Technische Universität München

Fakultät für Chemie

Lehrstuhl für Biotechnologie

Mechanistic and structural characterization of the molecular chaperone Spy

Philipp Koldewey

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende(r): Univ.-Prof. Dr. Ville R. I. Kaila

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Johannes Buchner
2. Univ.-Prof. Dr. Matthias J. Feige
3. Univ.-Prof. Dr. James C. A. Bardwell

Die Dissertation wurde am 07.11.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 24.11.2016 angenommen.

.

To my family and friends, who supported me throughout my Ph.D.

Summary

Understanding how chaperones work is an important question in basic biology. Given their function as guardians of protein folding and homeostasis, chaperones not only play roles in protein folding diseases such as Alzheimer's and cystic fibrosis, but, due to their involvement in signal transduction pathways, are also important in the pathogenesis of cancer. Here, the *Escherichia coli* chaperone Spy was used as a model system aimed at (1) identifying the kinetic, thermodynamic, and structural aspects that allow molecular chaperones to promote folding of its client folding, and (2) determining how these aspects affect the folding landscape of the client protein. The three publications summarized in this thesis illustrate the mechanism by which Spy facilitates the folding of its client protein Im7 in unprecedented detail. The fundamental basis of Spy's chaperoning function is the "folding-friendly" amphiphilic and flexible nature of its client binding site, which allows Im7 (the client protein used in this study) to fold while continuously bound to Spy. Binding of the unfolded state of Im7 is mediated through a combination of hydrophobic and hydrophilic interactions that allow Im7 to explore its folding landscape while bound. Despite this conformational freedom, Im7's conformational transitions are slowed down while bound to Spy, potentially providing a means to avoid the formation of non-native tertiary contacts and hence prevent misfolding. Although Spy slows Im7's folding kinetics, its folding pathway apparently remains relatively unaffected by the chaperone environment. Folding of Im7 results in the formation of a hydrophobic core and thereby reduces stabilizing hydrophobic contacts between Spy and Im7. This destabilizes the complex enough to trigger Im7 release. Such an amphiphilic binding and folding mechanism may explain how Spy can facilitate the folding of various unrelated proteins, as it doesn't require client-specific folding instructions. More importantly, rather than being dictated by the chaperone, client folding itself regulates client binding and release. Hence, Spy acts by providing a sanctuary for folding proteins that prevents aggregation and acts to slow folding kinetics, potentially to avoid misfolding. The folding pathway remains dictated by the sequence of the client protein. The electrostatic interactions formed between Spy and Im7 are a central component of this mechanism. Not only do these interactions increase the rate of client binding and therefore kinetically prevent protein aggregation, they also keep Im7 bound while it folds and hence eliminate the need for pre-native client release, often considered essential for successful client folding by chaperones. Given that the amphiphilic client binding surface exhibited by Spy is a common feature of chaperones, it is possible that this mode of self-guided folding may be the underlying mechanism for many folding chaperones.

Zusammenfassung

Wie molekulare Chaperone ihre zelluläre Funktion in der Proteinfaltung und –homöostase ausüben, ist eine der fundamentalen Fragen der Biologie. Chaperone spielen eine wichtige Rolle bei Proteinfaltungskrankheiten, wie Alzheimer oder zystischen Fibrose, und sind auf Grund ihrer Beteiligung in Signaltransduktionswegen auch involviert in der Krebspathogenese. In dieser Arbeit wird das molekulare Chaperon Spy als Modellsystem verwendet, um (1) kinetische, thermodynamische und strukturelle Prinzipien zu untersuchen, nach welchen Chaperone Proteinen bei der Faltung unterstüzten, und (2) welche Auswirkungen Chaperone auf die Faltungsenergielandschaft ihrer Substratproteine haben. Grundlage für die Chaperonfunktion von Spy ist die amphiphile, flexible Oberfläche, die es dem Substratprotein Im7 ermöglicht zu falten, während es an Spy gebunden bleibt. Ungefaltetes Im7 bindet and Spy durch eine Kombination von hydrophoben und hydrophilen Interaktionen, welche es Im7 ermöglichen Konformationsänderungen entlang der Faltungsenergielandschaft frei durchzuführen, während es an Spy gebunden bleibt. Jedoch werden Konformationsübergänge durch die Bindung an Spy kinetisch verlangsamt. Dies bietet möglicherweise einen Mechanismus, die Ausbilding von nicht-nativen, tertiären Bindungen und damit Fehlfaltung von Substratproteinen zu verhindern. Abgesehen von der verlangsamten Faltungskinetik bleibt der Faltungsweg von Im7 unverändert. Die Bildung des hydrophoben Kerns während der Faltung von Im7 führt zu einer Abnahme von stabilisierenden hydrophoben Bindungen zwischen Spy und Im7, was die Freisetzung von gefaltenen Im7 von der Substratbindungsstelle hervorruft. Solch ein amphiphiler Substratbindungs- und –faltungsmechanismus bietet eine Erklärung dafür, wie Spy strukturell unterschiedlichen Proteinen bei der Faltung unterstützen kann, da dieser Mechanismus keine Substratspezifität benötigt. Demnach ist die Bindung und Freisetzung des Substrats auch nicht durch das Chaperon, sondern durch die Faltung des Substratproteins selber kontrolliert. Spy agiert demnach, indem es eine Zuflucht für faltende Proteine bietet, deren Faltung verlangsamt und dabei Fehlfaltung und zelltoxische Aggregation vermeidet. Der Faltungsweg bleibt dabei von der Primärstruktur des Substratproteins beschrieben. Die elektrostatischen Bindungen zwischen Spy und Im7 übernehmen eine zentrale Rolle in diesem Mechanismus. Sie beschleunigen nicht nur die Bindung von Im7 und inhibieren damit Proteinaggregation, sie sind auch dafür verantwortlich, dass Im7 an Spy während der Faltung gebunden bleibt. Von daher ist eine Freisetzung von unvollständig gefalteten Substratprotein, wie es oft für die erfolgreiche Faltung von Substratproteinen durch Chaperone beschrieben wird, nicht notwendig. Da eine amphiphile und flexible Substratbindungsstelle eine gängige Eigenschaft von Chaperonen ist, ist es durchaus möglich, dass ein solcher vom Substratprotein selbst regulierter Faltungsmechanismus die Grundlage von vielen Chaperonen bildet.

Table of Contents

Table of figures

1 Introduction

1.1 Proteins are essential building blocks of all living matter

Proteins are among the most abundant and ubiquitous macromolecules in the biological world, controlling the vast majority of processes upon which life depends (Balchin et al., 2016). Proteins are synthesized by the ribosome as linear chains of up to a few thousand amino acids linked through peptide bonds (Noeske and Cate, 2012). Although generally made out of only 20 different amino acids, these polypeptides can form an enormous variety of three-dimensional structures with varying degrees of conformational freedom. This structural and conformational diversity generates the enormous functional and phenotypical diversity of living organisms found in a broad range of different environments on this planet (Goncearenco and Berezovsky, 2015).

