_1 and its active state H_2 from the detailed equilibrium constant of the 4 states model:

$$K_{1,S}^d = \frac{[H_1][S]}{[H_1 S]} = \frac{([H] + [H_{\text{SBD}}] + [H_{\text{NBD}}])[S]}{[HS] + [H_{\text{SBD}}S] + [H_{\text{NBD}}S]} = \frac{(1+K_{e,S}+K_{e,J})}{(\frac{1}{K_S^d} + \frac{K_{e,S}}{K_S'^d} + \frac{K_{e,J}}{K_S^d})} \tag{4.31}$$

$$K_{2,S}^d = \frac{[H_2][S]}{[H_2 S]} = \frac{[H_{\text{SBD·NBD}}][S]}{[H_{\text{SBD·NBD}}S]} = K_S'^d \tag{4.32}$$

Similarly, the Hsp70:DnaJ dissociation constants in the two states model can be computed from the 4 states model ones:

$$K_{1,J}^d = \frac{(1+K_{e,J}+K_{e,S})}{(\frac{1}{K_J^d} + \frac{K_{e,J}}{K_J'^d} + \frac{K_{e,S}}{K_J^d})} \tag{4.33}$$

$$K_{2,J}^d = K_J'^d \tag{4.34}$$

where K_J^d is the dissociation constant of H and H_{SBD} with J and $K_J'^d$, the dissociation constant of H_{NBD} and $H_{\text{SBD·NBD}}$ with J. The other parameters, ρ, and $K_{j,S}^d$, are not affected by the model coarse-graining as they are not directly linked to Hsp70 allosteric states.

Results with the 4-state model

The ATPase activity of DnaK WT and I483D, the mutant stuck in Hendrickson's stimulating state, were measured in single turn-over experiments in [69] in function of the concentration of the NRLLLTG peptide as substrate. While the increase in peptide concentration leads to a higher hydrolysis rate for the DnaK WT, no impact of the peptide is found on I483D which keeps a roughly constant high hydrolysis rate.

While fitting the data with our 4-state model, we also took into account the I483D data. Indeed, according to [71], the mutant SBD is blocked in its active state. In our model, this state is characterized by H_{SBD} and the $H_{\text{SBD·NBD}}$ states (blue area on Figure 4.11). Assuming that I483D can still fluctuate between these two states, the hydrolysis rate is given by:

$$k_{hy}^{\text{mutant}} = \frac{[H_{\text{SBD·NBD}}]}{[H_{\text{SBD}}] + [H_{\text{SBD·NBD}}]} k^{max} \tag{4.35}$$

$$= \frac{K_{e,J}}{1 + K_{e,J}} k^{max}. \tag{4.36}$$

And the acceleration is simply given by:

$$a^{\text{mutant}} = \frac{k_{hy}^{\text{mutant}}}{k_{hy}^0} \frac{K_{e,J} k^{max}}{1 + K_{e,J}} \frac{(1 + K_{e,J})(1 + K_{e,S})}{K_{e,J} K_{e,S} k^{max}} = \frac{1 + K_{e,S}}{K_{e,S}} \tag{4.37}$$

The presence of a substrate, indeed, cannot modify the hydrolysis rate as the hydrolysis rate depends on the equilibrium between the $[H_{\text{SBD}}]$ and the $[H_{\text{SBD·NBD}}]$ states.

$$k_{hy}^{S,\text{mutant}} = \frac{[H_{\text{SBD·NBD}}] + [H_{\text{SBD·NBD}} S]}{[H_{\text{SBD}}] + [H_{\text{SBD}} S] + [H_{\text{SBD·NBD}}] + [H_{\text{SBD·NBD}} S]} k^{max} \tag{4.38}$$

$$= \frac{[H_{\text{SBD}}](K_{e,J} + \frac{K_{e,J} S}{K_S'^d})}{[H_{\text{SBD}}](1 + \frac{S}{K_S'^d}) K_{e,J} + \frac{K_{e,J} S}{K_S'^d})} k^{max} \tag{4.39}$$

$$= \frac{K_{e,J}}{1 + K_{e,J}} k^{max} := k_{hy}^{\text{mutant}}. \tag{4.40}$$

Therefore, we imposed an additional constraint to (4.21) to minimize the error between the data from the I483D mutant and the model prediction.

The optimal parameters $\{\beta_{4s}\}_{opt}$ to reproduce the data in Figure 4.12 were obtained by minimizing the following expression:

$$SSE = \sum_i (f_{4s}(\{\beta_{4s}\}, c_i)) - a_i^{WT})^2 + \sum_i (\frac{1 + K_{e,S}}{K_{e,S}} - a_i^{\text{I483D}})^2 \tag{4.41}$$

where f_{4s} is the function of the 4-state model reproducing the acceleration of the hydrolysis rate, $\{\beta_{4s}\}$ the set of parameters describing the system, c_i the concentration of substrate, a_i^{WT} and a_i^{I483D} are the measured acceleration of the hydrolysis rate for DnaK WT, I483D respectively, at substrate concentration c_i.

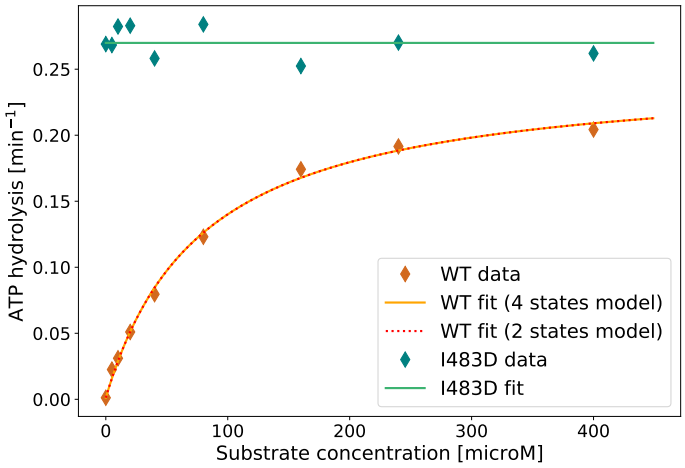

Figure 4.12: Substrates induced ATPase stimulation of DnaK WT and DnaK I483D (data from [69]). The fits were obtained with the 4-state model (solid lines) and the 2-state model (dashed line)

4-state model			2-state model		
$K_{e,S}$ [-]	K_S^d [μM]	$K_S'^d$ [μM]	K_e [-]	$K_{1,S}^d$ [μM]	$K_{2,S}^d$ [μM]
0.004	1064.3	0.37	10^{-5}	78.9	0.37

Table 4.4: Fitted parameters obtained from fitting Hendrickson's data presented on Figure 4.12 with the 4-state model and the 2-state model

The 4-state model easily reproduces both the wild type and the mutant ATPase activity as depicted in Figure 4.12 using the parameters from Table 4.4. The value of $K_{e,S}$ is strongly dictated by the ATPase measurement of the I483D mutant and in agreement with Hendrickson's approach, a very low affinity is found between the peptide and Hsp70 in state H ($K_S^d = 1064.3\mu$M).

The 2-state model gives exactly the same curve reproducing the data for the DnaK WT, but does not contain enough details to correctly represent the state of the mutant. The equivalence between the 4-state and the 2-state models is explicit when looking at the parameters both models used.

As expected from (4.32), both models obtain the same exact value for $K_S'^d$ and $K_{2,S}^d$. Using (4.30) to compute $K_{e,J}$ ($K_{e,J} = 0.0023$), we can also verify if the values of K_S^d and $K_{1,S}^d$ are consistent

with each other. From (4.31):

$$\frac{(1+K_{e,J}+K_{e,S})}{(\frac{1}{K_J^d}+\frac{K_{e,J}}{K_J'^d}+\frac{K_{e,S}}{K_J^d})}=78.87 \tag{4.42}$$

which indeed corresponds to the fitted value of $K_{1,S}^d$ found by fitting the data with the 2-state model.

4.2.5 Link between acceleration and selection

Over two decades ago, a hypothesis was put forth that DnaJ enables Hsp70 to distinguish between short peptides and long proteins [28], due to short peptides inability to interact synergistically with DnaJ in promoting Hsp70 ATPase activity.

We can now explain that a short peptide may not be able to bind simultaneously to DnaJ and Hsp70, due to a lack of available binding sites, which prevents the formation of the fully connect trimer and therefore limits the hydrolysis rate. On the other hand, a long protein with multiple binding sites will allow the formation of the fully connected trimer and its collaboration with DnaJ will maximally stimulate Hsp70 ATPase activity.

This mechanism is equivalent to the catalytic discrimination introduced in Chapter 3, where correct substrates are preferentially accelerated into the selection pathway over incorrect ones. Furthermore, the substrate selection also relies on its binding affinity with DnaJ and Hsp70. In the end, the collaboration between Hsp70 and DnaJ can be seen as a discriminatory mechanism, where the substrate needs to be accepted by the two proteins to be processed. This mechanism is part of DnaJ Hsp70 recruitment mechanism [84] and is at the base of the multi-functionality of the Hsp70 system.

In this context, the HscA/HscB/IscU system is of particular interest as a unique example of an Hsp70/DnaJ/substrate system with limited versatility. Indeed, these proteins interact exclusively with each other, implying that this system may not require a strong discriminatory mechanism.

This is suggested by the relatively low HscA ATPase stimulation obtained in presence of both HscB and IscU (Figure 4.7B), which is only about 30-fold compared to the basal hydrolysis rate. This relatively low level of stimulation stands in contrast to other Hsp70 systems which have demonstrated acceleration rates of up to 300 [29] to 1000-fold [28].

Additionally, as observed in Figure 4.9, the predicted structure of HscA is in a conformation that is highly similar to the Hendrickson stimulating state, in contrast to the conventional ATP-state relaxed state seen in the standard DnaK prediction (Figure A.3).

This suggests that HscA natural equilibrium is titled toward the H_{SBD} state. Binding to the substrate IscU has therefore little to no effect on its ATPase activity, as observed in Figure 4.7A, and the overall maneuvering space to improve the hydrolysis rate with both HscB and IscU is limited.

In the extreme scenario where the substrate-binding domain (SBD) of HscA is always in its

active state, the maximal hydrolysis rate is limited to the impact of DnaJ. To demonstrate this, we can compute the basal hydrolysis rate and the maximal hydrolysis rate in this particular scenario:

$$k_{hy}^{0} = \frac{[H_{\text{SBD·NBD}}]}{[H_{\text{SBD}}] + [H_{\text{SBD·NBD}}]} k^{max} = \frac{K_{e,J}}{1 + K_{e,J}} k^{max} \tag{4.43}$$

$$\lim_{S,J \gg K_S^d K_J^d} k_{hy}^{JS} = \frac{\frac{[H_{\text{SBD·NBD}}][J][S]\rho}{K_S^{\prime d} K_J^{\prime d} K_{J,S}^d}}{\frac{[H_{\text{SBD}}][J][S]\rho}{K_S^{\prime d} K_J^d K_{J,S}^d} + \frac{[H_{\text{SBD·NBD}}][J][S]\rho}{K_S^{\prime d} K_J^{\prime d} K_{J,S}^d}} k^{max} = \frac{K_{e,J} K_J^d}{K_J^{\prime d} + K_{e,J} K_J^d} k^{max} \tag{4.44}$$

where we consider that the maximal hydrolysis rate is reached when all Hsp70 are in the fully connected trimer configuration.

The maximal acceleration is therefore given by:

$$a_{\underline{JS}}^{max} = \frac{K_J^d (1 + K_{e,J})}{K_J^{\prime d} + K_J^d K_{e,J}} \simeq \frac{K_J^d}{K_J^{\prime d}} \tag{4.45}$$

where we suppose $K_{e,J} \ll 1$.

In this hypothetical scenario, the maximal acceleration of the hydrolysis rate is limited to the influence of HscB. As a result, HscB would be the only component capable of affecting HscA ATPase activity, while IscU sole function would be to facilitate the formation of the fully connected trimer in non-saturating concentrations conditions.

Although HscA is not as extreme as this scenario, the low hydrolysis rate observed with the substrate alone, the similarity between the IscU/HscB titration curves in Figure 4.7B and the modest value of the maximal acceleration suggests that HscA SBD may be naturally tilted towards its active state.

4.2.6 Summary

Hsp70 ATPase activity is a key element in driving the Hsp70 cycle to ensure a proper response to protein aggregation. It is regulated by the presence of substrates and/or DnaJ which trigger the ATP hydrolysis when bound to Hsp70.

In our model, based on the presence of allosterically active and inactive states, the observed hydrolysis rate is directly linked to the proportion of Hsp70 in the active state, able to hydrolyze ATP. In this framework, the increase in the hydrolysis rate upon binding to a substrate or DnaJ is explained by their ability to shift the equilibrium in between the allosteric states, toward the active state. With this hypothesis, corroborated by recent experimental observations highlighting the presence of different allosteric structures of Hsp70·ATP, we perfectly reproduced experimental ATPase measurement from titration of substrates and DnaJ individually.

In addition, we have provided insight into the synergistic effect between DnaJ and the substrate. The high hydrolysis rate, observed when DnaJ and the substrate can interact with each other, is explained by the proteins co-localization promoting the formation of the very stable fully

connected trimer formed by Hsp70, DnaJ and a substrate. A decrease in the affinity between DnaJ and substrates, as expected, leads to a decrease in their synergistic effect, lowering the hydrolysis rate. Furthermore, we propose that the synergistic effect's main purpose is not to reach a higher hydrolysis rate but simply to reach that high hydrolysis rate at low physiological concentrations.

Moreover, we show how our model of two allosteric states could be extended to a more complex model with four different allosteric states, assuming DnaJ and the substrate trigger different configurational changes in the Hsp70 structure. This model allows to better adapt to particular cases, such as Hsp70 mutant or HscA.

Finally, we highlight the role of the collaboration between Hsp70 and DnaJ in substrate discrimination. Their combined action favors the processing of longer protein substrates bound to DnaJ over short peptides or proteins in isolation. This highlights the two-signal discrimination mechanism of the Hsp70 system, which requires both the recognition of the substrate by DnaJ and its subsequent binding by Hsp70 for efficient processing.

4.3 DnaJ

The J-domain protein (JDP) family, to whom the DnaJ family belongs, gathers a large number of proteins of very different structures, all containing a highly conserved domain called the J-domain. Due to the large variation in the family, JDPs are usually separated into three categories, class A, B, and C. Class A and B proteins are characterized by their J-domain located at the N-terminal while class C proteins J-domain can be situated anywhere on the protein. Class A and B proteins have in common a glycine/phenylalanine (G/F) rich region following immediately after the J-domain and class A proteins also sports a Cysteine-rich region (CRR) in their C-terminal domain. The precise differences between class A and class B JDPs and their unique specificity are not of prime importance for the understanding of the following work and, therefore, will not be developed here. It is sufficient to know that there are some slight structural and functional differences between class A and B JDPs. Class C proteins present a large variety in sequences [85], and their complex structure and function is outside of the scope of this study.

The J-domain's function is to interact with Hsp70 and promote its ATPase activity, while the rest of the protein allows Hsp70 to carry a particular function. The large variation of JDPs allows Hsp70 to perform many different tasks [76; 84] under the principle that each combination of specific DnaJ and Hsp70 leads to a specific task [86]. As an illustration, the human body only expresses 11 different types of Hsp70 for 50 different DnaJ [76]. Although there are more different JDPs proteins than Hsp70, the total concentration of JDPs in vivo is usually about 10-fold lower than that of Hsp70s [85].

The difference between JDPs and DnaJ-like proteins comes from their ability to bind substrates. While the main role of all DnaJ-like proteins is to bind to substrates and to bring them to Hsp70, JDPs may recruit Hsp70 in other cases. For example, the JDP involved in translocation, i.e. displacement of proteins across a membrane, acts as a targeting device to attract Hsp70 close to the translocation pore, ready to bind incoming substrates and pull them across the membrane. In this case, the JDP is not directly interacting with the substrate but assisting Hsp70 interaction with the substrate.

In the following sections, we will only consider DnaJ-like proteins, interacting with both Hsp70 and substrates.

4.3.1 Self-binding DnaJ

The model presented in this section is a small part of a larger project carried out by Pierre Goloubinoff, Paolo De Los Rios, Mathieu Rebeaud, Bruno Fauvet and Satyam Tiwari on DnaJ "stop-start" mechanism. Special thanks to Mathieu Rebeaud who performed all the experimental work in this section and introduced me to a nice modeling problem.

DnaJ is commonly recognized as playing a dual role in the Hsp70 cycle. First DnaJ is known to assist Hsp70's unfolding process by bringing misfolded proteins to it. Secondly, the presence

of DnaJ, even without a substrate, stimulates Hsp70 ATPase activity, increasing the rate at which the Hsp70-bound ATP is hydrolyzed.

A recent study [87] challenges that description by showing that a truncated class B JDP, containing only the J-domain and the following domain, the G/F rich region, was not able to stimulate ATP hydrolysis. This observation was explained by structural changes in this truncated JDP protein where the J-domain was found to bind to the G/F-rich region, preventing it to interact with Hsp70. However, the full JDP was able to interact with Hsp70 and promote ATP hydrolysis.