To fulfill its destined function, a synthesized polypeptide chain generally must fold into a three-dimensional structure. Although it may seem possible that a protein with a given amino acid sequence could assume a vast number of structures, in reality, proteins normally only fold into a single discrete, biologically active structure. The mechanism by which folding occurs on a biologically relevant time scale remains one of the most puzzling phenomena in biological sciences (Dill and MacCallum, 2012; Gruebele et al., 2016). This extraordinary process is the result of a billion years of evolutionary pressure that has optimized the amino acid sequence of each protein to maximize not only its folding efficiency but also its stability and functionality in the context of its specific environment (DePristo et al., 2005; Gruebele, 2005). Nevertheless, despite years of evolutionary optimization, the process of folding is inherently error-prone for most proteins, and hence the cell must invest a considerable amount of energy into a network of folding helpers to ensure proper and timely protein folding (Balchin et al., 2016). Molecular chaperones take a central role in this network, as they directly assist in protein folding (Kim et al., 2013). Failure to maintain healthy protein homeostasis is closely linked to many human diseases, including Alzheimer's, cystic fibrosis, and Parkinson's (Hipp et al., 2014; Labbadia and Morimoto, 2015; Morawe et al., 2012) In view of the central role of molecular chaperones in proper protein folding and the link between human diseases and misfolding, it is crucial to gain a better understanding of the mechanism of chaperone action. Although it has been three decades since the discovery of molecular chaperones, it is still not entirely clear how these folding helpers affect protein folding. The first section of the Introduction discusses why and how proteins fold, and examines why the process of protein folding is so error-prone. The second part focuses on the concept of chaperones and describes what is known about how chaperones assist in protein folding. The mechanisms of chaperone function are also illustrated with a few prominent examples.

1.2 Folding is a fundamental process of protein biology

To fulfill their biological function, most proteins must fold into a thermodynamically stable and most often compact, globular structure called the "native state" or "folded state" (Gruebele et al., 2016). This state may represent a single conformation or a small ensemble of energetically and structurally similar conformations. Because proteins are synthesized by the ribosome as linear chains of amino acids, folding into the native state must occur following or during translation of the nascent polypeptide (Balchin et al., 2016). However, the fundamental mechanism that makes a protein capable of folding into its native state, and how the cellular environment influences the folding process are not yet completely understood (Gruebele et al., 2016).

In 1961, Christian Anfinsen discovered that the information necessary for a protein to fold into its distinct 3-dimensional, active state is encoded solely in its amino acid sequence—a discovery for which he was awarded the Nobel Prize in 1973 (Anfinsen, 1973; Anfinsen et al., 1961). Anfinsen demonstrated that chemically denatured and fully unfolded bovine ribonuclease (RNaseA), upon dilution into a buffer optimized for folding, could refold spontaneously into its active state *in vitro*, indicating that extrinsic factors are not absolutely required for a successful folding process (Anfinsen et al., 1961). Based on his experiments, Anfinsen postulated that under physiological conditions where folding occurs spontaneously, the folded structure of a protein is a unique state that is kinetically accessible and thermodynamically significantly more stable than any other state the protein may sample during folding. This idea is also known as the thermodynamic hypothesis of protein folding (Anfinsen, 1973). With his work, Anfinsen created the basis for decades of protein folding research. At first, it seemed that protein folding is basically a two-state reaction, where the unfolded, linear polypeptide chain suddenly and quickly collapses into a distinct 3-dimensional structure, measurable as a unimolecular event. However, it became clear that folding could not occur by a random thermal search through all possible backbone conformations (Karplus, 1997; Levinthal, 1968, 1969). Levinthal estimated that if folding of a small protein such as RNase A occurred by random search, it would take longer than the lifetime of the universe for it to find its biologically active state. However, experimentally, RNaseA only needs a few minutes to fold (Anfinsen et al., 1961). The discrepancy between calculations and experimental evidence is known as Levinthal's paradox. Inspired by his calculations, Levinthal hypothesized that for protein folding to occur on a biologically relevant time scale, a folding protein must follow a pathway with distinct, but very transiently populated and therefore kinetically invisible folding intermediates that guide the folding protein to its native state (Levinthal, 1968, 1969). Given the complexity and size of many proteins, how does the polypeptide chain manage to fold into a single native state in such a short time? In the following section, I summarize some principles that govern the folding process of proteins. In a later section, I discuss these principles in the context of the cellular environment.

1.2.1 Principles of protein folding and stability *in vitro*

The folded state of a protein is comprised of structured elements, such as α-helices, β-sheets, and turns, as well as random coils, which together make up the secondary structure. These secondary structure elements are arranged into an overall 3-dimensional structure, also called the tertiary structure. While the secondary structure elements are stabilized through hydrogen bonds between the backbone carbonyl oxygen and amine hydrogens that lie in relatively close proximity, the tertiary structure is mainly stabilized through many weak non-covalent bonds between amino acid side chains that are either close to or far apart from each other in the sequence of the protein. These non-covalent bonds, which include salt bridges, hydrogen bonds, van der Waals (VDW) interactions, stacking interactions between aromatic residues, and other interactions (Brockwell and Radford, 2007), must form during protein folding. As for any other chemical reaction, fundamental thermodynamic and kinetic principles govern the formation of the above described bonds and thus protein folding. From a thermodynamic point of view, a reaction occurs spontaneously if the Gibb's free energy difference (ΔG) between reactant and product is negative:

$$\Delta G = \Delta H - T * \Delta S = -R * T * lnK$$

where ΔH is the reaction enthalpy, which describes the thermal energy as well as the volume work exchanged with the environment; T is the absolute temperature at which the reaction occurs; ΔS is the reaction entropy, which describes the amount of order or disorder the reaction creates; R is the ideal gas constant; and K is the folding equilibrium constant, which describes the ratio of folded over unfolded protein in a single population (Privalov and Makhatadze, 1993). In the unfolded state, a protein mostly lacks tertiary structure and contains little to no secondary structure and intramolecular bonds and hence enjoys a high degree of conformational flexibility or entropy (Bowler, 2012; Kazmirski et al., 2001). In contrast, the folded state is often a tightly packed state with little flexibility, stabilized cooperatively through many weak non-covalent bonds (Dill, 1985; Horovitz and Fersht, 1992). The free energy difference (ΔG) between the unfolded and folded state, defined by the structural elements and bonds formed, drives a protein to fold. The pathway that has the lowest energy barrier to the folded state is the one taken by the folding protein. The conformational space a protein can sample as a function of the corresponding thermodynamic stability of accessible conformational states is often visualized as a funnel shaped folding energy landscape that the protein navigates during the folding process (Figure 1) (Dill and MacCallum, 2012; Karplus, 2011; Onuchic et al., 1995). Each coordinate on the surface of the folding funnel corresponds to a conformer of specific secondary and tertiary structure with a distinct number of non-covalent bonds within the polypeptide and with the solvent. For most globular, compactly folded proteins, a gain in structure results in a heat release and a creation of entropy, which renders ΔG for folding negative and therefore favors folding. The shape of this energy landscape is highly dependent on the individual protein and the environment