From that observation, we propose the existence of an auto-regulatory mechanism of DnaJ proteins in which the G/F-rich region inhibits the DnaJ protein's ability to interact with Hsp70 by binding to the J-domain.

Mathematical description of the models

In order to model the auto-regulatory mechanism of DnaJ, we assume that the protein can fluctuate between two allosteric states, an auto-inhibited and inactive state, J_-, and an active state, J_+. Whereas DnaJ in the active state, with a free J-domain, is able to bind to Hsp70·ATP, the auto-inhibited state is characterized by the binding of the J-domain to the G/F-rich region, preventing any binding to Hsp70·ATP and therefore any stimulation of the hydrolysis rate.

The simplest model of interaction between DnaJ and Hsp70 is represented in Figure 4.13. Hsp70, in its ATP state, can bind to DnaJ in the J_+ state to form $H_T J_+$. The binding of DnaJ promotes ATP hydrolysis and the configurational changes in Hsp70 leading to the state $H_{D\cdot Pi}$ and immediately ejecting DnaJ. Finally, the phosphate is released (k_m) and the nucleotide exchange (k_x) allows Hsp70 to go back to its initial state. In addition to this cycle, powered by DnaJ, Hsp70 can also hydrolyze ATP by itself with a low hydrolysis rate $k^h < k_J^h$, where k_J^h corresponds to the hydrolysis rate stimulated by DnaJ.

Using a Michaelis-Menten-like model, we are assuming that the interactions between Hsp70·ATP and the JDPs are in rapid equilibrium. This assumption is justified by the rate-limiting aspect of the ADP to ATP exchange. Furthermore, to be as concise as possible, the two allosteric states of Hsp70, H_1 and H_2 presented in Section 4.2 during the investigation of the synergy between substrates and DnaJs, are coarse-grained into one and a unique hydrolysis rate due to the presence of DnaJ is considered.

At steady state, the balance between the active and inactive state of DnaJ is characterized by the equilibrium constant K_j such that

$$[J_+] = K_j [J_-]. \tag{4.46}$$

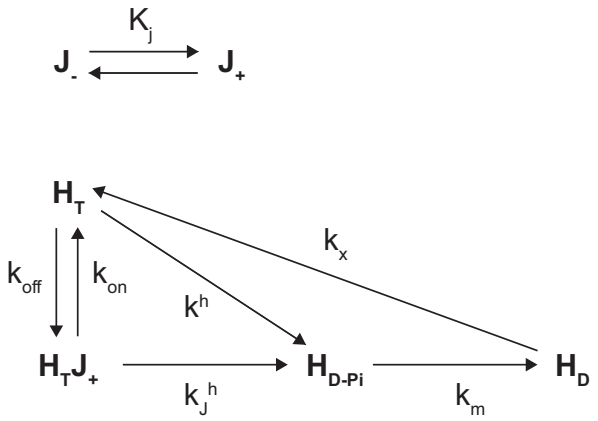

Figure 4.13: Simplified model of the Hsp70 ATP cycle stimulated by DnaJ, with the addition of DnaJ in an auto-inhibited state J_-.

Similarly, the steady-state concentrations of the different states of the model needs to satisfy the following equations:

$$[HJ_+] = \frac{[H][J_+]}{K_M} \qquad [H_{D\cdot P_i}] = \frac{[HJ_+]k_J^h + [H]k^h}{k_m} \qquad [H_D] = \frac{[H_{D\cdot P_i}]k_m}{k_x} \tag{4.47}$$

The total concentrations of DnaJ and Hsp70 are therefore given by

$$[J_{tot}] = [J_-] + [J_+] + [HJ_+] = [J_-]\left(1 + K_j + K_j\frac{[H]}{K_M}\right) \tag{4.48}$$

$$[H_{tot}] = [H] + [HJ_+] + [H_{D\cdot P_i}] + [H_D] \tag{4.49}$$

$$= \frac{[H]}{K_M k_m k_x}\left([J_+](k_m k_x k_J^h(k_m + k_x)) + K_M(k_m k_x k^h(k_m + k_x))\right) \tag{4.50}$$

$$= \frac{[H]}{K_M}\left([J_+](1 + \frac{k_J^h}{\hat{k}}) + K_M(1 + \frac{k^h}{\hat{k}})\right) \tag{4.51}$$

$$= [H]\left(z_J\frac{[J_+]}{K_M} + z\right) \tag{4.52}$$

where $K_M = \dfrac{k_{\text{off}}}{k_{\text{on}} + k_J^h}$ is the Michaelis-Menten constant, $\hat{k} = \dfrac{k_x k_m}{k_x + k_m}$ is the coarse grained exchange rate and $z = 1 + \dfrac{k^h}{\hat{k}}$ and $z_J = 1 + \dfrac{k_J^h}{\hat{k}}$. In the assumption of slow exchange rate, $k_m \gg k_x$, we can assume that $\hat{k} \sim k_x$.

Using $[J_-] = \frac{[J_{tot}]}{1+K_j+K_j\frac{[H]}{K_M}}$, we found the following quadratic equation for $[H]$:

$$[H]^2 + [H](K_M\frac{1+K_j}{K_j} + [J_{tot}]\frac{z_J}{z} - \frac{[H_{tot}]}{z}) - K_M\frac{1+K_j}{K_j}\frac{[H_{tot}]}{z} = 0 \tag{4.53}$$

The flux of hydrolyze ATP is given by

$$f = \frac{[HJ_+]k_J^h + [H]k^h}{H_{tot}} \tag{4.54}$$

Solving (4.53) for $[H]$ and replacing its solution in (4.54), we are finally obtaining

$$\begin{aligned} f = \frac{1}{2H_{tot}}\Bigg(&\frac{H_{tot}}{z}(\frac{z}{z_J}k_J^h + k^h) + (\frac{z}{z_J}k_J^h - k^h)(\frac{z_J}{z}J_{tot} + \frac{1+K_j}{K_j}K_M \\ &-\sqrt{2\frac{1+K_j}{K_j}K_M(\frac{H_{tot}}{z} + \frac{z_J}{z}J_{tot}) + (\frac{H_{tot}}{z} - \frac{z_J}{z}J_{tot})^2 + (\frac{1+K_j}{K_j})^2K_M^2)}\Bigg) \end{aligned} \tag{4.55}$$

This expression, although quite complex, simply corresponds to the quadratic rate equation of the Michaelis-Menten scheme f_{MM}^q, as introduced in [88], but with a re-scaling such that $H_{tot} \to \frac{H_{tot}}{z}$, $J_{tot} \to \frac{z_J}{z}J_{tot}$, $K_M \to \frac{1+K_j}{K_j}K_M$, and $k_J^h \to \frac{z}{z_J}k_J^h$. Compared to a proper Michaelis-Menten scheme, this re-scaling comes from of the consideration of the auto-regulatory mechanism and the nucleotide exchange transition.

$$f(H_{tot}, J_{tot}, K_M, k_J^h) = z f_{MM}^q(\frac{H_{tot}}{z}, \frac{z_J}{z}J_{tot}, \frac{1+K_j}{K_j}K_M, \frac{z}{z_J}k_J^h). \tag{4.56}$$

In detail, the Michaelis-Menten model of the ATP hydrolysis stimulation would ignore the release of phosphate, the exchange back to the ADP state and the auto-regulation of DnaJ, leading to the following rate equation:

$$\begin{aligned} &f_{MM}^q(H_{tot}, J_{tot}, K_M, k_J^h) \\ &= \frac{H_{tot}(k_J^h + k^h) + (k_J^h - k^h)(J_{tot} + K_M - \sqrt{2K_M(H_{tot} + J_{tot}) + (H_{tot} - J_{tot})^2 + K_M^2})}{2H_{tot}} \end{aligned} \tag{4.57}$$

Results

The ATPase activity induced by DnaJ was evaluated for two different proteins with highly similar sequences but from two distinct DnaJ classes: a class A DnaJ, Ydj1 (YY), and a class B, Sis1 (SS). These two proteins are yeast cytosolic DnaJs interacting with the cytosolic Hsp70, Ssa1. Despite their close resemblance in sequence, illustrated in Figure 4.14, Sis1 and Ydj1 show large differences in efficiency, with Sis1 exhibiting higher ability over Ydj1 in protein refolding [89]. This difference, however, cannot be attributed to variations in their stimulation of Ssa1 ATPase activity when Ydj1 and Sis1 are purified [90]. Furthermore, it was demonstrated that the proteins exhibit a higher efficiency when present in solution together, as opposed to alone [91]. This could imply that Sis1 and Ydj1 alone are not operating at their full potential.

To investigate the potential regulatory mechanisms of these DnaJ proteins, two chimera proteins, formed by swapping the J-domains of Ydj1 and Sis1, were also examined in the experiment, similarly to [90]. These chimeras, denoted as YS and SY, correspond to the Ydj1 J-domain attached to the remainder of the Sis1 protein and the Sis1 J-domain attached to the

Ydj1 protein, respectively, as pictured in Figure 4.14.

Due to intra-proteins co-evolution, those chimeras are considered to not be able to form the auto-inhibited state. Indeed, while Ydj1 J-domain can bind to its G/F-rich region, it may not be the case for Sis1 G/F-rich region and vice-versa for Sis1 J-domain and Ydj1 G/F-rich region.

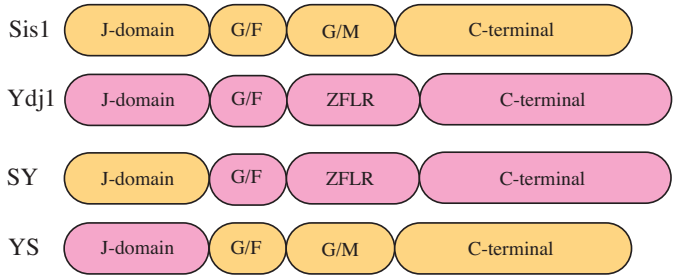

Figure 4.14: Schematic representation of Sis1, Ydj1 [92] and the chimeras YS and SY sequences. G/F characterizes the G/F-rich region, G/M the Glycine/Methionine-rich region, and ZFLR the zinc-finger-like region. Both the C-terminal region and the ZFLR define the protein activity [93; 94; 95].

Assuming that only the J-domain interacts with Hsp70, YY and YS, and SS and SY respectively, should share the same binding and stimulating ability, apart from the presence of the auto-inhibited state.

Therefore, any differences in the ATP hydrolysis rate can be attributed solely to the presence or absence of the auto-inhibited state J-. The ability of YY to stimulate ATP hydrolysis in Hsp70 was then compared to that of YS and SS to SY, with the assumption that YY and SS can be in the active or auto-inhibited state, while YS and SY can only be in their active state. The variations in the ATPase stimulation ability of YY, YS, SS and SY are depicted in Figure 4.15, where it is clear that both chimeras (YS and SY) have much larger ATPase stimulation abilities than the wild type (WT) SS and YY.

The ATPase activity measurements were fitted with the model outlined in (4.55). Parameters related to the J-domain, K_M and k_J^h, were considered identical for DnaJs with the same J-domain, thus K_M^Y and $k_J^{h,Y}$ were associated to YY and YS and K_M^S and $k_J^{h,S}$ to SS and SY. In addition, to account for the particularity of the chimeras, i.e. no regulatory mechanism, we set $K_j \to \infty$ for the SY and YS curves, while introducing K_j^Y for YY and K_j^S for SS. Finally, the coarse-grained rate of phosphate release and nucleotide exchange, $\hat{k}$, depending only on Hsp70, has to be the same for all curves.

K_M^Y [μM]	K_j^Y [-]	$k_J^{h,Y}$ [s^{-1}]	K_M^S [μM]	K_j^S [-]	$k_J^{h,S}$ [s^{-1}]	$\hat{k}$ [s^{-1}]
23.4	0.033	4.6	145.2	0.52	16.7	0.026

Table 4.5: Fitted parameters, fixed value of K^d to assure reasonable rates and $k_J^{h,S}$ was limited to a reasonable range of values.

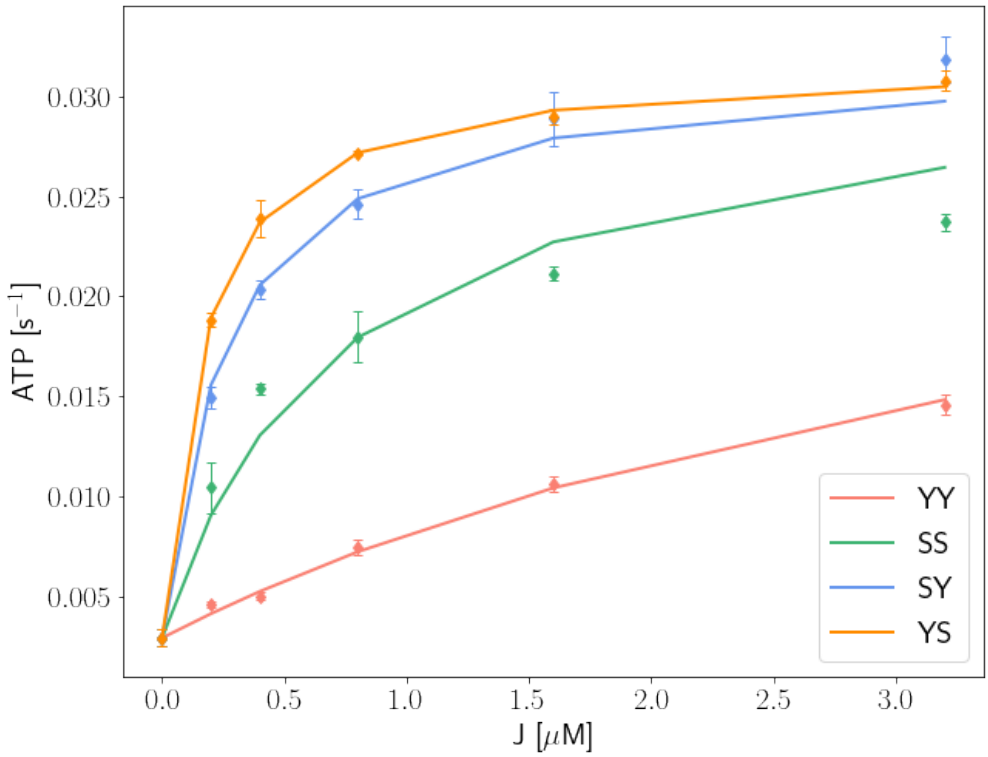

Figure 4.15: Measurement and fit on the three individual repetitions of the Hsp70 ATPase activity for WT DnaJ (YY and SS) and the chimeras (YS and SY). Data is represented as mean±SD. Experiments were performed with a fixed concentration of Hsp70 at 6μM, basal rate measured at 2.9×10^{-3} s^{-1}.

While YS and YY share the same J-domain, their impact on Hsp70 ATPase activity is drastically different; YS is the JDP allowing for the highest stimulation while YY displays the lowest stimulation ability (Figure 4.15). Although surprising, this large difference is fully explained by the strong tendency of YY toward the auto-inhibited state due to a low equilibrium value of K_j ($K_j = 0.03$ in Table 4.5).

To a lesser extent, the same behavior is observed for the pair SS and SY, where the smaller difference between the two curves is due to an equilibrium less titled toward the inactive state for SS ($K_j = 0.5$) compared to YY.

We observe that both chimeras curves are reaching a maximum value for the rate of ATP hydrolysis around 0.03 s^{-1}, which roughly corresponds to the nucleotide exchange rate $\hat{k} \sim k_x$ of Hsp70. Indeed, in ATPase steady-state experiments like the one performed here, ATP cannot be hydrolyzed faster than it is recovered in the Hsp70 cycle. The whole Hsp70 cycle is therefore limited by the exchange rate.

Despite this fact, our model is able to estimate the intrinsic hydrolysis rate due to the DnaJs, $k_J^{h,Y} = 4.6\ \text{s}^{-1}$ and $k_J^{h,S} = 16.6\ \text{s}^{-1}$, corresponding to the rate the system would reach if the hydrolysis rate was the limiting transition in the Hsp70 cycle.

Furthermore, the exchange as a rate-limiting step is also allowing the curves to reach saturation even when the DnaJs are present in substoichiometric quantities ($[H_{tot}] = 6\mu$M vs. max $[J_{tot}] = 3\mu$M). This is due to the difference in time scale between the nucleotide exchange and the binding of DnaJ to Hsp70·ATP. While one Hsp70 is going back to its initial ATP state, DnaJ has the opportunity to bind multiple Hsp70 and trigger their ATP hydrolysis. This highlights the catalytic role of DnaJ in the process.

Extension to HscA/HscB

Our model, as depicted in Figure 4.13, is not limited to the wild-type and chimeras experiment with Ydj1 and Sis1 and can be applied to other Hsp70/DnaJ pairing. We can therefore apply it on the data presented in Section 4.2, specifically the ATPase activity measurement of HscA during the titration of HscB. The data and fit are represented in Figure 4.16.