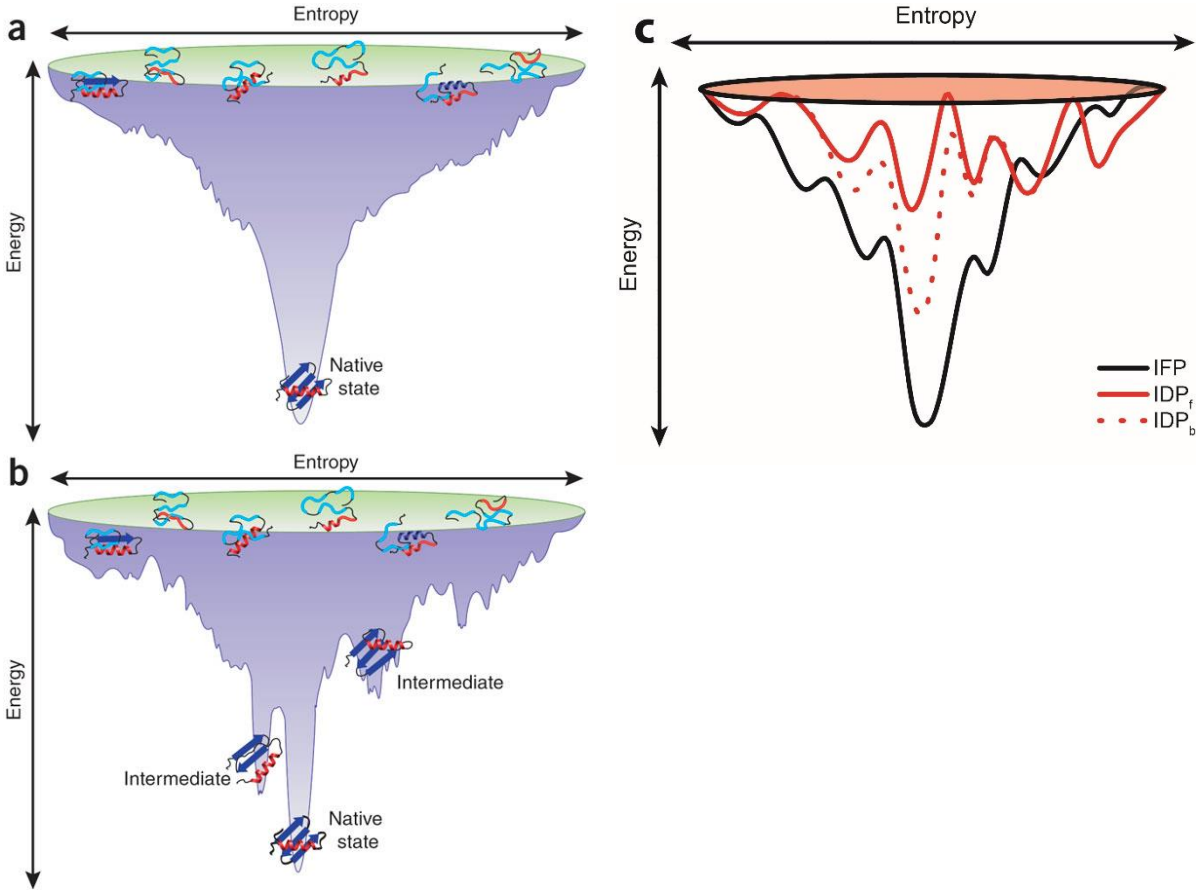

Figure 1: The energy landscape of folding proteins. Idealized folding landscapes of (a) two- or (b) multi-state transition folding of a spontaneously folding protein in absolute isolation. Folding energy (or free energy) is plotted as a function of polypeptide configurational entropy. A two-state transition is the simplest folding mechanism and has been observed for many small, globular, fast-folding proteins (a). For most proteins, however, the folding landscape contains local minima with high energy barriers that trap the protein in intermediate folding states that can be structurally more or less close to the native conformation (b). In contrast to the native state (which corresponds to a very narrow energy minimum as it is comprised of a single or an ensemble of relatively few biologically active conformations), in the unfolded state, a protein populates a large ensemble of conformations with high energy, conformational freedom, and little native and non-native residual structure. A loss in conformational entropy upon folding is compensated for by an energetic stabilization (Chodera and Mobley, 2013) (image adapted from (Bartlett and Radford, 2009) with permission from the journal). (c) Overlay of an idealized folding landscape for a spontaneously folding protein (IFP, black) and an intrinsically disordered protein ($IDP_{f,}$, red,). IDPs are mostly disordered in solution, sampling different lowly populated secondary structure-containing conformations. These conformations are usually stabilized upon binding to a target (IDP_b, dotted line); in other words, IDP folding is induced by target binding.

in which it folds (Gershenson et al., 2014). Moreover, the conformational state that represents the energetic minimum on this energy landscape and as a consequence is the most populated state at equilibrium is determined by: (1) the amino acid composition, as each of the 20 amino acids has unique chemical as well as steric properties (Grahnen et al., 2011; Ilardo et al., 2015); (2) the order in which the amino acids occur in the linear polypeptide chain, the so-called protein sequence or primary

structure; and (3) the environmental conditions, such as temperature, pressure, solvent, pH, redox potential, and ionic strength (Khorasanizadeh et al., 1996).

1.2.1.1 What drives a protein to fold? Some thermodynamic aspects

Where does the favorable energy for folding in the form of enthalpy and entropy come from? The cellular milieu of every organism is an aqueous environment (Spitzer and Poolman, 2009; Takei, 2015). Through evolution, nature has optimized the amino acids used by every organism to a minimum set of 20 chemically and structurally unique residues (Ilardo et al., 2015). These are sufficient not only to construct a vast variety of three-dimensional structures through different non-covalent and covalent bonds and steric constraints (Ramachandran et al., 1963), but also to provide chemical reactivity for all sorts of catalyzed metabolic reactions, signal sensing, and transduction. About one third of the 20 canonical amino acids are hydrophobic; i.e., their side chains lack polar groups and therefore are not very soluble in the aqueous environment of the cell (Kellis et al., 1988). As a result, water forms cages around solvent-exposed hydrophobic residues. These cages due to their rigidity gain crystal like properties. This process involves work; it is called cavity work because it takes energy for water to form a cavity around the solubilized hydrophobic residue. Because this process takes energy, the ΔG of solubilization of hydrophobic residues is positive. The energy penalty that renders ΔG positive is entropic due to cage formation, which is a highly ordered process (Southall et al., 2002).

Hence, one major factor that destabilizes the unfolded over the folded state, and therefore drives folding is the solvent exposure of hydrophobic residues in the unfolded state (Baldwin, 2007; Bartlett and Radford, 2010; Chandler, 2005; Dill, 1990, 1985; Kellis et al., 1988; Pace et al., 2011; Southall et al., 2002). These residues are buried in the hydrophobic cores of globularly folded proteins. Thus, upon folding, the highly ordered water molecules get released, which creates a large increase in entropy (ΔS). As the number of hydrophobic residues in a protein sequence increases, this entropic effect upon folding becomes stronger. This so-called hydrophobic effect was first described by Tanford (Tandford, 1979). Apart from the entropic gain, the formation of the hydrophobic core also results in heat release. This is because the tight packing of the hydrophobic core results in a high density of cooperatively acting VDW interactions, the formation of which generates heat. The summarized enthalpic gain due to the tight packing outweighs the energy needed to break the crystal water hydrogen bonds and the VDW interactions between hydrophobic residues and the water molecules. The formation of the hydrophobic core delivers most of the energy that drives folding. In addition, other non-covalent bonds between amino acid residues, such as hydrogen bonds and salt bridges, contribute to the overall stability of the folded state. Their contribution to the overall folding energy is minor compared to the formation of the hydrophobic core mainly because the residues involved in these interactions can also form equivalent bonds with water and solubilized ions (Pace et al., 2011; Yesselman et al., 2015). However, these intramolecular hydrophilic interactions are

important for specificity and guidance of the folding process as they may form very early in the folding process (Anfinsen, 1973; Rose et al., 2006). In this context, some of these bonds create a more stabilizing effect than others of the same kind; this is especially true for backbone hydrogen bonds, as they kinetically favor the formation of structural elements (Baldwin, 2007; Hu et al., 2013; Strop and Mayo, 2000). A more complete list of energetic contributors to folding can be found in Table 1.

There are almost as many thermodynamically destabilizing effects associated with the process of folding as there are stabilizing ones. As a result, apart from a few exceptions (Xia et al., 2012), most proteins are only marginally stable (DePristo et al., 2005). The ΔG of protein folding is generally only 12–42 kJ mol^{-1}, which is equivalent to the energy contained in only one to a few hydrogen bonds. As a consequence, the folding equilibrium constant (K) is rather low, and thus an individual protein molecule folds and unfolds many times during its lifetime (Gruebele, 2005; Gruebele et al., 2016). Genetic and biochemical studies have revealed that this low stability has been evolutionarily selected for in many proteins as it allows for a certain degree of conformational freedom/flexibility that is needed for biological activity and protein turnover (Bloom et al., 2007; Dellus-Gur et al., 2013; Foit et al., 2009; Gosavi, 2013; Gosavi et al., 2008; Jäger et al., 2006).

Table 1: Processes occurring during folding that thermodynamically favor or disfavor folding*

Contribution/Penalty	***ΔH***	***ΔS***
Dehydration of hydrophilic residues	Positive ΔH: due to breaking of dipole interactions and hydrogen bonds with the water	Positive ΔS: Release of water molecules creates little entropy
Dehydration of hydrophobic residues	Positive ΔH: due to breaking of hydrogen bonds amongst the crystal water; breaking of VDW interactions between residues and water	Positive ΔS: Release of water molecules from hydrophobic cages creates large amounts of entropy
Formation of secondary structure through hydrogen bonds	Negative ΔH: formation of backbone hydrogen bonds results in heat release	Negative ΔS: creation of local symmetry results in loss of freedom
Formation of the hydrophobic core	Negative ΔH: tight packing of hydrophilic residues results in strong cooperative VDW interactions, therefore heat release	Negative ΔS: tight packing results in loss of conformational freedom of side chains and backbone.
Other bonds important for the stability of the tertiary structure	Negative ΔH: formation of bonds result in heat release	Negative ΔS: additional loss of conformational freedom

*Thermodynamic contributions or penalties are listed as for a two-state folding model, comparing the unfolded and folded state (Makhatadze, 2005). Note that a negative ΔH drives protein folding, whereas a negative ΔS counteracts folding and *vice versa*.

Due to the hydrophobic effect, the intrinsic tendency for a protein sequence to spontaneously fold in aqueous solution depends very much on the ratio of hydrophobic to hydrophilic residues. In recent years, evidence has emerged that a significant percentage of the proteome in every branch of life (bacteria, archaea, viruses, and eukaryotes) consists of proteins that do not spontaneously fold on their

own in solution, the so-called intrinsically disordered proteins (IDPs) (Burger et al., 2016). IDPs have a high percentage of hydrophilic (i.e., water soluble) residues, including an excess of uncompensated charge, which leads to intramolecular repulsion and energetically favorable solvation of the extended polypeptide chain (Uversky, 2016). In contrast, intrinsically folded proteins have a high percentage of hydrophobic amino acids (about 1 hydrophobic reside for every 3.3 residues) (Tompa and Tompa, 2002). Hence, IDPs, free in solution, do not fold spontaneously. However, upon interaction with their biological target, which changes the local environment sufficiently to favor a more structured state, they may fold (Figure 1) (Gianni et al., 2016; Maghawry et al., 2014; Shoemaker et al., 2000; Uversky et al., 2008; Xue et al., 2012). Contrary to the unique folded state of intrinsically folded proteins, IDPs may adopt many environment-dependent structures, which is why these proteins can also be called conditionally folded proteins (Bardwell and Jakob, 2012; Burger et al., 2016).

1.2.1.2 Which pathway does a folding protein choose? Some kinetic aspects of on-pathway protein folding

As indicated in Figure 1, in addition to proteins that populate only two states during folding (the unfolded and folded state), there are proteins that populate obligatory intermediate states along their folding pathway (Baldwin, 2008; Brockwell and Radford, 2007). There is evidence that many apparently two-state folding proteins also transiently populate folding intermediate(s) that are difficult to resolve experimentally (Beauchamp et al., 2012; Khorasanizadeh et al., 1996). In the cases of both two-state and multi-state folding, it is not yet entirely clear in which order all the thermodynamically stabilizing non-covalent bonds and structural elements described above are formed, and it is therefore unclear how a protein navigates on its folding energy landscape (Baldwin, 1995; Englander and Mayne, 2014; Udgaonkar, 2008). Protein folding is often a very rapid process, occurring in some cases within a few microseconds. This is due to the low energy barriers for folding exhibited for many two-state folders (Baker, 1998; Bryngelson et al., 1995; Kubelka et al., 2004). This speed indicates that transitions between populated conformational states (i.e., from unfolded to folded) do not occur directly via random thermal motions, but rather involve transitions between very transient discretely populated conformational intermediates with increasing native-like structural content that guide the folding process (Dill et al., 2008; Karplus, 1997; Levinthal, 1968, 1969; Sekhar and Kay, 2013). Moreover, it is still unclear if the native state can be reached through multiple pathways or if folding is a highly hierarchical process, with well defined on pathway transient intermediates (Englander and Mayne, 2014).

The hydrophobic effect not only has thermodynamic consequences, but also dictates the kinetics of the folding process. Due to the entropically very unfavorable formation of water cages around exposed hydrophobic residues in the unfolded state, the first folding step (occurring during the first few microseconds of folding) is the formation of a hydrophobic core through a collapse of the elongated

polypeptide chain. This is termed the hydrophobic collapse (Bartlett and Radford, 2010; Felitsky et al., 2008; Gruebele, 1999; Haran, 2012; Ziv and Haran, 2009). The immediate result of this collapse is a more compact conformation that restricts the conformational space that must be searched during folding. Since the hydrophobic collapse brings distant residues close together, it allows the formation of tertiary hydrophobic interactions as well as additional stabilizing hydrogen or ionic bonds (Capaldi et al., 2002; Dinner et al., 2000; Fersht and Sato, 2004). The state following hydrophobic collapse is sometimes referred to as the molten globular state due to its increased compactness but still rather high degree of conformational plasticity compared to the native state (Arai and Kuwajima, 2000; Ptitsyn, 1995). A molten globule may be stabilized through native-like but also non-native interactions. The hydrophobic collapse is accompanied by the formation of native-like secondary structure through backbone hydrogen bonds, which are favored by the local high density of order-promoting residues (Bowler, 2012; Haran, 2012; Lindorff-Larsen et al., 2011; Uversky, 2016). These secondary structure elements, which generally consist of one to a few α-helices or β-sheets, are lowly populated in the unfolded state, but become stabilized through non-covalent tertiary bonds with other residues as a result of the hydrophobic collapse. This leads to a cooperative stabilization of the native-like structural elements, which acts to guide the folding process. Native-like bonds formed this early during the folding process are often important for the stability of the native state and when mutated may render the protein incapable of folding (Baldwin, 2007; Hu et al., 2013; Pashley et al., 2012; Strop and Mayo, 2000). The structural elements are seen as the building blocks in the folding process and are sometimes also referred to as "foldons" (Bai and Englander, 1996; Fersht, 1997).