It is worth recalling that HscA and HscB are distinct from typical Hsp70 and DnaJ proteins, as they are known to interact exclusively with each other [79], whereas most Hsp70s and DnaJs are characterized by their ability to interact with various partner.

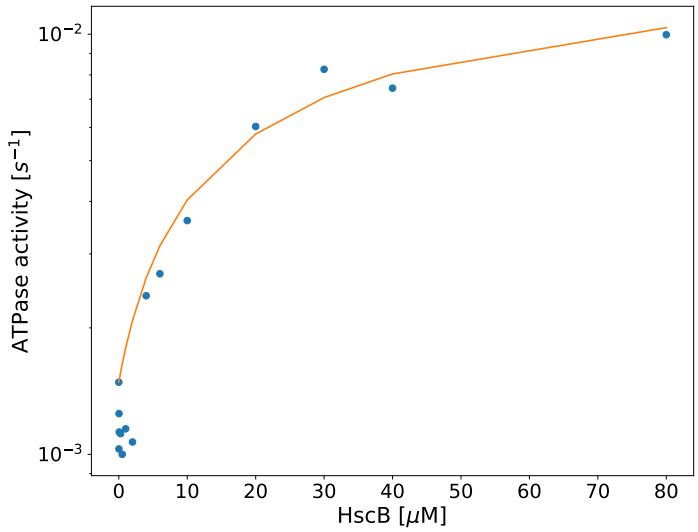

Figure 4.16: Fit of the HscA (Hsp70) ATPase activity during HscB (DnaJ) titration using the model presented on Figure 4.13 considering the presence of an auto-inhibited state J_{-}. Data point, (blue dots): Average of three steady-state experiments with a fixed concentration of HscA, $[H_{tot}] = 0.5\mu$M.

As observed in Table 4.6, HscB displays a very weak auto-regulatory mechanism with $K_j > 1$, indicating that the system is favoring the active state over the auto-inhibited state.

k^{on} [$\mu\text{M}^{-1}\text{s}^{-1}$]	K^d [μM]	K_j [-]	k_J^h [s^{-1}]	$\hat{k}$ [s^{-1}]
0.062	68.4	6.38	0.026	0.029

Table 4.6: Parameters obtained by fitting HscA ATPase activity during HscB titration in Figure 4.16. The basal hydrolysis rate is $k^h = 0.001485\ \text{s}^{-1}$

Furthermore, in contrast to previous observations for other Hsp70/DnaJ systems (Figure 4.15), the HscA cycle, triggered by HscB, is not limited by the exchange rate $\hat{k}$. Indeed, we found $\hat{k} \gtrsim k_J^h$. This result is in agreement with the biological knowledge of the HscA/HscB/IscU system as it is known to have a relatively fast exchange rate and to function properly without a nucleotide exchange factor (NEF).

The results obtained from these two observations (low regulatory mechanism and absence of NEF) demonstrate the uniqueness of the HscA/HscB/IscU system, in which the proteins have been specifically designed to interact exclusively with one another. This contrasts with the majority of Hsp70/DnaJ/substrate pairings, which exhibit greater versatility in their interactions and functions.

Finally, comparable values were found for the intrinsic hydrolysis rate k_J^h in Table 4.6 and the maximal acceleration induced by HscB obtained with the H_1/H_2 model introduced in Section 4.2. Indeed, from Table 4.6 we found that $k_J^h/k^h \sim 17$ while the maximal acceleration induced by HscB with the 2-state model is given by $a_J^{\max} = K_{1,J}^d/K_{2,J}^d \sim 10$. Overall, these results, although not equal, are relatively close to each other.

However, the estimation of the dissociation constant K^d between HscA and HscB is not consistent with the dissociation constant found in Section 4.2 when considering the same system. The estimated K^d is approximately 70μM, while with the H_1/H_2 model we find $K_{1,J}^d \sim 26\mu$M. This difference highlights that, despite providing an accurate description of the observed phenomena, our models may not necessarily provide precise estimations of all biological system parameters.

4.3.2 Effect of 3-body interactions

One key element of the Hsp70 cycle is the formation of a trimer involving Hsp70, DnaJ and the target protein, also known as substrate. However, the specificity of 3-body interaction has never been, to our knowledge, properly addressed in the context of Hsp70 chaperones and the different configurations of a trimer have simply been ignored.

Part of this issue was addressed in Section 4.2 when we considered the three different configurations of 2-bond trimers and the facilitated transition from a 2-bond trimer to a fully connected trimer due to the high local concentration ρ between proteins. We now want to propose a more complete description that also takes into account the differences between

binding a single protein to another single protein and binding a single protein to a dimer to form a 2-bond trimer.

The statistical properties of a system containing interacting particles can be obtained by its partition function Z. To describe the pair-wise interaction between a unique protein i and a dimer jk when the protein i bind to one of the elements of the dimer, j, we define the interaction potential between i and j as $V_{i,jk}$. The partition function thus becomes:

$$\begin{aligned} Z &= \frac{1}{\lambda^{3N} N! V^N} \int_V \prod_i^N d\vec{r}_i \; e^{-\beta \sum_i^N V_{i,jk}} = \frac{1}{\lambda^{3N} N! V^N} \left(\int_V d\vec{r} e^{-\beta V_{i,jk}} \right)^N \\ &= \frac{1}{\lambda^{3N} N! V^N} \left(\int_V d\vec{r}\, 1 + (e^{-\beta V_{i,jk}} - 1) \right)^N = \frac{1}{\lambda^{3N} N! V^N} \left(V + \int_V d\vec{r} (e^{-\beta V_{i,jk}} - 1) \right)^N \end{aligned} \tag{4.58}$$

where V is the volume available to the system and N the number of particles i.
From the virial theorem, this quantity can be expanded such that: c

$$Z = \frac{1}{\lambda^{3N} N! V^N} \left(V^N + N V^{N-1} \int_V d\vec{r}\, (e^{-\beta V_{.,jk}} - 1) + \frac{N(N-1)}{2} V^{N-2} \left(\int_V dr\, (e^{-\beta V_{i,jk}} - 1) \right)^2 + ... \right) \tag{4.59}$$

Stopping at the second term, the expansion becomes:

$$Z \simeq \frac{1}{\lambda^{3N} N!} \left[1 + \frac{N}{V} \int_V d\vec{r}\, (e^{-\beta V_{.,jk}} - 1) \right] = \frac{1}{\lambda^{3N} N!} \left[1 + \frac{B}{V} \right] \tag{4.60}$$

where B is the second virial coefficient linked to pair-wise interaction between particles.

Due to the presence of k, the interaction between i and j can be modified as such: $V_{i,jk}(r) = V_{i,j}(\vec{r}) + V_{rep}(\vec{r})$, where $V_{i,j}$ is the interaction potential between i and j and is defined in an area Ω_j around the protein j and V_{rep} is a repulsive potential around the protein k, in a space defined as Ω_k. Those two areas, Ω_j and Ω_k are partially superimposed due to the proximity between j and k.
Therefore the integral defining B can be separated into three distinct regions:

$$\begin{aligned} \int_V d\vec{r}\, (e^{-\beta V_{.,jk}} - 1) &= \int_{\Omega_j \cap \Omega_k'} d\vec{r}\, (e^{-\beta V_{i,j}} - 1) + \int_{\Omega_j \cap \Omega_k} d\vec{r}\, (e^{-\beta (V_{i,j} + V_{rep})} - 1) + \int_{\Omega_k \cap \Omega_j'} d\vec{r}\, (e^{-\beta V_{rep}} - 1) \\ &= \int_{\Omega_j} d\vec{r}\, (e^{-\beta V_{i,j}} - 1) - \int_{\Omega_j \cap \Omega_k} d\vec{r}\, (e^{-\beta V_{i,j}} - e^{-\beta (V_{i,j} + V_{rep})} + e^{-\beta V_{rep}} - 1) + \int_{\Omega_k} d\vec{r}\, (e^{-\beta V_{rep}} - 1) \end{aligned} \tag{4.61}$$

We can therefore express B as $B = B_{i,j} - B_{\Omega_j \cap \Omega_k} - |B_{rep}|$, where B_{ij} represent the interaction between i and j in absence of k, B_{rep} the absence of interaction due to the volume take by k and $B_{\Omega_j \cap \Omega_k}$ the impact of k on the interaction of i with j.
Linking the second virial coefficient to the dissociation constant with the equation $K^d =$

$-\frac{1}{2N_AB}$, we found that

$$\begin{aligned} K^d_{i,jk} &= -\frac{1}{2N_A(B_{i,j} - B_{\Omega_j\cap\Omega_k} - |B_{rep}|)} \\ &= -\frac{1}{2N_AB_{i,j}}\frac{1}{1 - B_{\Omega_j\cap\Omega_k}/B_{i,j} - |B_{rep}|/B_{i,j})} \\ &= K^d_{i,j}\exp(-\beta k_B T\ln(1 - B_{\Omega_j\cap\Omega_k}/B_{i,j} - |B_{rep}|/B_{i,j})) \\ &= K^d_{i,j}\exp(\beta\epsilon) \end{aligned} \tag{4.62}$$

where $\epsilon = |k_B T\ln(1 - B_{\Omega_j\cap\Omega_k}/B_{i,j} - |B_{rep}|/B_{i,j})|$.
Assuming that the unbinding rate of i with j is not impacted by the presence of the protein k, we finally found that:

$$k^{\text{on}}_{i,jk} = k^{\text{on}}_{ij}\exp(-\beta\epsilon) \tag{4.63}$$

The $|B_{rep}|/B_{i,j}$ parameter can easily be computed. Using the hard sphere assumption:

$$V_{rep}(\vec{r}) = \begin{cases} \infty & \forall\vec{r}\in\Omega_k \\ 0 & \forall\vec{r}\notin\Omega_k \end{cases} \tag{4.64}$$

Therefore, we found $|B_{rep}|/B_{i,j} = \Omega_k N_A K^d_{i,j}$. Assuming that Hsp70 is trying to bind to the substrate, already attached to DnaJ, we have that:

$$\Omega_k = \frac{4}{3}\pi R^3 = \frac{4}{3}\times\pi\times(60\times10^{-9})^3 \sim 1\times10^{-21}L \tag{4.65}$$

$$N_A K^d_{i,j} = 6\times10^{23}\times4.5\times10^{-6} \sim 3\times10^{18}L^{-1} \tag{4.66}$$

where $R = 60$Å is the radius of gyration of DnaJ and $K^d = 4.5\mu$M is the dissociation constant between Hsp70 and the substrate. Therefore, we find

$$|B_{rep}|/B_{i,j} \sim 10^{-3} \tag{4.67}$$

which becomes negligible. Therefore, the excluded volume due to the presence of DnaJ around the substrate does not significantly impact the interaction between Hsp70 and the substrate.

However, the presence of DnaJ "hides" part of the substrate volume, impairing its interaction with Hsp70. This impact is characterized by $B_{\Omega_j\cap\Omega_k}$. Therefore, we obtain that $\epsilon \sim |k_B T\ln(1 - B_{\Omega_j\cap\Omega_k}/B_{i,j})|$.

Hsp70 cycle model

To ensure a full description of the Hsp70 cycle, we consider all the possible interactions between the three proteins, Hsp70, DnaJ and substrate, from dimers and trimers formation to ATP/ADP exchange and hydrolysis/synthesis.

To ease the description, the two Hsp70·ATP allosteric states H_1 and H_2, presented above, are

coarse-grained in a unique state, H_T, and the Hsp70·ADP state is represented by H_D. DnaJ, marked as J and the substrate, as S, can interact with H_T and H_D independently or as the dimer JS. As in Section 4.2, H_TJS and H_DJS refer to a 2-bond trimer, with either H, J or S bound to the two proteins and $\boxed{H_TJS}$, $\boxed{H_DJS}$ to the fully connected trimer. A more detailed description of the model and the specific reactions and reaction rates can be found in the appendix. In this section, we will provide only the essential information needed to understand the model features.

The transitions between ATP and ADP states are as follows and dictated by both the exchange rates (2.25) and the hydrolysis and synthesis rates:

$$H_T \rightleftharpoons H_D, \quad H_TS \rightleftharpoons H_DS, \quad H_TJ \rightleftharpoons H_DJ, \quad H_TJS \rightleftharpoons H_DJS, \quad \boxed{H_TJS} \rightleftharpoons \boxed{H_DJS}. \tag{4.68}$$

Here and in all the following description H_TJS and H_DJS refer to three combinations of 2-bond trimer, i.e. H bound to S and J, S bound to H and J and J bound to H and S.

The dimers formation transitions simply depend on the binding and unbinding rates between the two proteins and follow:

$$J + S \rightleftharpoons JS \tag{4.69}$$
$$H_T + J \rightleftharpoons H_TJ \tag{4.70}$$
$$H_T + S \rightleftharpoons H_TS \tag{4.71}$$
$$H_D + J \rightleftharpoons H_DJ \tag{4.72}$$
$$H_D + S \rightleftharpoons H_DS \tag{4.73}$$
$$\tag{4.74}$$

The reactions[2] of the 2-bond trimers formation where the binding rate is penalized by a cost ϵ due to the excluded volume of the pre-existing dimer, such that $k^{\text{on}} \longrightarrow e^{-\beta\epsilon}k^{\text{on}}$ are given below:

$$JS + H_T \rightleftharpoons H_TJS \qquad H_TS + J \rightleftharpoons H_TJS \qquad H_TJ + S \rightleftharpoons H_TJS \tag{4.75}$$
$$JS + H_D \rightleftharpoons H_DJS \qquad H_DS + J \rightleftharpoons H_DJS \qquad H_DJ + S \rightleftharpoons H_DJS \tag{4.76}$$

Finally, the formation of the fully connected trimer is promoted by the co-localization of the two remaining proteins not bound to each other, mimicking a high concentration ρ.

$$H_TJS \rightleftharpoons \boxed{H_TJS}, \qquad H_DJS \rightleftharpoons \boxed{H_DJS} \tag{4.77}$$

[2] It is important to acknowledge that the equations presented here for the formation of the Hsp70-DnaJ-substrate trimer have been simplified for the sake of clarity. However, the different configurations of the trimer have been considered in the analysis.

Results

Choosing physiological concentrations of Hsp70, DnaJ and substrates ($[H_{tot}] = 10\mu$M, $[J_{tot}] = 1\mu$M, $[S_{tot}] = 0.3\mu$M) and binding and unbinding rates from the literature [74; 96; 97], we are trying to reproduce the actual dynamics of the Hsp70 cycle in biological systems.

As already mentioned, the ternary complex formed by Hsp70, DnaJ and the substrate is the key element of the Hsp70 cycle, promoting the hydrolysis rate to drive the cycle in a specific direction and acting on misfolded substrates. This trimer is mostly formed by an important flux of the dimer JS into the Hsp70 cycle to bind H_T as depicted in Figure 4.17. Increasing ϵ at first has a positive outcome on the flux of JS into the cycle and reaches a maximal value for $\epsilon \simeq 5k_BT$. However, higher values of ϵ start to have a negative impact on the system and the flux of JS entering the cycle decreases.

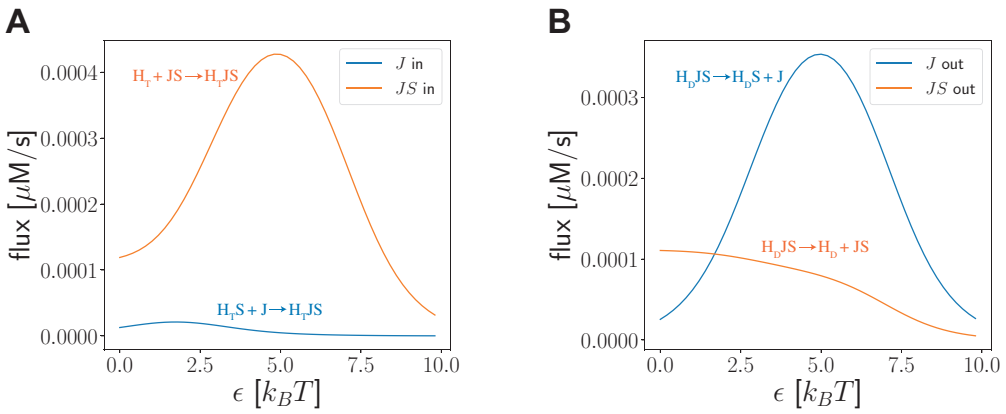

Figure 4.17: Net flux of DnaJ in and out of the Hsp70 cycle in function of the cost ϵ. **A**: Inward flux of J and JS to form the trimer H_TJS before the hydrolysis rate. **B**: Outward flux of J and JS from the H_DJS trimer.

After the hydrolysis, both J alone and JS can leave the trimer formed now with H_D (Figure 4.17B). While a low value of ϵ favors the release of the dimer JS, increasing the parameter allows the expulsion of DnaJ alone, allowing Hsp70 to act on the substrate. Indeed, for Hsp70 to act on the substrate it needs to hold onto it in its ADP form and release it after the nucleotide exchange with the opening of the substrate binding domain.