Although these early folding steps are well accepted, due to technical limitations, it is still under debate in which order these foldons form, and if a protein can reach its folded state through multiple pathways (Figure 2) (Englander and Mayne, 2014). On the residue level, stochastic, thermal search may lead to the formation of energetically more favorable local structures (foldons). However, macroscopically, a random search through possible structural elements would take too long for a protein to reach its native state (Levinthal, 1968, 1969). It seems plausible that folding is a discrete process, described by a single or a few optional pathways. The formation of the first native-like foldon at the state of the molten globule likely builds the foundation of the folding process and guides the sequential formation of a series of foldons in a cooperative fashion through energetically favorable native-like interactions (Baldwin and Rose, 1999; Englander and Mayne, 2014). Evidence for such a hierarchal folding process is provided by experiments as well as through molecular dynamic simulations, illustrating that even apparently two-state folding proteins can fold through distinct but transient folding intermediates (Bai et al., 1995; Englander et al., 2007; Krishna et al., 2007; Lindorff-Larsen et al., 2011; Sekhar and Kay, 2013). The fact that slow- as well as fast-folding proteins have similar folding transition times further points towards a cooperative, hierarchical folding mechanism (Chung et al., 2012). In the case of a multi-pathway model, the identity of the first foldon formed may dictate the optional folding pathway on the energy landscape (Englander and Mayne, 2014). Examples

of proteins that have been found to fold through a distinct or at least major pathway include cytochrome C (Englander et al., 2007), lysozyme (Krishna et al., 2007), maltose binding protein (MBP) (Walters et al., 2013), RNase H (Hu et al., 2013), and *E. coli* immunity protein 7 (Im7) (Bartlett and Radford, 2010). Im7 is used as a folding model protein in this thesis and hence its folding pathway is discussed in detail in the following section.

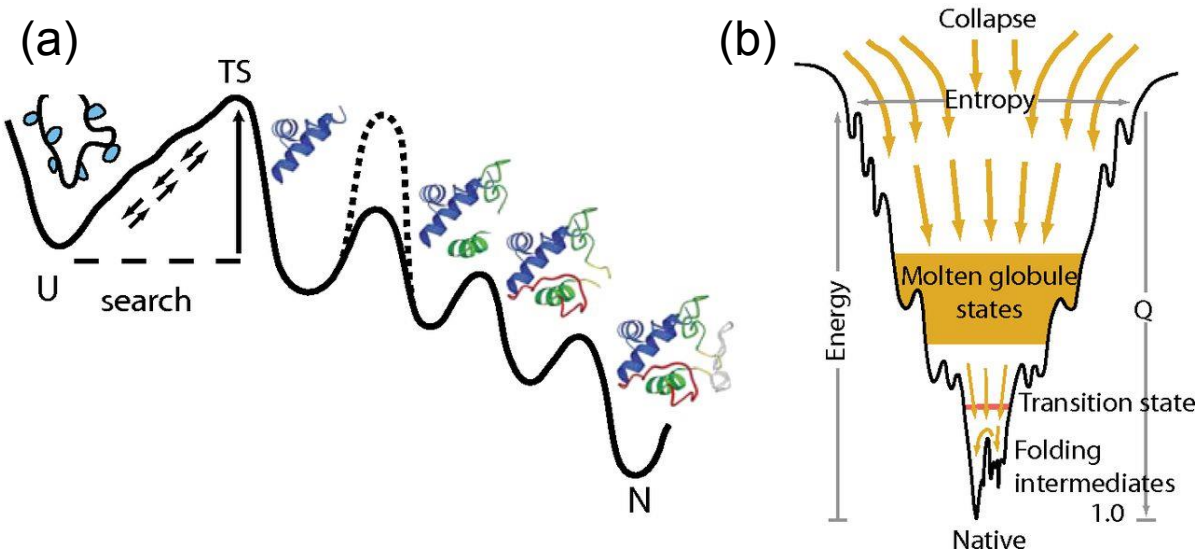

Figure 2: A possible hierarchical nature of protein folding pathways. Whether or not proteins fold through a defined pathway (a) or multiple parallel pathways (b) is not entirely clear. Although there is some experimental evidence that proteins can fold through multiple pathways and the process of folding is therefore asynchronous, higher resolution data from hydrogen-deuterium-exchange mass spectrometry (HX MS) experiments have recently pointed towards a hierarchical model that includes one or a few distinct states that are transiently populated on the protein's way to achieving the native state. (a) Random thermal search through secondary structure and a collapse of the chain lead to the first stabilizing contacts and native-like secondary structure, which forms the basis for a hierarchical folding pathway. The dotted energy barrier represents optional off-pathway misfolding events, such as proline mis-isomerization, incorrect disulfide bond formation, or non-native hydrophobic clustering, that kinetically trap part of the protein's population. Low resolution methods often applied to characterize the folding process of a protein, such as fluorescence spectroscopy, may misleadingly lead to the conclusion of multi-pathway folding (Bédard et al., 2008). (b) Alternative view of multi-pathway protein folding. The highly entropic unfolded state can collapse into a multitude of different conformational states, all of which eventually lead to the native state at the global free energy minimum, shown at the bottom of the figure. (Image adapted from Englander and Mayne, 2014).

In summary, while unfolded, the protein population samples a vast range of conformations (Bowler, 2012; Dyson and Wright, 2002; Voelz et al., 2010). With each folding step, the accessible conformational space narrows down, and as a result, its conformational entropy decreases. This entropic penalty is offset entropically by the desolvation of hydrophobic residues and enthalpically by the formation of more stable native-like non-covalent interactions between amino acid residues (Englander and Mayne, 2014; Gruebele et al., 2016). As a result of this hierarchal, cooperative folding

mechanism, the folding energy landscape of small proteins can often be represented as a relatively smoothly shaped funnel that lacks high energy barriers (Figure 1). As a consequence, some small proteins fold on the timescale of microseconds, also referred to as downhill folding (Dill and MacCallum, 2012).

1.2.1.3 Im7 as a protein folding model

The *E. coli* immunity protein 7 (Im7) has one of the most well characterized folding pathways, and is a good example of a protein that employs a hierarchical folding process (Figueiredo et al., 2013; Wang et al., 2016). Im7 belongs to the family of bacterial immunity proteins. It is co-expressed with and functions as an inhibitor of the bactericidal DNase colicin E7, to which it binds very tightly and very rapidly (Dennis et al., 1998; James et al., 1996; Ko et al., 1999; Li et al., 2004). Im7 is a globular protein that contains four α-helices that form a hydrophobic core (Chak et al., 1996). Over the past two decades, experimental as well as computational work has led to a detailed picture of Im7's folding pathway (Figure 3). The most striking characteristic of Im7 is that despite its small size of 87 amino acids, it folds through an on-pathway intermediate state (I) (Capaldi et al., 2001; Ferguson et al., 1999). Although the transition from the unfolded state (U) to (I) occurs quite rapidly (sub millisecond, folding rate constant k_{UI} = 3000 s^{-1}, folding equilibrium constant K_{UI} = 140 at 10 C and 0.4 M Na_2SO_4), the transition to the native state (N) is slower (k_{IN} = 300 s^{-1}, K_{IN} = 720 at 10 C and 0.4 M Na_2SO_4) and hence is the rate limiting step of Im7's folding pathway (Capaldi et al., 2001). As a consequence, the intermediate state is sufficiently populated during refolding to be characterized experimentally. In addition, protein engineering in conjunction with biochemical and biophysical approaches have delivered detailed structural information on all folding states of Im7 that are populated along the folding trajectory, including the two transition states.

The unfolded state ensemble of Im7 consists of moderately extended species with low helical propensity and without any persistent long range stabilizing tertiary interactions. However, it does contain local hydrophobic clusters that act to conformationally constrain the regions of the protein that ultimately form helices 1, 2, and 4 in the native state. These constraints represent an early local collapse driven by hydrophobic interactions that guide the folding protein towards the intermediate state (Capaldi et al., 2002; Le Duff et al., 2006; Pashley et al., 2012; Pugh et al., 2010). The transition to the intermediate state is described by the first transitions state (TS1). Although TS1 was found to be rather extended, i.e., structurally similar to the unfolded state, there is experimental evidence for early persistent native and non-native tertiary interactions between helices 1 and 2 that lead to a further collapse in this region (Friel et al., 2009). In other words, TS1 is structurally closer to (U) than (I), and hence the activation energy needed for its formation is rather small. Conformational changes from TS1 to (I) mark the key step in the Im7 folding pathway as they involve the cooperative formation of a hydrophobic core as well as some native-like secondary structure (Bartlett and Radford, 2010; Friel et al., 2009). Although this intermediate state ensemble contains mostly native-like secondary structure,

as three of the helices are partially (N-terminal half of helix 2) or completely (helices 1 and 4) formed, its tertiary structure is distorted through non-native but energy minimizing interactions that act to shield otherwise exposed hydrophobic side chains (Capaldi et al., 2002; Chen and Chan, 2015; Friel et al., 2009; Gorski et al., 2004; Gsponer et al., 2006; Whittaker et al., 2011). One example of a non-native interaction is residue W75, which is more buried in the hydrophobic core of the intermediate state than it is in the native or unfolded state. As a result, the intermediate state is more fluorescent than the other two states. This is an experimentally important feature, as it allows for the rapid and accurate discrimination between these three states experimentally. Consequently, the fluorescence emission of W75 has been used to study the kinetics of the Im7 folding pathway (Gsponer et al., 2006; Rodriguez-Mendieta et al., 2005). The non-native arrangement of formed helices in the intermediate state prevents the formation and docking of helix 3 onto the hydrophobic core. Therefore, the transition to the native state requires both conformational rearrangements and the disruption of these non-native stabilizing interactions in order to form a docking site for helix 3. (Capaldi et al., 2002; Friel et al., 2009; Gsponer et al., 2006; Knowling et al., 2009). Therefore, the energy barrier between (I) and (N) is higher than the one between (U) and (I), and thus the (I) to (N) transition represents the rate limiting step (Figure 3). A similar folding mechanism has been described for barnase (Fersht, 1993).

Im7, like many other proteins, is only moderately stable (ΔG = -17 kJ mol^{-1}) (DePristo et al., 2005; Ferguson et al., 1999). A genetic selection for increased thermodynamic stability *in vivo* resulted in most stabilizing mutations being found at the colicin E7 binding site. This was taken as evidence that for Im7, stability is compromised by an evolutionary selection for tight and fast binding to its biological target colicin E7 (Foit et al., 2009). Along these lines, not only the stability of Im7, but also its folding efficiency is compromised to yield optimal functionality. Many of the hydrophobic residues that form non-native contacts early during folding and thus lead to the formation of the rate-limiting intermediate state are crucial for the formation of the E7 binding site (Friel et al., 2009). Directed mutation of these residues is predicted to yield a faster, two-state folding mechanism (Sutto et al., 2007). Together, these results highlight the evolutionary competition between biological functionality, thermodynamic stability, and the kinetic efficiency of protein folding (Gershenson et al., 2014).

Im7 is also an excellent example to illustrate the importance of the hydrophobic effect, as well as the importance of order-promoting amino acid resides in the early stages of protein folding. The mutation of three key leucine residues to alanine (L18A L19A L37A), which are important for the formation of helices 1, 2, and 4 and the formation of early hydrophobic contacts, and ultimately, the hydrophobic core, renders Im7 incapable of folding. In other words, this triple alanine mutant predominantly populates the unfolded state under folding conditions, despite the fact that all other residues that stabilize the native state are still present (Pashley et al., 2012; Sutto et al., 2007). This is because these residues, together with other hydrophobic residues in the region of helices 1 and 2, form tertiary

stabilizing interaction during TS1 and therefore are essential for initiation of the folding process (Friel et al., 2009; Rodriguez-Mendieta et al., 2005).

Detailed phi-value analyses (Zarrine-Afsar and Davidson, 2004) of Im7's folding pathway have helped to generate an arsenal of structurally, thermodynamically, and kinetically well-characterized mutants. Different mutations trap the protein in the unfolded state, the intermediate state, a combination of states, or simply destabilize the native state (Bartlett and Radford, 2010; Capaldi et al., 2002; Friel et al., 2009). Thus, Im7 is not only an excellent model for studying protein folding, but it also provides a useful tool for investigating factors that influence the complex process of protein folding and stability *in vivo* (Foit et al., 2009; Quan et al., 2011). Unlike many other proteins, it is possible to find conditions in which Im7 stays soluble regardless of its folding state. Thus, Im7 it can also be used as a chaperone client protein without complications associated with protein aggregation. Using a genetic selection system that directly links thermodynamic protein stability with antibiotic resistance *in vivo*, Quan *et al.* discovered a new periplasmic molecular chaperone called Spy by selecting for stabilization of an unstable variant of Im7 (L53A I54A) that traps Im7 in the intermediate state.