As for the in-flux, the out-flux of DnaJ is optimized for $\epsilon \sim 5k_BT$. In this case, DnaJ enters in the Hsp70 cycle while in the JS dimer, therefore bringing the substrate to Hsp70 and exits as J alone immediately after the hydrolysis. A larger ϵ (over 5 k_BT) leads to a decrease in these mechanisms due to the difficulty in forming any kind of trimers.

Understanding how and why the addition of an energy cost to the formation of trimers can, to some extent, increase the concentration of certain trimers and the release rate of DnaJ requires a deeper understanding of the underlying mechanisms involved.

The key element is the change in affinity between Hsp70 and DnaJ during the Hsp70 cycle.

Indeed, in its ATP state, Hsp70 has a strong affinity for DnaJ ($K^d_{H_TJ} = 0.07\mu$M [74]), while the configurational changes due to the ATP hydrolysis in Hsp70 decrease strongly its affinity with DnaJ. As no references are reporting any values for the dissociation constant, we assumed it to be large enough to ensure DnaJ unbinding with H_D ($K^d_{H_D,J}$=500 μM).

This decrease in affinity allows DnaJ to detach from H_D from the fully connected trimer right after the ATP hydrolysis. However, for very low values of ϵ, DnaJ is easily maintained in the trimer due to its high affinity for the substrate, as illustrated in Figure 4.18.

Increasing ϵ affects the trimer stability by strongly impairing the balance between the binding and unbinding of elements in the trimer. This effect will prevent any rebinding of DnaJ exiting the trimer, hence increasing the net flux of DnaJ out of the cycle, leaving the fully connected trimer.

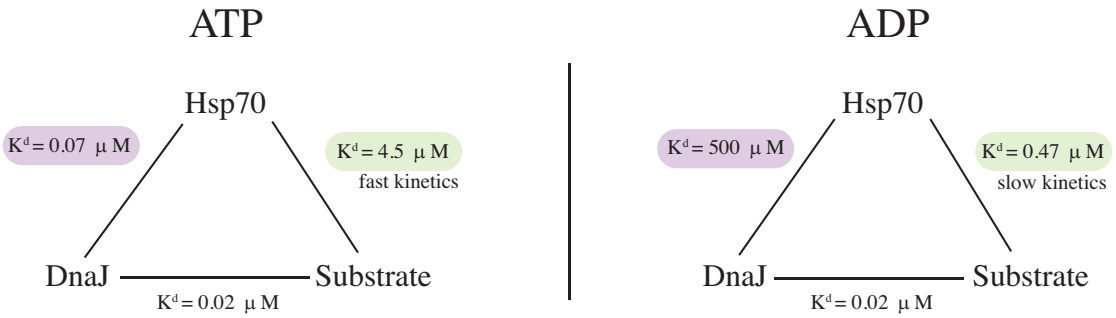

Figure 4.18: Schematic description of the affinity between the three proteins involved in the Hsp70 cycle and their modification due to the nucleotide change. Marked in purple, the dissociation constant of DnaJ and Hsp70 which increases drastically when Hsp70 is in ADP state, and in green the dissociation constant of Hsp70 with the substrate which increases strongly in the ADP state.

Increasing ϵ does not have the same effect on the flux of JS out of the trimer as seen in Figure 4.17B. Indeed, most of this flux comes from the unbinding of DnaJ to H_D from the 2-bond trimer configuration where S and H_D are not bound to each other. As the substrate is already unbound from, or never was bound to, Hsp70, the sudden decrease in affinity between DnaJ and Hsp70 immediately pushes the dimer JS away. Increasing ϵ does not reinforce this mechanism but actually leads to greater instability of those 2-bond trimers where the substrate is not bound to Hsp70. This leads to a decrease in their concentrations, thus implying a decrease in the release of JS.

As the total outward flux of DnaJ increases, the total inward flux must also increase due to mass conservation. While Hsp70 has a higher binding affinity for J compared to S, the fast kinetics of its interaction with S allows for rapid binding to the substrate, even as ϵ increases. Thus, it favors the binding of Hsp70 to the substrate in the JS dimer. Once the H_TSJ complex is formed, characterized by the substrate being bound to both H_T and J, the fast unbinding rate between H_T and S is compensated by the co-localization of H_T and J promoting their binding. Therefore, the co-localization is able to stabilize the trimer in the fully connected

state $\boxed{H_T JS}$ before the unbinding of Hsp70 and the substrate.

Without DnaJ, the substrate can easily bind and unbind to Hsp70, leading to a low concentration of $H_T S$ which implies a constant low flux of J binding to $H_T S$ for all ϵ.

Therefore, it seems that a high enough value of ϵ is able to select and amplify a precise selection of reactions while decreasing the others.

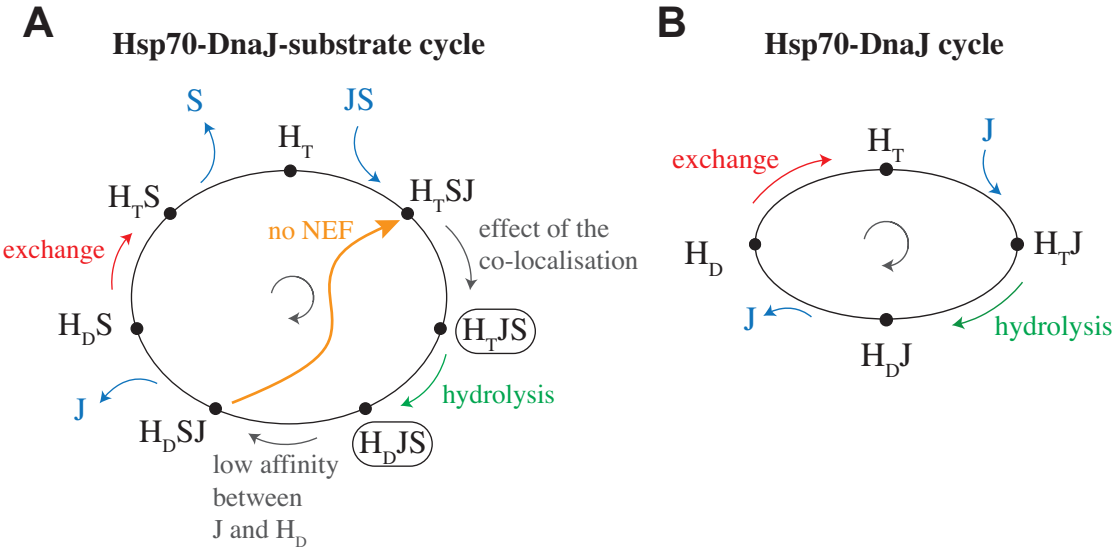

Figure 4.19: **A**: Hsp70-DnaJ-substrate (HJS) cycle where Hsp70 act on a substrate, S, assisted by DnaJ (J). **B**: Hsp70-DnaJ (HJ) cycle where Hsp70 ATPase activity is only stimulate by DnaJ.

Overall, those selected reactions, displayed in Figure 4.19A, seem to correspond to the ones presented in the usual description of the Hsp70 cycle in the literature. DnaJ selects the substrate and brings it to Hsp70. The three proteins bind together forming a ternary complex and the presence of both DnaJ and the substrate easily stabilizes Hsp70 in its allosterically active state leading to a high ATPase activity. The hydrolysis provokes the closure of the SBD, trapping the substrate in place, while other configurational changes decrease the affinity between DnaJ and Hsp70·ADP. DnaJ is then expelled from the complex and a nucleotide exchange allows the SBD to open, leading to the release of the substrate, resetting Hsp70 in its original state H_T. However, the release of DnaJ and then of the substrate is slowed down by the exchange rate between $H_D JS$ and $H_T JS$ resetting the trimer at the beginning of the cycle, before the hydrolysis rate and keeping the substrate bound to Hsp70 for a long time.

This phenomenon is likely due to the absence of nucleotide exchange factors (NEF) in our simulation. The NEF is known to promote the unbinding of ADP and the binding of ATP when it is bound to Hsp70. While the proper mechanisms of its function are not fully known, we can reasonably assume the NEF would bind with more ease to Hsp70 alone or bound to J or S than directly to a trimer. Indeed, as the formation of a trimer is already impaired by steric effects, a tetramer, a complex made of four proteins, will be even less probable. Therefore, we can

hypothesize that the exchange between H_DJS and H_TJS will not benefit from the presence of the NEF, contrary to the reaction $H_DS \rightarrow H_TS$, pushing the system toward releasing substrates faster.

Confirming our deductive reasoning, experimental observations have shown that a decrease in the release of substrate correlates with a decrease in the NEF activity [98]. For example, during a heat shock, Hsp70 and DnaJ efficiently target misfolded proteins and prevent their aggregation but the refolding only starts once the system is back in normal condition [99]. This is likely due to the denaturation of the NEF during the heat shock which leads to "the sequestering of protein substrates at heat-shock temperatures" [100].

While the three proteins indeed interact with each other according to the usual Hsp70 cycle as displayed in Figure 4.19A, the surplus in DnaJ concentration compared to the substrates results in a second and more prevalent cycle. The large amount of free DnaJ binds to Hsp70·ATP, promotes the ATP hydrolysis, and exits the cycle without presenting a substrate to Hsp70, as illustrated in Figure 4.19B.

Although the Hsp70-DnaJ cycle does not impair Hsp70 ability to act on substrates, due to the excess concentration of Hsp70 and DnaJ compared to the substrates, it is energetically costly. Despite the presence of high levels of ATP in both our simulation and in vivo, where the concentration of ATP ranges from 1-10 mM, it is unlikely that biological systems would have evolved to waste such an important resource through useless Hsp70 cycles.

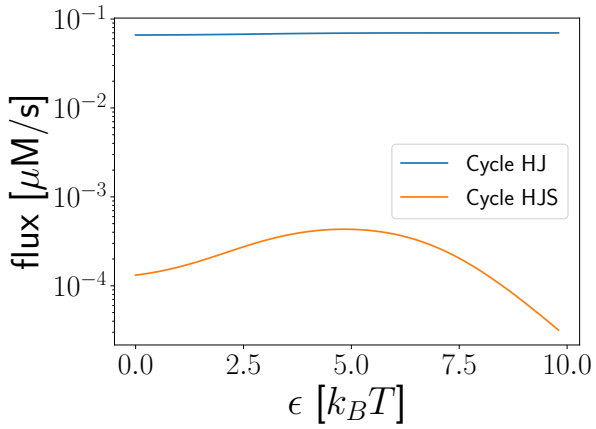

Figure 4.20: Total fluxes of the cycle presented in Figure 4.19, in blue the HJ cycle and Hsp70 and in orange the HJS cycle.

This observation highlights the need for DnaJ regulation mechanism in an effort to reduce the hydrolysis of ATP outside of cycles targeting a substrate.

4.3.3 DnaJ's auto-regulatory mechanism in the full Hsp70 cycle

Considering the auto-inhibited state of DnaJ, J_-, introduced in Section 4.3.1 we add to our model the following reaction

$$J \rightleftharpoons J_- \tag{4.78}$$

dictated by the equilibrium constant K_j. As before, J_- cannot bind to Hsp70.

It is not yet clear whether DnaJ in its auto-inhibited state is able to bind to the substrate, nor what the effect of the substrate is, if it can actually bind.

Assuming that an interaction between J_- and S is possible, the following reaction

$$\begin{aligned} JS &\rightleftharpoons J_-S \\ J_- + S &\rightleftharpoons J_-S \end{aligned} \tag{4.79}$$

dictated by K_j^S the equilibrium constant between JS and J_-S and the dissociation constant $K^d_{J_-,S}$, are also added to the Hsp70 cycle model presented above.

Due to detailed balance condition (interaction cycle presented in Figure 4.21), the equilibrium constant between JS and J_-S, K_j^S is given by:

$$K_j^S := \frac{[JS]}{[J_-S]} = \frac{K^d_{J_-,S}}{K^d_{J,S}} K_j \tag{4.80}$$

where $K^d_{J,S}$ is the dissociation constant between J and S.

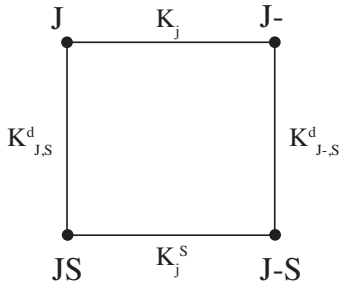

Figure 4.21: Interaction cycle of the substrate and DnaJ in active state and auto-inhibited state.

An efficient regulatory mechanism will push most of the individual DnaJ in the state J_-, implying $K_j < 1$, while promoting the JS dimer, $K_j^S > K_j$, to assure substrates are passed to and processed by Hsp70.

This condition requires J_-S to be less stable than the JS, meaning that $K^d_{J_-,S} > K^d_{J,S}$. In the limits of $K^d_{J_-,S} \to \infty$, the equilibrium between JS and J_-S will be completely titled toward the JS dimer:

$$\lim_{K^d_{J_-,S} \to \infty} K_j^S \to \infty \tag{4.81}$$

This means that to maximize the substrate impact on DnaJ regulation, J_- should not be able to bind to the substrate.

Due to the lack of evidence on the interactions between the substrate and DnaJ in its auto-inhibited state, we will consider two opposite and extreme cases.

The first case assumes that the substrate has no impact on the auto-regulatory mechanism of DnaJ $K_j^S = K_j \Rightarrow K_{J_-,S}^d = K_{J,S}^d$, such that the substrate has the same affinity for J than for J_-.

The second case proposes that the substrate is strongly pushing DnaJ toward its active state: $K_j^S \to \infty \Rightarrow K_{J_-,S}^d \to \infty$, therefore the substrate cannot bind to J_-.

To fully represent the behavior and performance of the Hsp70-DnaJ-substrate cycle, we are defining two parameters of particular interest; ϕ_S^{out}, the net flux of substrate out of the Hsp70-DnaJ-substrate cycle, and η, the efficiency of the cycle in releasing substrate, as defined below.

Assuming the HJS cycle behaves as proposed in Figure 4.19A, the substrate is released from H_TS state after the ADP/ATP exchange. This allows us to define ϕ_S^{out} as:

$$\phi_S^{out} = [H_TS]\, k_{H_TS}^{\text{off}} - [H_T][S]\, k_{H_TS}^{\text{on}} \tag{4.82}$$

The efficiency of the cycle measures the ratio between the released of substrates ϕ_S^{out} and the use of ATP, ϕ_{ATP}, such that

$$\phi_{\text{ATP}} = [H_T]\, k^h - [H_D]\, k^s + [H_TS]\, k_S^h - [H_DS]\, k_S^s + [H_TJ]\, k_J^h - [H_DJ]\, k_J^s + [H_TJS]\, k_{JS}^h - [H_DJS]\, k_{JS}^s \tag{4.83}$$

Therefore, the efficiency η is defined as

$$\eta = \frac{\phi_S^{out}}{\phi_{\text{ATP}}} \tag{4.84}$$

The fluxes of substrate out of the ATP cycle, represented by ϕ_S^{out} and the efficiency, represented by η, are directly impacted by variations in K_j and ϵ as demonstrated in Figures 4.22 and 4.23. The impact of the substrate on the DnaJ regulatory mechanism can be fully appreciated when comparing Figure 4.22 where the presence and binding of the substrate to DnaJ does not impact the equilibrium between the auto-inhibited and the active state of DnaJ and Figure 4.23 where the substrate can fully counter the effect of the regulation.

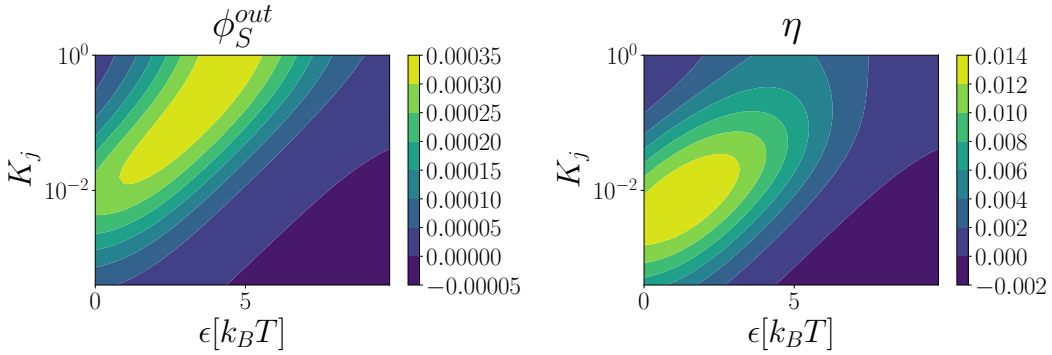

Figure 4.22: ϕ_S^{out} and η in function of ϵ and K_j, in the assumption that the substrate doesn't influence the regulation of DnaJ.

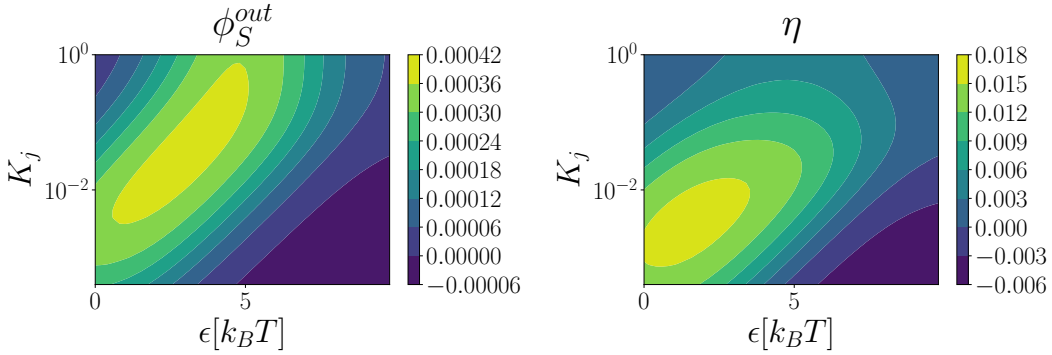

Figure 4.23: ϕ_S^{out} and η in function of ϵ and K_j, in the assumption of no binding between J_- and S

As previously established (Figure 4.17), there is an optimal value for ϵ for which the HJS cycle is promoted. This also leads to an optimization of substrate flux out of the ATP cycle. However, this optimal ϵ changes in the function of DnaJ regulatory ability (K_j). A low value of K_j implies a lower value for the optimal ϵ whether the substrate modifies the equilibrium between the auto-inhibited state and the active state of DnaJ (Figure 4.23) or not (Figure 4.22).

The maximal value for ϕ_S^{out} in Figure 4.22, $\max(\phi_S^{out}) = 0.00034\mu$M/s, is found at $\epsilon = 4.4k_BT$ and K_j=0.8, while in Figure 4.23 we found $\max(\phi_S^{out}) = 0.00038\mu$M/s at $\epsilon = 3k_BT$ and $K_j = 0.03$.

Therefore, the substrate can improve a little bit the regulation of DnaJ by avoiding binding to its auto-inhibition form but the impact is not found to be large.

In both cases, for lower values of K_j than represented ($K_j < 10^{-3}$), the auto-inhibition mechanism would be too strong, preventing any interaction of DnaJ with Hsp70, therefore canceling the HJS cycle.

Interestingly, the optimal conditions for the release of substrates are not the same as the optimal conditions for the efficiency η. Indeed, η maximal values are found for lower K_j and ϵ. In Figure 4.22, $\max(\eta) = 0.014$ at $\epsilon = 1.4 k_B T$ and $K_j = 0.008$ and in Figure 4.23, $\max(\eta) = 0.016$ at $\epsilon = 1.6 k_B T$ and $K_j = 0.003$.

This difference in optimal conditions is explained by K_j ability to reduce the use of ATP in useless cycles, i.e. Hsp70 cycles not acting on the substrate. Indeed, by pushing free DnaJ in the auto-inhibited state, a low K_j is preventing them to interact with Hsp70 in an HJ cycle.

The optimal values for K_j, are found to be within the range of biologically relevant values. Indeed, in the previous section, we identify $K_j = 0.52$ for Sis1 and $K_j = 0.033$ for Ydj1 (in Table 4.5) which are nicely within the range of the optimal values for ϕ_S^{out}.

Finally, it should be noted that the negative flux and efficiency that arise for very low K_j and high ϵ is due to the system's inability to form trimers under these conditions. While the low K_j pushes a large proportion of DnaJ into the J_- state, the high ϵ impairs the formation of trimers. Those two phenomena together prevent the functioning of the HJS cycle. Therefore, Hsp70 alone acts on the substrate in an HS cycle, similar to the HJ cycle in Figure 4.19B. This results in the substrate exiting the cycle from Hsp70·ADP and entering it by binding to Hsp70·ATP, leading to negative out-flux ϕ_S^{out}.

4.3.4 Summary

In this section, we investigated the role and impact of DnaJ in the Hsp70 cycle. We first introduced a very basic model of Hsp70 and DnaJ interaction during the Hsp70 cycle taking into account an auto-regulatory mechanism for DnaJ. Through experimentation with wild-type and chimeric forms of Sis1 and Ydj1, we were able to demonstrate the existence of an auto-inhibited state of DnaJ and highlight its impact on Hsp70 ATPase activity.

Applying the same model to the HscA/HscB system demonstrates the specificity of that same system. Indeed, the HscA/HscB system was found to have a nearly absent regulatory mechanism and a fast exchange rate, in comparison to experiments with Sis1 and Ydj1.

To fully represent the role of DnaJ in the Hsp70 cycle, we employed a kinetic model considering all possible interactions between Hsp70, DnaJ, and the substrate. Among the multitude of possible interactions, our model successfully reproduced the role of DnaJ in Hsp70 cycle as described in the literature, in which DnaJ brings the substrate to Hsp70 and leaves on its own right after the ATP hydrolysis. Interestingly, the presence of a cost impairing the trimer formation initially promotes the functioning of this cycle.

However, a more prevalent cycle where DnaJ alone interacts with Hsp70 was identified, due to the higher concentration of DnaJ and Hsp70 compared to the substrate concentration. This cycle highlighted the necessity of an auto-regulatory mechanism for DnaJ, by minimizing the unnecessary consumption of ATP in cycles where Hsp70 is not interacting with the substrate and therefore improves the efficiency of the protein refolding process.

Both the cost associated with trimer formation and the auto-regulatory mechanism of DnaJ

impacted the function of the HJS cycle. A strong interplay between these two parameters was found such that a stronger auto-regulatory mechanism implied a lower cost for the system to operate in optimal conditions.

Overall, we identified two novel parameters associated with the presence of DnaJ which may have a significant impact on the functioning of the Hsp70 cycle. Experimental studies are now needed to gain more insight into the role of these parameters and confirm their importance for the chaperone function of Hsp70.

5 Concluding remarks

Our work brings new insights into the physics of Hsp70 and its interaction with its co-chaperone DnaJ during its ATP cycle. The models we developed were intended to be generic enough to be applied to various Hsp70 systems, rather than being specific to a particular system. This approach is designed to highlight the fundamental characteristics of Hsp70 and DnaJ, and the nature of their interactions, which are essential to gain a better comprehension of protein disaggregation and refolding mechanisms.

As Hsp70 typically targets specific proteins, particularly misfolded proteins or protein aggregates, we first investigated the general mechanism of substrate selection in biological systems. Using a minimal model of enzymatic substrate-to-product transformation and with explicit energy consumption transitions, we studied the impact of different selection mechanisms in out-of-equilibrium dynamics. Our results showed that accurate and fast substrate selection could not only rely on the difference in affinity, as proposed in the kinetic proofreading scheme [25; 26], but should also rely on catalytic discrimination, where the recognition of the correct substrate induces faster transitions than an incorrect one. This phenomenon is well-known in systems renowned for their accuracy, such as tRNA translation, where the induced-fit mechanism has been overly studied but has not been fully considered in other systems, such as the Hsp70 cycle.

In the Hsp70 chaperone cycle, the ATPase activity of Hsp70 varies depending on the presence of a suitable substrate or its co-chaperone DnaJ. Furthermore, the joint interaction of DnaJ and the substrate amplifies drastically the ATP hydrolysis rate. We proposed a model that explains both the mechanism of the acceleration of the ATPase activity and the synergistic effect DnaJ and the substrate have on the ATPase rate. Our findings are also in agreement with experimental observations showing that impairing the affinity between DnaJ and the substrate results in a lower hydrolysis rate in the same concentration condition. Finally, we proposed that the synergy is a feature allowing to reach a fast hydrolysis rate at relatively low physiological concentrations of DnaJ and substrates.

Considering the acceleration of the ATP hydrolysis rate as a selection step, we further proposed that the collaboration between Hsp70 and DnaJ plays a crucial role in substrate selection. Indeed, to be efficiently processed through the Hsp70 cycle, a substrate needs to be able to interact individually with both proteins but also simultaneously, eliminating short proteins

such as peptides out of this process.

This idea was reinforced by the counterexample of the HscA/HscB/IscU system which did not display a very high acceleration of the hydrolysis rate. Because this system is designed to be highly specialized, it is plausible that it may not require a particularly strong selection process, and therefore may not necessitate a large acceleration of the ATP hydrolysis rate.

While considering the interactions of Hsp70, DnaJ and the substrate in the Hsp70 cycle, we found that steric effects were an important element that should not be neglected. While the absence of steric effects would lead to the release of the JS dimer right after ATP hydrolysis, preventing Hsp70 to act on the substrate, strong steric effects completely cancel the formation of trimers, preventing the DnaJ-bound substrate from even entering the Hsp70 cycle. It is only for moderate effect, associated to an energetic cost of a few $k_B T$ that the steric effects promote the binding of the dimer JS to Hsp70·ATP and release of DnaJ alone after ATP hydrolysis, in agreement with the general description of the Hsp70 cycle.

Furthermore, based on experimental evidence, we proposed that DnaJ exhibits a regulatory mechanism that influences its interaction with Hsp70, and that this mechanism plays a critical role in preventing wasteful cycles where no substrates are being treated by Hsp70. We suggested that this regulatory mechanism may be influenced by the presence of a substrate, adding an additional layer to the selection mechanism of the Hsp70 cycle. Indeed, while DnaJ allows Hsp70 to discriminate substrates by bringing them to it, the appropriate substrate could activate DnaJ by impacting its regulatory mechanism.

In our models, we predict the existence of four parameters, namely K_e, ρ, K_j, ϵ, which are essential to the function of the Hsp70 cycle but have not been experimentally validated yet. The first parameter, K_e, is at the base of Hsp70 ATPase activity and governs the Hsp70·ATP equilibrium between its allosteric states.

The parameter ρ corresponds to the apparent concentration of the proteins during the binding of the three proteins together and enables the synergy between DnaJ and the substrate. This mechanism seems to be at the root of Hsp70 selection due to DnaJ recruitment. Moreover, we observed that this parameter is essential for stabilizing the binding of the substrate onto Hsp70.

The regulation of the DnaJ protein is illustrated by K_j, which describes the equilibrium constant between DnaJ active and inactive states.

Finally, ϵ denotes the cost associated with the steric effects of combining three proteins to form a trimer and is associated with the volume those proteins occupy in space.

Experimental validation of these parameters should be performed to confirm our findings and would be a crucial step in understanding the dynamics and interactions in the Hsp70 cycle. While our proposed models provide important insights into the Hsp70 system and its relation with its co-chaperone DnaJ, there are still many aspects that require further investigation and improvement.

First, our synergy model could be extended to explicitly consider the ATP hydrolysis and

the nucleotide exchange transitions, in order to have a model adapted to any Hsp70 system, including those characterized by slower nucleotide exchange rates. It would also be important to test both experimentally and through simulations whether the synergy between Hsp70 and DnaJ is indeed a selection mechanism.

Regarding DnaJ regulatory mechanism, more investigation is needed to characterize the substrate impact, and how these elements play out in ensuring accurate substrate selection.

Finally, the nucleotide exchange factor could be a significant element that has not yet been included in our proposed models and requires further studies, through experiments and simulations, to better understand its role in the system and its impact on the mechanism we studied in this work.

A Appendices

A.1 Selection mechanisms

A.1.1 Rates for the simulations

The Tables A.1 to A.3 gather the rates used for each simulation presented in the section. All rates were carefully chosen to respect detailed balance at equilibrium ($\alpha = \alpha_{eq}$), as stated above. Therefore, among the following rates, some were freely chosen while others were constrained by the equilibrium condition. In the following tables, KP is associated to the kinetic proofreading system, and the rates presented were used for Figure 3.3, C corresponds to the catalytic discrimination and Figures 3.6 and 3.7, finally KP+C is associated to the figures where the two selection systems could be turn on and off namely Figures 3.8 and 3.9.

		KP	C	Comparison
Conc. [μM]	$[R]$	1	10	5
	$[W]$	1	10	5
On rates $[s^{-1}\mu\text{M}^{-1}]$	k_X^{on}	95	95	100
	$k_{X^*}^{\text{on}}$	10^{-3}	10^{-3}	10^{-3}
Off rates $[s^{-1}]$	k_R^{off}	45	varies	4.5
	k_W^{off}	450	varies	$k_R^{\text{off}} \times \lambda$
	$k_{R^*}^{\text{off}}$	84	84	8.4
	$k_{W^*}^{\text{off}}$	840	84	$k_{R^*}^{\text{off}} \times \lambda$
	$k_{P_s}^{\text{off}}$	10^{-5}	10^{-5}	10^{-5}
	$k_{P_s^*}^{\text{off}}$	10^{3}	10^{3}	10^{3}

Table A.1: Concentration and binding and unbinding rates used for the simulation of the Hopfield's like scheme (KP), the catalysis discrimination (C), and when comparing the two (Comparison), with $\lambda \in [1, 100]$.

		KP	C	Comparison
Production rates $[s^{-1}]$	k_{RP}	1	1	100
	k_{WP}	1	1	100
"Reverse production" rates $[s^{-1}]$	k_{PR}	1.2×10^{-5}	1.2×10^{-5}	0.012
	k_{PW}	1.2×10^{-6}	1.2×10^{-5}	$0.012/\lambda$
Hydrolysis rates $[s^{-1}]$	k_R^{hy}	varies	varies	$k_W^{hy} \times \omega$
	k_W^{hy}	varies	varies	0.1
Synthesis rates $[s^{-1}]$	k_R^{syn}	varies	$\frac{(7.98 k_R^{hy})}{k_S^{\text{off}}}$	$0.19\ k_R^{hy}$
	k_S^{syn}	varies	$\frac{(7.98 k_W^{hy})}{k_S^{\text{off}}}$	0.019

Table A.2: Rates of products formation and destruction and NTP hydrolysis and synthesis in a Hopfield's like scheme (KP), in the catalysis discriminating case (C) and when comparing the two (Comparison), with $\lambda \in [1, 100]$, $\omega \in [1, 1000]$.

Enzyme bound to:		$\varnothing$	P_s	S (KP)	S (C)	S (Comparison)
On rates $[s^{-1}\mu\text{M}^{-1}]$	k^{D+}	10	2×10^{-5}	0.0018	$5 \times 10^{-4} (k_S^{\text{off}})^{1/3}$	0.0018
	k^{T+}	10	4.6×10^{6}	5.7×10^{4}	$2 \times 10^{-5} (k_S^{\text{off}})^{-1/3}$	5.7×10^{4}
Off rates $[s^{-1}]$	k^{D-}	10	10^{-10}	10^{-5}	10^{-5}	10^{-5}
	k^{T-}	10	2×10^{-5}	0.0018	$5 \times 10^{-4} (k_S^{\text{off}})^{1/3}$	0.0018

Table A.3: Rates of binding and unbinding NTP and NDP for the enzyme alone ($\varnothing$) or bound to the produced (P_s) for each simulation.

Finally, their rate are explicitly identified on Figure A.1.

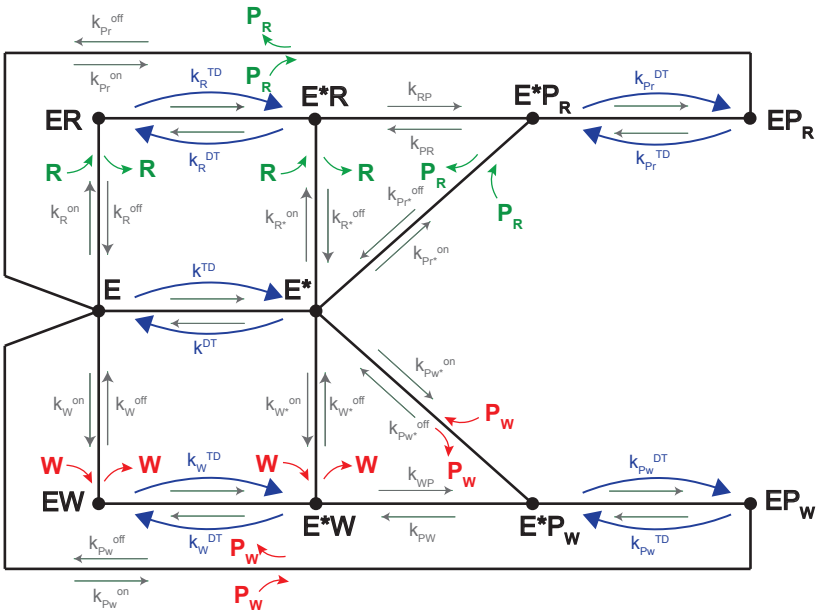

Figure A.1: Additional scheme of the proofreading system with explicit mention of the different rates

A.2 Hsp70 - Regulation of the ATPase activity

A.2.1 Additional details from the model

In addition to the five different states presented in the section, namely H, HS, HJ, HJS and $\boxed{HJS}$, we also considered complexes containing more than three proteins. Indeed, we assumed that HS could bind to JS to form the $HJSS$ complex such that the J-protein from the dimer will bind to HS in the following configuration: $S-J-H-S$. Similarly, a $HJJS$ complex can exist implying $J-S-H-J$ and a $HJJSS$ as $J-S-H-J-S$. The presence of those complexes only has an impact for really large concentrations of S and or J and allows us to avoid observing a Hook effect when drastically increasing the concentration of S or J in Figure 4.5 and Figure 4.6. The equilibrium concentrations of those states are given by:

$$\begin{aligned}
[H_iJJS] &= \frac{[H_i][J][J][S]}{K^d_{i,J}K^d_{i,S}K^d_{J,S}} \\
[H_iJSS] &= \frac{[H_i][J][S][S]}{K^d_{i,J}K^d_{i,S}K^d_{J,S}} \\
[H_iJJSS] &= \frac{[H_i][J][J][S][S]}{K^d_{i,J}K^d_{i,S}K^d_{J,S}K^d_{J,S}}
\end{aligned} \tag{A.1}$$

Furthermore, the actual computation for the estimation of the hydrolysis rate considers that the total concentration of substrate and/or DnaJ is different from their free concentrations. We actually have to find the free concentration of H_1, J and S by considering the total concentrations $[H]_{tot}$, $[S]_{tot}$ and $[J]_tot$ as constant. However, this system doesn't always have an algebraic solution [101] or a very unreadable one.