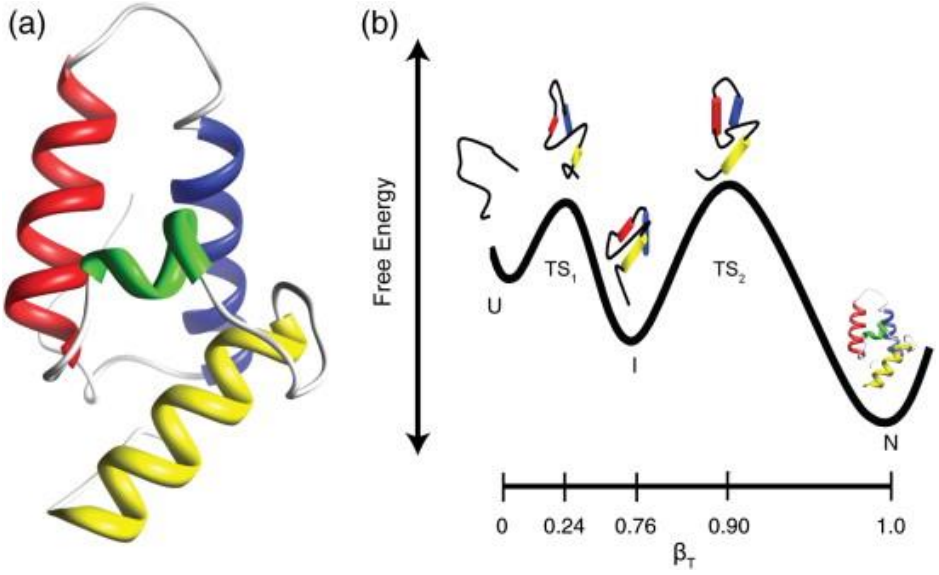

Figure 3: The α-helical protein Im7 folds via an on-pathway intermediate. (a) Cartoon representation of the crystal structure of Im7 (PDB ID: 1AYI). The four helices are color coded as follows: helix 1 (blue), helix 2 (red), helix 3 (green), and helix 4 (yellow). (b) Energy diagram for the folding pathway of Im7. The free energy of populated conformational states is plotted as a function of compactness β_T (β_T of 0 = unfolded state, β_T of 1 = folded state). Schematic structures of the intermediate state as well as both transition states are based on experimentally restrained molecular dynamic simulations. Helical color code is identical to (a). This figure was taken from (Knowling et al., 2009).

1.2.1.4 Off-pathway protein folding events

To uncover the principles underlying the complex process of protein folding, the scientific community has mostly studied small (< 200 amino acids), well behaved proteins that exhibit a relatively simple

folding pathway with little to no on-pathway folding intermediates (Barrick, 2009; Braselmann et al., 2013; Brockwell and Radford, 2007). These proteins reversibly fold in the environment of the test tube, and hence they are ideal protein folding models. Our understanding of such simple folding pathways is advanced enough that they can even be partially recapitulated through molecular dynamics simulations (Lindorff-Larsen et al., 2011). However, such proteins do not represent the majority of all cellular proteins, also called the cellular proteome. For example, about 85% of *E. coli* proteins are significantly larger, reveal more complex, often multidomain structures, and may even assemble into multi-protein complexes. In addition, the size and complexity of proteins increases from prokaryotes to eukaryotes (Balchin et al., 2016). Eukaryotic proteins exhibit more complex folding pathways, including multiple on-pathway intermediates (Braselmann et al., 2013; Dahiya and Chaudhuri, 2013; Pirchi et al., 2011; Scholl et al., 2014). The energy folding landscape of eukaryotic proteins generally isn't smooth, but rather rugged, including local minima with high energy barriers that kinetically trap the folding protein in an intermediate state (Figure 1) (Gershenson et al., 2014). Consequently, eukaryotic proteins often fold slower and are more prone to off-pathway misfolding events, such as non-native hydrophobic clustering (Klein-Seetharaman et al., 2002), proline mis-isomerization (Brandts et al., 1975; Wedemeyer et al., 2002), and incorrect disulfide bond formation (Fass, 2012; Song and Scheraga, 2000). Consequently, folding energy barriers can be entropic and/or enthalpic in nature because kinetically trapped folding intermediates or misfolded states may exhibit extensive conformational flexibility and/or non-native bonds (see Table 1). This, in addition to their generally low thermodynamic and kinetic stability, causes these more complex proteins to fold inefficiently in the test tube (Figure 4) (Braselmann et al., 2013). One reason for this is the hydrophobic effect—although it is the major kinetic and thermodynamic driving force of protein folding, it comes with a dangerous side effect. Most folding intermediates of a protein on the way to its native state expose hydrophobic residues. Complex folding pathways and off-pathway misfolding events may prolong the time in which a protein stays in a hydrophobic surface-exposed conformation. Due to this thermodynamically unfavorable exposure of hydrophobic surfaces, folding intermediates or misfolded states tend to engage in non-native intermolecular interactions, a process that leads to aggregation (Figure 4 and Figure 6) (Chandler, 2005; Chiti and Dobson, 2006; Ciryam et al., 2013a). Protein aggregation is an off-pathway reaction characterized by the formation of non-native but energetically stabilizing intermolecular interactions between the hydrophobic surfaces of two or more proteins. When buried in an aggregate, the protein is generally unable to fold and therefore is unable to fulfill its biological function. Aggregation is most often characterized as being amorphous or fibrillar in nature (Figure 6) (Chiti and Dobson, 2006; Ciryam et al., 2013a). Shifts in temperature, pH, or the addition of denaturants (such as organic solvents) may further shift the fragile conformational equilibrium to partially folded or unfolded states, which are prone to misfold, and hence increase the aggregation propensity of proteins. In contrast to folding, which is a reversible and concentration-independent process, protein aggregation is highly concentration dependent and is essentially

irreversible. Thus, not only in the test tube, but especially in the crowded environment of the cell, protein aggregation is in constant competition with folding (Ellis and Minton, 2006; Espargaró et al., 2008). This causes practical problems not just for experimental scientists, but also in the pharmaceutical industry, as many therapeutic proteins are quite aggregation prone (Estes et al., 2015). Due to the loss in biological activity and altered structural propensities that proteins trapped in an aggregate exhibit, including the exposure of hydrophobic surfaces, protein oligomers and aggregates are often toxic to the cell (Bolognesi et al., 2010; Hipp et al., 2014; Luheshi et al., 2008). To combat this, the cell has developed strategies that support protein folding and that reduce the amount of protein aggregates.

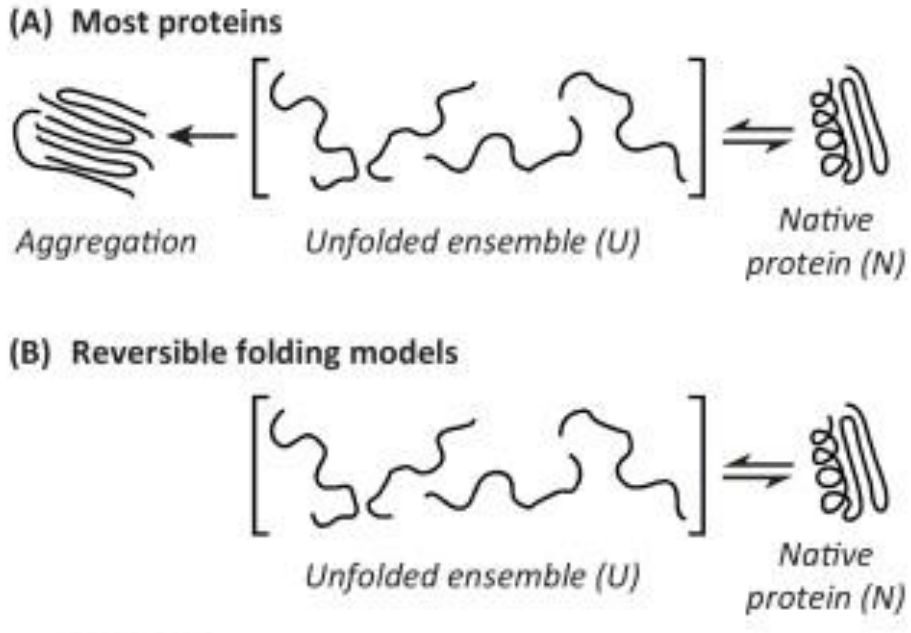

Figure 4: Protein folding for most proteins is an aggregation-prone process. The unfolded state (U) of most proteins is aggregation prone. Hence, upon the removal of denaturant, kinetically reversible folding to the native state (N) competes with kinetically irreversible aggregation (A). The slower a protein folds to the stable native state, the larger the aggregation potential. Protein aggregation interferes with the determination of the thermodynamic stability as well as the kinetics of protein folding. Consequently, most proteins used as models to study protein folding *in vitro* are not aggregation prone under conditions optimized for folding (B) (Figure was adapted from (Braselmann et al., 2013) with permission from the journal)