A.2.2 Distributions of the parameters estimated for the HscA/HscB/IscU system

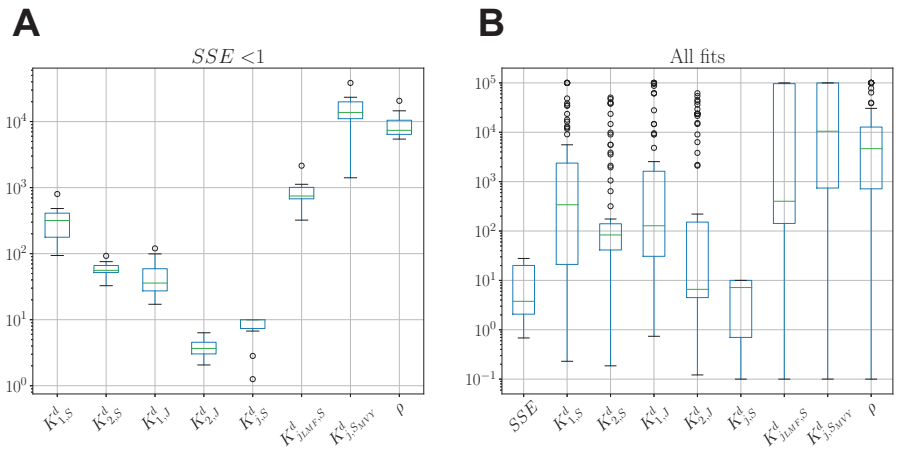

Figure A.2: Distribution of the fitted parameters for the parameters inducing a good fit **A** and all parameters **B**

A.2.3 Experimental data

The averaged data from the steady-state ATPase experiment of the HscA/HscB/IscU system presented in Figure 4.7 are given in Tables A.4 to A.10.

HscA [μM]	IscU [μM]	V_0 [s−1]
0.5	0	0.00182
0.5	0.015625	0.00151
0.5	0.031250	0.00159
0.5	0.125	0.0018
0.5	0.25	0.00171
0.5	0.5	0.00157
0.5	1	0.00147
0.5	2	0.00162
0.5	6	0.00159
0.5	10	0.00179
0.5	20	0.00272
0.5	40	0.00289
0.5	80	0.00387
0.5	100	0.00374
0.5	200	0.00442

Table A.4: Data from the titration of IscU

HscA [μM]	HscB [μM]	V_0 [s−1]
0.5	0	0.001485
0.5	0.015625	0.00103
0.5	0.03125	0.00125
0.5	0.0625	0.00113
0.5	0.25	0.00112
0.5	0.5	0.001
0.5	1	0.00115
0.5	2	0.00107
0.5	4	0.002385
0.5	6	0.00269
0.5	10	0.0036
0.5	20	0.00603
0.5	30	0.00825
0.5	40	0.00744
0.5	80	0.00997

Table A.5: Data from the titration of HscB

HscA [μM]	IscU [μM]	HscB [μM]	V_0 [s−1]
0.5	5	0	0.00233
0.5	5	0.01563	0.00308
0.5	5	0.03125	0.003795
0.5	5	0.0625	0.006875
0.5	5	0.125	0.0092
0.5	5	0.25	0.0148
0.5	5	0.5	0.0150
0.5	5	1	0.0216
0.5	5	2	0.0230
0.5	5	4	0.0331
0.5	5	6	0.0395
0.5	5	10	0.04325
0.5	5	20	0.0471
0.5	5	30	0.0465
0.5	5	40	0.0502

Table A.6: Data from the titration of HscB WT with a fixed 5 μM concentration of IscU

HscA [μM]	IscU [μM]	HscB [μM]	V_0 [s−1]
0.5	10	0	0.00199
0.5	10	0.015625	0.00344
0.5	10	0.03125	0.0041
0.5	10	0.0625	0.005735
0.5	10	0.125	0.009635
0.5	10	0.25	0.0215
0.5	10	0.5	0.01773
0.5	10	1	0.0259
0.5	10	2	0.03275
0.5	10	4	0.03405
0.5	10	6	0.0366
0.5	10	10	0.0501
0.5	10	20	0.0582
0.5	10	30	0.0536
0.5	10	40	0.0525

Table A.7: Data from the titration of HscB WT with a fixed 10 μM concentration of IscU

HscA [μM]	IscU [μM]	HscB [μM]	V_0 [s−1]
0.5	20	0	0.00234
0.5	20	0.015625	0.00416
0.5	20	0.03125	0.003997
0.5	20	0.0625	0.005443
0.5	20	0.125	0.00868
0.5	20	0.25	0.0304
0.5	20	0.5	0.02204
0.5	20	1	0.03307
0.5	20	2	0.04003
0.5	20	4	0.04547
0.5	20	6	0.0491
0.5	20	10	0.0519
0.5	20	20	0.05153
0.5	20	30	0.05315
0.5	20	40	0.0512

Table A.8: Data from the titration of HscB WT with a fixed 20 μM concentration of IscU

HscA [μM]	IscU [μM]	HscB LMF [μM]	V_0 [s−1]
0.5	20	0	0.002235
0.5	20	0.01563	0.002855
0.5	20	0.0313	0.002975
0.5	20	0.063	0.003175
0.5	20	0.13	0.003515
0.5	20	0.3	0.00428
0.5	20	0.5	0.00455
0.5	20	1	0.006005
0.5	20	2	0.006615
0.5	20	4	0.00914
0.5	20	6	0.0123
0.5	20	10	0.01675
0.5	20	20	0.0265
0.5	20	30	0.0355

Table A.9: Data from the titration of HscB LMF with a fixed 20 μM concentration of IscU

HscA [μM]	IscU MVY [μM]	HscB [μM]	V_0 [s−1]
0.5	20	0	0.002790
0.5	20	0.01563	0.00275
0.5	20	0.0313	0.00298
0.5	20	0.0625	0.00324
0.5	20	0.25	0.00338
0.5	20	0.5	0.003105
0.5	20	1	0.003265
0.5	20	2	0.00421
0.5	20	4	0.00679
0.5	20	6	0.00762
0.5	20	10	0.015
0.5	20	20	0.01455
0.5	20	30	0.0194
0.5	20	40	0.0201

Table A.10: Data from the titration of HscB WT with a fixed 20 μM concentration of IscU MVY

A.2.4 Additional structures

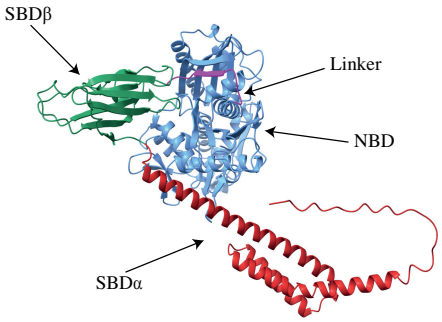

Figure A.3: AlphaFold prediction of the DanK structure [82; 83]

A.3 DnaJ

A.3.1 Auto-inhibition mechanism

Table A.11 collect all the parameters estimated to fit the data presented in Figure 4.15 and the raw data are displayed in Tables A.12 and A.13.

Table A.11: All parameters obtained from fitting the curve in Figure 4.15

k_Y^{on}	k_S^{on}	K_Y^d	K_S^d	K_j^Y	K_j^S	$k_J^{h,Y}$	$k_J^{h,S}$	$\hat{k}$
11.8 $\mu M^{-1}s^{-1}$	7.7 $\mu M^{-1}s^{-1}$	23 μM	143 μM	0.033	0.521	4.6 s^{-1}	16.7 s^{-1}	0.026 s^{-1}

JDP [μM]	YY1 [s^{-1}]	YY2 [s^{-1}]	YY3 [s^{-1}]	SS1 [s^{-1}]	SS2 [s^{-1}]	SS3 [s^{-1}]
0.0	0.002453	0.002982	0.003284	0.002453	0.002982	0.003284
0.2	0.004491	0.004567	0.004793	0.010379	0.011738	0.009171
0.4	0.005020	0.005171	0.004793	0.015134	0.015285	0.015663
0.8	0.007284	0.007888	0.007133	0.018758	0.018682	0.016493
1.6	0.010152	0.010907	0.010756	0.020720	0.021324	0.021324
3.2	0.014002	0.014681	0.014983	0.023211	0.023890	0.024041

Table A.12: Raw data from the WT ATPase experiments

JDP [μM]	SY1 [s^{-1}]	SY2 [s^{-1}]	SY3 [s^{-1}]	YS1 [s^{-1}]	YS2 [s^{-1}]	YS3 [s^{-1}]
0.0	0.002453	0.002982	0.003284	0.002453	0.002982	0.003284
0.2	0.014380	0.014908	0.015512	0.018456	0.019210	0.018833
0.4	0.020418	0.019814	0.020720	0.022834	0.024268	0.024570
0.8	0.024419	0.024041	0.025400	0.027287	0.027136	0.027061
1.6	0.028193	0.028042	0.030457	0.028646	0.028872	0.029401
3.2	0.031665	0.030835	0.033099	0.030382	0.030608	0.031363

Table A.13: Raw data from the chimeras ATPase experiments

A.3.2 The full Hsp70 cycle

When simulating the full Hsp70 cycle (Figures 4.17 and 4.20), we consider the following elements in the cycle: a substrate S, Hsp70 either in its ATP state, marked as T or in ADP-state, D and DnaJ, J. Those proteins can interact together to form dimers: JS, TS, TJ, DS and DJ. For each Hsp70, four different trimers can be formed: SHJ, SJH, HSJ and $\boxed{HJS}$, $H = T$ or D, such that: SHJ corresponds to the 2-bond trimer with H bound to both S and J, SJH is the 2-bond trimer with J bound to S and H and HSJ, S is bound to H and J. Finally, $\boxed{HJS}$ corresponds to the fully connected trimer with all proteins interacting with each other.

The values of reaction rates used in the simulation are given in Tables A.14 and A.15.

Elements	$H_T - S$ [64; 73]	$H_T - J$[74]	$H_D - S$ [74]	$H_D - J$	$J - S$ [74]
On-rate [$\mu M^{-1} s^{-1}$]	0.45	2.3×10^{-2}	10^{-3}	2.3×10^{-3}	0.33
Off-rate [s^{-1}]	2	1.6×10^{-3}	4.7×10^{-4}	1.15	6.2×10^{-3}

Table A.14: Binding and unbinding rates between the three proteins

k^{T-}	$1.33 \times 10^{-4}\ s^{-1}$	k^h	$0.0006\ s^{-1}$
k^{T+}	$0.13\ M^{-1} s^{-1}$	k^h_S	$0.007\ s^{-1}$
k^{D-}	$0.022\ s^{-1}$	k^h_J	$0.79\ s^{-1}$
k^{D+}	$0.267\ M^{-1} s^{-1}$	k^h_{JS}	$1.8\ s^{-1}$

Table A.15: References for the nucleotide exchange rate [97]and for the hydrolysis rate [28; 27]

Additionally, we define some useful constants:

$$C_S = \frac{k^{\text{on}}_{SD} k^{\text{off}}_{ST}}{k^{\text{off}}_{SD} k^{\text{on}}_{ST}} \qquad C_J = \frac{k^{\text{on}}_{DJ} k^{\text{off}}_{TJ}}{k^{\text{off}}_{DJ} k^{\text{on}}_{TJ}} \tag{A.2}$$

$$\rho = 1000\mu\text{M} \qquad \alpha_{eq} = 10^{-6} \qquad \alpha = 10 \tag{A.3}$$

Respecting the detailed balance condition, the nucleotide binding and unbinding rates when

Hsp70 is bound to S and/or J are given by:

$$\begin{aligned}
k_S^{T-} &= k^{T-} C_S^{1/4}, & k_J^{T-} &= k^{T-} C_J^{1/4}, & k_{JS}^{T-} &= k^{T-} C_S^{1/4} C_J^{1/4} \\
k_S^{T+} &= k^{T+} C_S^{-1/4}, & k_J^{T+} &= k^{T+} C_J^{-1/4}, & k_{JS}^{T+} &= k^{T+} C_S^{-1/4} C_J^{-1/4} \\
k_S^{D-} &= k^{D-} C_S^{-1/4}, & k_J^{D-} &= k^{D-} C_J^{-1/4}, & k_{JS}^{D-} &= k^{D-} C_S^{-1/4} C_J^{-1/4} \\
k_S^{D+} &= k^{D+} C_S^{1/4}, & k_J^{D+} &= k^{D+} C_J^{1/4}, & k_{JS}^{D+} &= k^{D+} C_S^{1/4} C_J^{1/4}
\end{aligned} \tag{A.4}$$

$$\begin{aligned}
k^{DT} &= \frac{k^{D-} k^{T+} \alpha}{k^{D+} + \alpha k^{T+}}, & k^{TD} &= \frac{k^{T-} k^{D+}}{k^{D+} + \alpha k^{T+}} \\
k_X^{DT} &= \frac{k_X^{D-} k_X^{T+} \alpha}{k_X^{D+} + \alpha k_X^{T+}}, & k_X^{TD} &= \frac{k_X^{T-} k_X^{D+}}{k_X^{D+} + \alpha k_X^{T+}} \\
k_{JS}^{DT} &= \frac{k_{JS}^{D-} k_{JS}^{T+} \alpha}{k_{JS}^{D+} + \alpha k_{JS}^{T+}}, & k_{JS}^{TD} &= \frac{k_{JS}^{T-} k_{JS}^{D+}}{k_{JS}^{D+} + \alpha k_{JS}^{T+}}
\end{aligned} \tag{A.5}$$

$$k^s = k^h \frac{k^{D-} k^{T+} \alpha_{eq}}{k^{D+} k^{T-}}, \quad k_X^s = k_X^h \frac{k^{D-} k^{T+} \alpha_{eq}}{k^{D+} k^{T-}} C_X^{-1}, \quad k_{JS}^s = k_{JS}^h \frac{k^{D-} k^{T+} \alpha_{eq}}{k^{D+} k^{T-}} C_S^{-1} C_J^{-1}. \tag{A.6}$$

All the equations characterizing the Hsp70 cycle interactions are presented below with the rates associated to the reactions.
Equations of dimers binding and unbinding:

$$\begin{aligned}
J + S &\to JS \quad (k_{SJ}^{\text{on}}), & JS &\to J + S \quad (k_{SJ}^{\text{off}}), \\
T + J &\to TJ \quad (k_{TJ}^{\text{on}}), & TJ &\to T + J \quad (k_{TJ}^{\text{off}}), \\
T + S &\to ST \quad (k_{ST}^{\text{on}}), & ST &\to T + S \quad (k_{ST}^{\text{off}}), \\
D + J &\to DJ \quad (k_{DJ}^{\text{on}}), & DJ &\to D + J \quad (k_{DJ}^{\text{off}}), \\
D + S &\to SD \quad (k_{SD}^{\text{on}}), & SD &\to D + S \quad (k_{SD}^{\text{off}}).
\end{aligned} \tag{A.7}$$

Equations for the trimers formation with the energetic cost associated to the steric effect:

$$\begin{aligned}
SJ+T &\rightarrow SJT \quad (k_{TJ}^{\text{on}} \times e^{-\beta\epsilon}), & SJT &\rightarrow SJ+T \quad (k_{TJ}^{\text{off}}),\\
SJ+T &\rightarrow TSJ \quad (k_{ST}^{\text{on}} \times e^{-\beta\epsilon}), & TSJ &\rightarrow SJ+T \quad (k_{ST}^{\text{off}}),\\
SJ+D &\rightarrow SJD \quad (k_{DJ}^{\text{on}} \times e^{-\beta\epsilon}), & SJD &\rightarrow SJ+D \quad (k_{DJ}^{\text{off}}),\\
SJ+D &\rightarrow DSJ \quad (k_{SD}^{\text{on}} \times e^{-\beta\epsilon}), & DSJ &\rightarrow SJ+D \quad (k_{SD}^{\text{off}}),\\
SD+J &\rightarrow SDJ \quad (k_{DJ}^{\text{on}} \times e^{-\beta\epsilon}), & SDJ &\rightarrow SD+J \quad (k_{DJ}^{\text{off}}),\\
SD+J &\rightarrow DSJ \quad (k_{SJ}^{\text{on}} \times e^{-\beta\epsilon}), & DSJ &\rightarrow SD+J \quad (k_{SJ}^{\text{off}}),\\
ST+J &\rightarrow STJ \quad (k_{TJ}^{\text{on}} \times e^{-\beta\epsilon}), & STJ &\rightarrow ST+J \quad (k_{TJ}^{\text{off}}),\\
ST+J &\rightarrow TSJ \quad (k_{SJ}^{\text{on}} \times e^{-\beta\epsilon}), & TSJ &\rightarrow ST+J \quad (k_{SJ}^{\text{off}}),\\
DJ+S &\rightarrow SDJ \quad (k_{SD}^{\text{on}} \times e^{-\beta\epsilon}), & SDJ &\rightarrow DJ+S \quad (k_{SD}^{\text{off}}),\\
DJ+S &\rightarrow SJD \quad (k_{SJ}^{\text{on}} \times e^{-\beta\epsilon}), & SJD &\rightarrow DJ+S \quad (k_{SJ}^{\text{off}}),\\
TJ+S &\rightarrow STJ \quad (k_{ST}^{\text{on}} \times e^{-\beta\epsilon}), & STJ &\rightarrow TJ+S \quad (k_{ST}^{\text{off}}),\\
TJ+S &\rightarrow SJT \quad (k_{SJ}^{\text{on}} \times e^{-\beta\epsilon}), & SJT &\rightarrow TJ+S \quad (k_{SJ}^{\text{off}}).
\end{aligned} \tag{A.8}$$

Formation of the fully connected trimer promoted by the apparent high concentration of the last not-bound proteins:

$$\begin{aligned}
STJ &\rightarrow \boxed{TJS} \quad (k_{SJ}^{\text{on}} \times \rho), & \boxed{TJS} &\rightarrow STJ \quad (k_{SJ}^{\text{off}}),\\
SDJ &\rightarrow \boxed{DJS} \quad (k_{SJ}^{\text{on}} \times \rho), & \boxed{DJS} &\rightarrow SDJ \quad (k_{SJ}^{\text{off}}),\\
SJT &\rightarrow \boxed{TJS} \quad (k_{ST}^{\text{on}} \times \rho), & \boxed{TJS} &\rightarrow SJT \quad (k_{ST}^{\text{off}}),\\
SJD &\rightarrow \boxed{DJS} \quad (k_{SD}^{\text{on}} \times \rho), & \boxed{DJS} &\rightarrow SJD \quad (k_{SD}^{\text{off}}),\\
TSJ &\rightarrow \boxed{TJS} \quad (k_{TJ}^{\text{on}} \times \rho), & \boxed{TJS} &\rightarrow TSJ \quad (k_{TJ}^{\text{off}}),\\
DSJ &\rightarrow \boxed{DJS} \quad (k_{DJ}^{\text{on}} \times \rho), & \boxed{DJS} &\rightarrow DSJ \quad (k_{DJ}^{\text{off}}).
\end{aligned} \tag{A.9}$$

Transitions between the ATP to ADP states through nucleotide exchange and hydrolysis /synthesis:

$$\begin{aligned}
T &\rightarrow D, \quad (k^{TD}+k^{h}) & D &\rightarrow T, \quad (k^{DT}+k^{s})\\
ST &\rightarrow SD, \quad (k_S^{TD}+k_S^{h}) & SD &\rightarrow ST, \quad (k_S^{DT}+k_S^{s})\\
TJ &\rightarrow DJ, \quad (k_J^{TD}+k_J^{h}) & DJ &\rightarrow TJ, \quad (k_J^{DT}+k_J^{s})\\
STJ &\rightarrow SDJ, \quad (k_{JS}^{TD}+k_{JS}^{h}) & SDJ &\rightarrow STJ, \quad (k_{JS}^{DT}+k_{JS}^{s})\\
SJT &\rightarrow SJD, \quad (k_J^{TD}+k_J^{h}) & SJD &\rightarrow SJT, \quad (k_S^{DT}+k_S^{s})\\
TSJ &\rightarrow DSJ, \quad (k_S^{TD}+k_S^{h}) & DSJ &\rightarrow TSJ, \quad (k_S^{DT}+k_S^{s})\\
\boxed{TJS} &\rightarrow \boxed{DJS}, \quad (k_{JS}^{TD}+k_{JS}^{h}) & \boxed{DJS} &\rightarrow \boxed{TJS}, \quad (k_{JS}^{DT}+k_{JS}^{s})
\end{aligned} \tag{A.10}$$

Finally, in Figures 4.22 and 4.23, DnaJ regulatory mechanism was also considered. Therefore

those additional transitions were added to the cycle:

$$\begin{aligned} J &\rightarrow J_- \quad (k^{+\rightarrow -}) & \qquad J_- &\rightarrow J \quad (k^{-\rightarrow +}) \\ SJ &\rightarrow SJ_- \quad (k_S^{+\rightarrow -}) & \qquad SJ_- &\rightarrow SJ \quad (k_S^{-\rightarrow +}) \end{aligned} \tag{A.11}$$

And the reaction rates are detailed in Table A.16. The value of $K^d_{J_-,S}$, the dissociation constant between J_- and S, was considered in Figure 4.22 as equal to the dissociation constant between J and S, such that $K^d_{J_-,S} = K^d_{J,S} = 0.019\mu$M and as very high to prevent any binding in Figure 4.23, such that $K^d_{J_-,S} = 10^6\mu$M.

$k^{+\rightarrow -}$	$k^{-\rightarrow +}$	$k_S^{+\rightarrow -}$	$k_S^{-\rightarrow +}$
$1\ \text{s}^{-1}$	$K_J k^{+\rightarrow -}$	$1\ \text{s}^{-1}$	$K_J \dfrac{K^d_{J_-,S}}{K^d_{J,S}} k_S^{+\rightarrow -}$

Table A.16: Added rates when considering the auto-regulation mechanism on DnaJ. $K^d_{J_-,S} = 10^6\mu$M.

Bibliography

[1] David Hartich, Andre C. Barato, and Udo Seifert. Nonequilibrium sensing and its analogy to kinetic proofreading. *New Journal of Physics*, 17(5):055026, 2015.

[2] Grégory Boël, Olivier Danot, Victor de Lorenzo, and Antoine Danchin. Omnipresent Maxwell's demons orchestrate information management in living cells. *Microbial Biotechnology*, 12(2):210–242, 2019.

[3] Solange Flatt, Daniel M. Busiello, Stefano Zamuner, and Paolo De Los Rios. ABC Transporters are billion-year-old Maxwell Demons. *bioRxiv*, 2021.

[4] Andrija Finka and Pierre Goloubinoff. Proteomic data from human cell cultures refine mechanisms of chaperone-mediated protein homeostasis. *Cell Stress and Chaperones*, 18(5):591–605, September 2013.

[5] Jens Tyedmers, Axel Mogk, and Bernd Bukau. Cellular strategies for controlling protein aggregation. *Nature Reviews Molecular Cell Biology*, 11(11):777–788, November 2010. Number: 11 Publisher: Nature Publishing Group.

[6] G. Merlini, V. Bellotti, A. Andreola, G. Palladini, L. Obici, S. Casarini, and V. Perfetti. Protein aggregation. *Clinical Chemistry and Laboratory Medicine*, 39(11):1065–1075, November 2001.

[7] Christopher J. Roberts. Non-native protein aggregation kinetics. *Biotechnology and Bioengineering*, 98(5):927–938, 2007. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bit.21627.

[8] Jevgenij A. Raskatov and David B. Teplow. Using chirality to probe the conformational dynamics and assembly of intrinsically disordered amyloid proteins. *Scientific Reports*, 7(1):12433, October 2017. Number: 1 Publisher: Nature Publishing Group.

[9] J. P. Hendrick and F. U. Hartl. The role of molecular chaperones in protein folding. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 9(15):1559–1569, December 1995.

[10] Paolo De Los Rios, Anat Ben-Zvi, Olga Slutsky, Abdussalam Azem, and Pierre Goloubinoff. Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proceedings of the National Academy of Sciences*,

103(16):6166–6171, April 2006. Publisher: Proceedings of the National Academy of Sciences.

[11] David Balchin, Manajit Hayer-Hartl, and F. Ulrich Hartl. In vivo aspects of protein folding and quality control. *Science*, 353(6294):aac4354, July 2016. Publisher: American Association for the Advancement of Science.

[12] F. Ulrich Hartl. Molecular chaperones in cellular protein folding. *Nature*, 381(6583):571–580, June 1996. Number: 6583 Publisher: Nature Publishing Group.

[13] Pierre Goloubinoff and Paolo De Los Rios. The mechanism of Hsp70 chaperones: (entropic) pulling the models together. *Trends in Biochemical Sciences*, 32(8):372–380, August 2007.

[14] Xuechao Gao, Marta Carroni, Carmen Nussbaum-Krammer, Axel Mogk, Nadinath B. Nillegoda, Anna Szlachcic, D. Lys Guilbride, Helen R. Saibil, Matthias P. Mayer, and Bernd Bukau. Human Hsp70 Disaggregase Reverses Parkinson's-Linked alpha-Synuclein Amyloid Fibrils. *Molecular Cell*, 59(5):781–793, September 2015.

[15] Anne S. Wentink, Nadinath B. Nillegoda, Jennifer Feufel, Gabrielė Ubartaitė, Carolyn P. Schneider, Paolo De Los Rios, Janosch Hennig, Alessandro Barducci, and Bernd Bukau. Molecular dissection of amyloid disaggregation by human HSP70. *Nature*, 587(7834):483–488, November 2020. Number: 7834 Publisher: Nature Publishing Group.

[16] Ana Maria Cuervo, Esther S. P. Wong, and Marta Martinez-Vicente. Protein degradation, aggregation, and misfolding. *Movement Disorders: Official Journal of the Movement Disorder Society*, 25 Suppl 1:S49–54, 2010.

[17] Abhisek Mukherjee, Diego Morales-Scheihing, Peter C. Butler, and Claudio Soto. Type 2 diabetes as a protein misfolding disease. *Trends in Molecular Medicine*, 21(7):439–449, July 2015.

[18] Christopher A. Ross and Michelle A. Poirier. Protein aggregation and neurodegenerative disease. *Nature Medicine*, 10(7):S10–S17, July 2004. Number: 7 Publisher: Nature Publishing Group.

[19] Justin M. Long and David M. Holtzman. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell*, 179(2):312–339, October 2019.

[20] John A. Hardy and Gerald A. Higgins. Alzheimer's Disease: The Amyloid Cascade Hypothesis. *Science*, 256(5054):184–185, April 1992. Publisher: American Association for the Advancement of Science.

[21] DailyMed - LEQEMBI- lecanemab injection, solution.

[22] Chad J. Swanson, Yong Zhang, Shobha Dhadda, Jinping Wang, June Kaplow, Robert Y. K. Lai, Lars Lannfelt, Heather Bradley, Martin Rabe, Akihiko Koyama, Larisa Reyderman,

Donald A. Berry, Scott Berry, Robert Gordon, Lynn D. Kramer, and Jeffrey L. Cummings. A randomized, double-blind, phase 2b proof-of-concept clinical trial in early Alzheimer's disease with lecanemab, an anti-Abeta protofibril antibody. *Alzheimer's Research & Therapy*, 13(1):80, April 2021.

[23] Mark S. Hipp, Sae-Hun Park, and F. Ulrich Hartl. Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends in Cell Biology*, 24(9):506–514, September 2014.

[24] Eugenia M. Clerico, Wenli Meng, Alexandra Pozhidaeva, Karishma Bhasne, Constantine Petridis, and Lila M. Gierasch. Hsp70 molecular chaperones: multifunctional allosteric holding and unfolding machines. *Biochemical Journal*, 476(11):1653–1677, 2019.

[25] J. J. Hopfield. Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity. *Proceedings of the National Academy of Sciences*, 71(10):4135–4139, 1974.

[26] J. Ninio. Kinetic amplification of enzyme discrimination. *Biochimie*, 57(5):587–595, 1975.

[27] John S. McCarty, Alexander Buchberger, Jochen Reinstein, and Bernd Bukau. The Role of ATP in the Functional Cycle of the DnaK Chaperone System. *Journal of Molecular Biology*, 249(1):126–137, 1995.

[28] Thomas Laufen, Matthias P. Mayer, Christian Beisel, Dagmar Klostermeier, Axel Mogk, Jochen Reinstein, and Bernd Bukau. Mechanism of regulation of Hsp70 chaperones by DnaJ cochaperones. *Proceedings of the National Academy of Sciences*, 96(10):5452–5457, 1999. Publisher: National Academy of Sciences Section: Biological Sciences.

[29] Wanjiang Han and Philipp Christen. cis-Effect of DnaJ on DnaK in ternary complexes with chimeric DnaK/DnaJ-binding peptides. *FEBS Letters*, 563(1-3):146–150, 2004. Publisher: John Wiley & Sons, Ltd.

[30] Xue-Juan Zhang, Hong Qian, and Min Qian. Stochastic theory of nonequilibrium steady states and its applications. Part I. *Physics Reports*, 510(1):1–86, January 2012.

[31] Long-Qing Chen. Chemical potential and Gibbs free energy. *MRS Bulletin*, 44(7):520–523, July 2019.

[32] Kenneth A. Johnson and Roger S. Goody. The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten Paper. *Biochemistry*, 50(39):8264–8269, October 2011. Publisher: American Chemical Society.

[33] Ron Milo and Rob Phillips. *Cell biology by the numbers*. Garland Science, Taylor & Francis Group, New York, NY, 2016.

[34] Pablo Sartori and Simone Pigolotti. Kinetic versus Energetic Discrimination in Biological Copying. *Physical Review Letters*, 110(18):188101, 2013.

[35] Terrell L. Hill. *Free Energy Transduction and Biochemical Cycle Kinetics*. Springer, New York, NY, 1989.

[36] Salvatore Assenza, Alberto Stefano Sassi, Ruth Kellner, Benjamin Schuler, Paolo De Los Rios, and Alessandro Barducci. Efficient conversion of chemical energy into mechanical work by Hsp70 chaperones. *eLife*, 8:e48491, 2019. Publisher: eLife Sciences Publications, Ltd.

[37] Adélaïde A. Mohr, Daniel M. Busiello, Stefano Zamuner, and Paolo De Los Rios. The individual and combined benefits of different non-equilibrium proofreading mechanisms, July 2022.

[38] Hani S. Zaher and Rachel Green. Fidelity at the molecular level: lessons from protein synthesis. *Cell*, 136(4):746–762, 2009.

[39] Kenneth A Johnson. Conformational coupling in DNA polymerase fidelity. *Annual Review of Biochemistry*, 62(1):685–713, 1993.

[40] Jeong Woong Lee, Kirk Beebe, Leslie A. Nangle, Jaeseon Jang, Chantal M. Longo-Guess, Susan A. Cook, Muriel T. Davisson, John P. Sundberg, Paul Schimmel, and Susan L. Ackerman. Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature*, 443(7107):50–55, September 2006.

[41] Arvind Murugan, David A. Huse, and Stanislas Leibler. Speed, dissipation, and error in kinetic proofreading. *Proceedings of the National Academy of Sciences*, 109(30):12034–12039, 2012.

[42] Kinshuk Banerjee, Anatoly B. Kolomeisky, and Oleg A. Igoshin. Elucidating interplay of speed and accuracy in biological error correction. *Proceedings of the National Academy of Sciences*, 114(20):5183–5188, 2017.

[43] Joel D. Mallory, Anatoly B. Kolomeisky, and Oleg A. Igoshin. Trade-Offs between Error, Speed, Noise, and Energy Dissipation in Biological Processes with Proofreading. *The Journal of Physical Chemistry B*, 123(22):4718–4725, 2019.

[44] Charles H. Bennett. Dissipation-error tradeoff in proofreading. *Biosystems*, 11(2):85–91, 1979.

[45] Riccardo Rao and Luca Peliti. Thermodynamics of accuracy in kinetic proofreading: dissipation and efficiency trade-offs. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(6):P06001, 2015.

[46] Qiwei Yu, Anatoly B. Kolomeisky, and Oleg A. Igoshin. The energy cost and optimal design of networks for biological discrimination. *Journal of The Royal Society Interface*, 19(188):20210883, January 2022. Publisher: Royal Society.

[47] Ganhui Lan, Pablo Sartori, Silke Neumann, Victor Sourjik, and Yuhai Tu. The energy–speed–accuracy trade-off in sensory adaptation. *Nature Physics*, 8(5):422–428, 2012.

[48] Magnus Johansson, Martin Lovmar, and Måns Ehrenberg. Rate and accuracy of bacterial protein synthesis revisited. *Current Opinion in Microbiology*, 11(2):141–147, April 2008.

[49] Ingo Wohlgemuth, Corinna Pohl, and Marina V Rodnina. Optimization of speed and accuracy of decoding in translation. *The EMBO Journal*, 29(21):3701–3709, 2010.

[50] Davide Chiuchiú, Yuhai Tu, and Simone Pigolotti. Error-Speed Correlations in Biopolymer Synthesis. *Physical Review Letters*, 123(3):038101, 2019.

[51] D. E. Jr. Koshland, G. Némethy, and D. Filmer. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits. *Biochemistry*, 5(1):365–385, January 1966. Publisher: American Chemical Society.

[52] Daniel E. Koshland Jr. The Key–Lock Theory and the Induced Fit Theory. *Angewandte Chemie International Edition in English*, 33(23-24):2375–2378, 1995.

[53] Jacques Monod, Jean-Pierre Changeux, and François Jacob. Allosteric proteins and cellular control systems. *Journal of Molecular Biology*, 6(4):306–329, April 1963.

[54] Chung-Jung Tsai and Ruth Nussinov. A Unified View of "How Allostery Works". *PLOS Computational Biology*, 10(2):e1003394, 2014. Publisher: Public Library of Science.

[55] Scott C. Blanchard, Ruben L. Gonzalez, Harold D. Kim, Steven Chu, and Joseph D. Puglisi. tRNA selection and kinetic proofreading in translation. *Nature Structural & Molecular Biology*, 11(10):1008–1014, 2004.

[56] Tillmann Pape, Wolfgang Wintermeyer, and Marina Rodnina. Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome. *The EMBO Journal*, 18(13):3800–3807, 1999.

[57] Kirill B. Gromadski and Marina V. Rodnina. Kinetic Determinants of High-Fidelity tRNA Discrimination on the Ribosome. *Molecular Cell*, 13(2):191–200, 2004.

[58] Chithra Hariharan, Linda B. Bloom, Sandra A. Helquist, Eric T. Kool, and Linda J. Reha-Krantz. Dynamics of Nucleotide Incorporation: Snapshots Revealed by 2-Aminopurine Fluorescence Studies. *Biochemistry*, 45(9):2836–2844, 2006. Publisher: American Chemical Society.

[59] Kenneth A. Johnson. Role of induced fit in enzyme specificity: a molecular forward/reverse switch. *The Journal of Biological Chemistry*, 283(39):26297–26301, September 2008.

[60] Roman Kityk, Jürgen Kopp, Irmgard Sinning, and Matthias P. Mayer. Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. *Molecular Cell*, 48(6):863–874, December 2012. Publisher: Elsevier.

[61] Eric B. Bertelsen, Lyra Chang, Jason E. Gestwicki, and Erik R. P. Zuiderweg. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proceedings of the National Academy of Sciences*, 106(21):8471–8476, May 2009. Publisher: Proceedings of the National Academy of Sciences.

[62] Anastasia Zhuravleva, Eugenia M. Clerico, and Lila M. Gierasch. An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. *Cell*, 151(6):1296–1307, 2012.

[63] Joanna F. Swain, Gizem Dinler, Renuka Sivendran, Diana L. Montgomery, Mathias Stotz, and Lila M. Gierasch. Hsp70 Chaperone Ligands Control Domain Association via an Allosteric Mechanism Mediated by the Interdomain Linker. *Molecular Cell*, 26(1):27–39, 2007.

[64] Daniel Schmid, Antonio Baici, Heinz Gehring, and Philipp Christen. Kinetics of Molecular Chaperone Action. *Science*, 263(5149):971–973, 1994. Publisher: American Association for the Advancement of Science.

[65] Matthias P. Mayer, Hartwig Schröder, Stefan Rüdiger, Klaus Paal, Thomas Laufen, and Bernd Bukau. Multistep mechanism of substrate binding determines chaperone activity of Hsp70. *Nature Structural Biology*, 7(7):586–593, 2000. Number: 7 Publisher: Nature Publishing Group.

[66] Claudia S. Gässler, Alexander Buchberger, Thomas Laufen, Matthias P. Mayer, Hartwig Schröder, Alfonso Valencia, and Bernd Bukau. Mutations in the DnaK chaperone affecting interaction with the DnaJ cochaperone. *Proceedings of the National Academy of Sciences*, 95(26):15229–15234, 1998. Publisher: National Academy of Sciences Section: Biological Sciences.

[67] Matthias P. Mayer, Thomas Laufen, Klaus Paal, John S. McCarty, and Bernd Bukau. Investigation of the Interaction between DnaK and DnaJ by Surface Plasmon Resonance Spectroscopy. *Journal of Molecular Biology*, 289(4):1131–1144, 1999.

[68] Rainer Schlecht, Annette H. Erbse, Bernd Bukau, and Matthias P. Mayer. Mechanics of Hsp70 chaperones enables differential interaction with client proteins. *Nature Structural & Molecular Biology*, 18(3):345–351, 2011.

[69] Wei Wang, Qinglian Liu, Qun Liu, and Wayne A. Hendrickson. Conformational equilibria in allosteric control of Hsp70 chaperones. *Molecular Cell*, page S1097276521006237, 2021.

[70] Roman Kityk, Jürgen Kopp, and Matthias P. Mayer. Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. *Molecular Cell*, 69(2):227–237.e4, 2018. Publisher: Elsevier.

[71] Wayne A. Hendrickson. Theory of Allosteric Regulation in Hsp70 Molecular Chaperones. *QRB Discovery*, 1:e7, 2020.

[72] Mateusz Dyla, Nicolás S. González Foutel, Daniel E. Otzen, and Magnus Kjaergaard. The optimal docking strength for reversibly tethered kinases. *Proceedings of the National Academy of Sciences*, 119(25):e2203098119, June 2022. Company: National Academy of Sciences Distributor: National Academy of Sciences Institution: National Academy of Sciences Label: National Academy of Sciences Publisher: Proceedings of the National Academy of Sciences.

[73] Serge M. Gisler, Ezra V. Pierpaoli, and Philipp Christen. Catapult mechanism renders the chaperone action of hsp70 unidirectional11Edited by A. R. Fersht. *Journal of Molecular Biology*, 279(4):833–840, June 1998.

[74] Won-Chul Suh, William F. Burkholder, Chi Zen Lu, Xun Zhao, Max E. Gottesman, and Carol A. Gross. Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. *Proceedings of the National Academy of Sciences*, 95(26):15223–15228, 1998. Publisher: National Academy of Sciences Section: Biological Sciences.

[75] A. Wali Karzai and Roger McMacken. A Bipartite Signaling Mechanism Involved in DnaJ-mediated Activation of the Escherichia coli DnaK Protein (). *Journal of Biological Chemistry*, 271(19):11236–11246, May 1996.

[76] Harm H. Kampinga and Elizabeth A. Craig. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology*, 11(8):579–592, 2010. Number: 8 Publisher: Nature Publishing Group.

[77] Jonathan J. Silberg and Larry E. Vickery. Kinetic Characterization of the ATPase Cycle of the Molecular Chaperone Hsc66 from Escherichia coli. *Journal of Biological Chemistry*, 275(11):7779–7786, 2000.

[78] Jonathan J. Silberg, Tim L. Tapley, Kevin G. Hoff, and Larry E. Vickery. Regulation of the HscA ATPase Reaction Cycle by the Co-chaperone HscB and the Iron-Sulfur Cluster Assembly Protein IscU. *Journal of Biological Chemistry*, 279(52):53924–53931, 2004. Publisher: American Society for Biochemistry and Molecular Biology.

[79] Larry E. Vickery and Jill R. Cupp-Vickery. Molecular Chaperones HscA/Ssq1 and HscB/-Jac1 and Their Roles in Iron-Sulfur Protein Maturation. *Critical Reviews in Biochemistry and Molecular Biology*, 42(2):95–111, 2007. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10409230701322298.

[80] Kevin G. Hoff, Jonathan J. Silberg, and Larry E. Vickery. Interaction of the iron–sulfur cluster assembly protein IscU with the Hsc66/Hsc20 molecular chaperone system of

Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 97(14):7790–7795, 2000.

[81] Wei Wang and Wayne A Hendrickson. Intermediates in allosteric equilibria of DnaK–ATP interactions with substrate peptides. *Acta Crystallographica Section D*, 77(5):606–617, 2021.

[82] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, August 2021.

[83] Mihaly Varadi, Stephen Anyango, Mandar Deshpande, Sreenath Nair, Cindy Natassia, Galabina Yordanova, David Yuan, Oana Stroe, Gemma Wood, Agata Laydon, Augustin Žídek, Tim Green, Kathryn Tunyasuvunakool, Stig Petersen, John Jumper, Ellen Clancy, Richard Green, Ankur Vora, Mira Lutfi, Michael Figurnov, Andrew Cowie, Nicole Hobbs, Pushmeet Kohli, Gerard Kleywegt, Ewan Birney, Demis Hassabis, and Sameer Velankar. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50(D1):D439–D444, January 2022.

[84] Elizabeth A. Craig and Jaroslaw Marszalek. How Do J-Proteins Get Hsp70 to Do So Many Different Things? *Trends in Biochemical Sciences*, 42(5):355–368, May 2017.

[85] Harm H. Kampinga, Claes Andreasson, Alessandro Barducci, Michael E. Cheetham, Douglas Cyr, Cecilia Emanuelsson, Pierre Genevaux, Jason E. Gestwicki, Pierre Goloubinoff, Jaime Huerta-Cepas, Janine Kirstein, Krzysztof Liberek, Matthias P. Mayer, Kazuhiro Nagata, Nadinath B. Nillegoda, Pablo Pulido, Carlos Ramos, Paolo De los Rios, Sabine Rospert, Rina Rosenzweig, Chandan Sahi, Mikko Taipale, Bratłomiej Tomiczek, Ryo Ushioda, Jason C. Young, Richard Zimmermann, Alicja Zylicz, Maciej Zylicz, Elizabeth A. Craig, and Jaroslaw Marszalek. Function, evolution, and structure of J-domain proteins. *Cell Stress and Chaperones*, 24(1):7–15, January 2019.

[86] Douglas M. Cyr, Thomas Langer, and Michael G. Douglas. DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. *Trends in Biochemical Sciences*, 19(4):176–181, April 1994.

[87] Ofrah Faust, Meital Abayev-Avraham, Anne S. Wentink, Michael Maurer, Nadinath B. Nillegoda, Nir London, Bernd Bukau, and Rina Rosenzweig. HSP40 proteins use class-specific regulation to drive HSP70 functional diversity. *Nature*, 587(7834):489–494, November 2020. Number: 7834 Publisher: Nature Publishing Group.

[88] J. F. Morrison. Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochimica et Biophysica Acta (BBA) - Enzymology*, 185(2):269–286, August 1969.

[89] Zhen Lu and Douglas M. Cyr. Protein Folding Activity of Hsp70 Is Modified Differentially by the Hsp40 Co-chaperones Sis1 and Ydj1. *Journal of Biological Chemistry*, 273(43):27824–27830, October 1998. Publisher: Elsevier.

[90] Chun-Yang Fan, Soojin Lee, Hong-Yu Ren, and Douglas M. Cyr. Exchangeable Chaperone Modules Contribute to Specification of Type I and Type II Hsp40 Cellular Function. *Molecular Biology of the Cell*, 15(2):761–773, February 2004.

[91] Nadinath B. Nillegoda, Janine Kirstein, Anna Szlachcic, Mykhaylo Berynskyy, Antonia Stank, Florian Stengel, Kristin Arnsburg, Xuechao Gao, Annika Scior, Ruedi Aebersold, D. Lys Guilbride, Rebecca C. Wade, Richard I. Morimoto, Matthias P. Mayer, and Bernd Bukau. Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature*, 524(7564):247–251, August 2015.

[92] Júlio C. Borges, Thiago V. Seraphim, David Z. Mokry, Fabio C. L. Almeida, Douglas M. Cyr, and Carlos H. I. Ramos. Identification of Regions Involved in Substrate Binding and Dimer Stabilization within the Central Domains of Yeast Hsp40 Sis1. *PLoS ONE*, 7(12):e50927, December 2012.

[93] S. Rüdiger, J. Schneider-Mergener, and B. Bukau. Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *The EMBO journal*, 20(5):1042–1050, March 2001.

[94] Soojin Lee, Chun Yang Fan, J. Michael Younger, Hongyu Ren, and Douglas M. Cyr. Identification of essential residues in the type II Hsp40 Sis1 that function in polypeptide binding. *The Journal of Biological Chemistry*, 277(24):21675–21682, June 2002.

[95] Jingzhi Li, Xinguo Qian, and Bingdong Sha. The crystal structure of the yeast Hsp40 Ydj1 complexed with its peptide substrate. *Structure (London, England: 1993)*, 11(12):1475–1483, December 2003.

[96] Bin Hu, Matthias P. Mayer, and Masaru Tomita. Modeling Hsp70-Mediated Protein Folding. *Biophysical Journal*, 91(2):496–507, 2006.

[97] Rick Russell, Robert Jordan, and Roger McMacken. Kinetic Characterization of the ATPase Cycle of the DnaK Molecular Chaperone. *Biochemistry*, 37(2):596–607, January 1998. Publisher: American Chemical Society.

[98] Verena Kohler and Claes Andréasson. Hsp70-mediated quality control: should I stay or should I go? *Biological Chemistry*, 401(11):1233–1248, October 2020. Publisher: De Gruyter.

[99] S. Diamant and P. Goloubinoff. Temperature-controlled activity of DnaK-DnaJ-GrpE chaperones: protein-folding arrest and recovery during and after heat shock depends on the substrate protein and the GrpE concentration. *Biochemistry*, 37(27):9688–9694, July 1998.

[100] J. P. Grimshaw, I. Jelesarov, H. J. Schönfeld, and P. Christen. Reversible thermal transition in GrpE, the nucleotide exchange factor of the DnaK heat-shock system. *The Journal of Biological Chemistry*, 276(9):6098–6104, March 2001.

[101] Eugene F. Jr. Douglass, Chad J. Miller, Gerson Sparer, Harold Shapiro, and David A. Spiegel. A Comprehensive Mathematical Model for Three-Body Binding Equilibria. *Journal of the American Chemical Society*, 135(16):6092–6099, April 2013. Publisher: American Chemical Society.

ADELAIDE MOHR

PhD Student in Statistical Biophysics

@ adelaide.mohr@epfl.ch Lausanne, Switzerland

ACADEMIC EXPERIENCE

Why do we need so many proteins? A physical insight into the collaboration of Hsp70 and DnaJ

PhD in Statistical Biophysics EPFL

Oct 2018 – Ongoing Lausanne

- Development of numerical simulations for the understanding of protein homeostasis (Mathematica®, Python)
- Analysis of biological data to infer biophysical model (machine learning, data analysis, Python)

Brain dysfunction and altered metabolism in diet-induced obesity: a BOLD fMRI and ^{1}H MRS study

EPFL Master thesis (performed at Lund University)

Mars 2018 – Sept 2018 Sweden

- Development of NMR experimental procedures for spectroscopy (^{1}H MRS) and functional imaging (BOLD fMRI)
- Analysis of the impact of glucose intake on diabetics mice (time series analysis, hypothesis testing, MATLAB, R)

Semester projects

EPFL

Sept 2016 – Feb 2018 Lausanne

- Assessing the impact of caffeine consumption on diabetic rodents (data analysis, modeling, statistical inference)
- Study of proteins structure and binding behaviors through multiple sequence alignment analysis (Python, statistics, molecular simulation)

Teaching Assistant in General Physics I, Analytical Mechanics and Statistical physics of biomacromolecules

EPFL

2018 – 2022 Lausanne

STRENGTHS

Communication skills | Self driven | Problem solving | Fast learner | Multidisciplinary teams

MATLAB | R | Python | Data analysis | Biostatistics | Stochastic simulation

LANGUAGES

French (Native) | English (C1/C2) | German (B1) | Italian (B1)

EDUCATION

Ph.D. in Statistical Biophysics

EPFL Oct 2018 – Ongoing

B.Sc. & M.Sc. in Physics

EPFL Sept 2012 – Sept 2018

REFEREES

Prof. Paolo de los Rios
@ EPFL
paolo.delosrios@epfl.ch

Prof. Joao Duarte
@ Lund University
joao.duarte@med.lu.se

Dr. Stefano Zamuner
@ ClearSky
stefano.zamuner@protonmail.com

PUBLICATIONS

Journal Articles

- Garcia-Serrano, A. M., Mohr, A. A., Philippe, J., Skoug, C., Spégel, P., & Duarte, J. M. (2022). Cognitive Impairment and Metabolite Profile Alterations in the Hippocampus and Cortex of Male and Female Mice Exposed to a Fat and Sugar-Rich Diet are Normalized by Diet Reversal. *Aging and Disease*, *13*(1), 267–283. doi:10.14336/AD.2021.0720
- Mohr, A. A., Busiello, D. M., Zamuner, S., & Rios, P. D. L. (2022). The individual and combined benefits of different non-equilibrium proofreading mechanisms. *bioRxiv*. doi:10.1101/2022.07.05.498791
- Mohr, A., Garcia-Serrano, A., Vieira, J., Skoug, C., Davidsson, H., & Duarte, J. (2020). A glucose-stimulated BOLD fMRI study of hypothalamic dysfunction in mice fed a high-fat and high-sucrose diet. *Journal of Cerebral Blood Flow & Metabolism*. doi:10.1177/0271678X20942397