1.2.2 Protein folding inside the living cell and the proteostasis network

The *in vivo* folding environment stands in stark contrast to the dilute environment of the test tube, where folding is usually observed without any intermolecular interactions (Braselmann et al., 2013). The most striking difference *in vivo* is the presence of high concentrations of co-solutes, such as salts, osmolytes, and up to 400 mg/mL of macromolecules, such as proteins and nucleic acids (McConkey, 1982; Zimmerman and Trach, 1991). Hence, the cytosol is a very dense, viscous solution, which promotes non-specific interactions between macromolecules. These interactions have a detrimental

effect on the folding landscape, as they drive folding proteins into aggregation (Espargaró et al., 2008; Ignatova et al., 2007) (Figure 4 and Figure 6). In addition, proteins are synthesized as unfolded linear polypeptide chains by the ribosome, starting at the N-terminus. Translation often kinetically delays folding and hence alters the folding landscape by prolonging the time a protein persists in an unfolded or partially folded state (Ciryam et al., 2013b). Hence, during synthesis, many proteins remain prone to misfolding and aggregation until their stable native tertiary or quaternary state is reached. Finally, changes in the composition and chemical properties of the cellular milieu affect protein folding and stability *in vivo*. Even though the cell is a tightly regulated environment, changes in temperature, redox potential, pH, or ionic strength may arise under stress conditions, induced either internally, for example during aging (Hipp et al., 2014; Sharma and Chatterji, 2010), or by changes in the external environment (Micevski and Dougan, 2013). We know from *in vitro* experiments that proteins are only marginally stable due to the importance of conformational flexibility in generating biological activity (DePristo et al., 2005; Gruebele, 2005). Hence, because protein folding and stability are evolutionarily adapted to the proteins natural localization and environmental conditions inside the cell (Grahnen et al., 2011; Levy et al., 2012), cellular stress unavoidably shifts the fragile conformational equilibrium of proteins by destabilizing the native state and thus tends to drive proteins into toxic aggregates (Figure 6).

In light of the constant threat of protein misfolding and aggregation, cells in all three branches of the tree of life have developed strategies to optimize protein folding efficiency and maintain a healthy conformational equilibrium of their proteome, even under stress conditions. The maintenance of protein homeostasis is also termed proteostasis (Roth and Balch, 2011). To ensure the integrity of its proteome, the cell invests in a complex protein quality control machinery that consists of a network of host factors that synergistically guide proteins from synthesis through degradation. These cost factors assist in folding, ensure conformational integrity, and control aggregation under stress conditions by complexing non-native proteins and/or sequestrating them into transient, reversible assemblies of proteins for temporary deposition, such as stress granules (Miller et al., 2015). These host factors can partially dissolve aggregates and mediate the degradation of damaged and irreversibly misfolded proteins. This so-called proteostasis network (PN) consists of three major components: (1) folding helpers, collectively called molecular chaperones, which mediate protein folding, trafficking, and conformational repair and therefore play the central role in the PN (Figure 6) (Kim et al., 2013); (2) folding catalysts that accelerate aforementioned slow-folding reactions that can lead to off-pathway misfolding events such as proline mis-isomerization (Brandts et al., 1975; Schmidpeter and Schmid, 2015; Wedemeyer et al., 2002) and incorrect disulfide bond formation (Fass, 2012; Song and Scheraga, 2000); and (3) factors that mediate protein degradation, such as the prokaryotic AAA+ proteases (Buchberger, 2013) or the eukaryotic ubiquitin-proteasome system and autophagy (Cohen-kaplan et al., 2016). Hence, the PN is a highly flexible system that not only maintains the cellular proteome but has the ability to remodel it, which allows cells to adapt to internal changes (e.g., those

that occur during different stages of the cell cycle) as well as changes in the environment (Powers and Balch, 2013; Powers et al., 2009).

In addition, cells can respond to harsh environmental conditions that cause increased protein aggregation such as high temperatures or oxidative stress by temporarily upregulating the expression of PN proteins through multiple stress response pathways. These stress response pathways sense a proteostasis imbalance. Examples are the unfolded protein response, the oxidative stress response, or the heat shock response (Akerfelt et al., 2010; Hetz et al., 2015; Pincus, 2016; Schumann, 2016; Truscott et al., 2011; Walter and Ron, 2011; Zhang et al., 2015). The proteostasis network is vital for cellular fitness and survival, and not surprisingly, cells dedicate a tremendous amount of energy into the PN, which comprises about 10-15 % of the entire mammalian proteome (Hipp et al., 2014). The knock-out of PN components or a decline in PN capacity, for example during aging, causes an imbalance in the conformational equilibrium of the cellular proteome, which leads to increased protein aggregation and eventually cell death (Hipp et al., 2014; Labbadia and Morimoto, 2015). Under circumstances of excessive protein aggregation inside the cell, amorphous or fibrillar, the PN capacity is massively diminished since most chaperones are trapped bound to aggregates. This is a cause and a consequence of several human chronic diseases (Labbadia and Morimoto, 2015). Hence, the PN constitutes a target for drug development (Assimon et al., 2013; Taldone et al., 2014).

In light of the central role of molecular chaperones in the PN and human diseases associated with the decline of the PN, it is crucial to gain a better understanding of the mechanisms of chaperone action. This thesis primarily focuses on interactions between chaperones and folding proteins and how chaperones affect the protein folding landscape. Therefore, the chaperone concept is discussed in further detail in the next sections. The mechanisms of chaperone function are also illustrated with a few prominent examples of molecular chaperones.

1.3 Mechanisms of chaperone function

Molecular chaperones are ubiquitous proteins that are conserved throughout all branches of life (Kim et al., 2013), suggesting that chaperones have existed since the very early stages of life. Although molecular chaperones are generally thought of as proteins, recently other molecules such as osmolytes, nucleic acids, and polyphosphate have been found to play roles in protein folding as well (Docter et al., 2016; Gray et al., 2014; Ignatova and Gierasch, 2006). Due to their role in the cellular stress response, chaperones were first discovered in cells that were exposed to heat stress. Cells respond to heat induced folding stress by upregulating the expression of a set of proteins, collectively called heat shock proteins (Hsps). Many heat shock proteins are molecular chaperones that are classified into different groups based on sequence or structural homology. These groups were initially named according to their monomeric molecular weights: Hsp40s, Hsp60s, Hsp90s, Hsp100s etc. (Figure 7, Figure 8, and Figure 9) (Kim et al., 2013). In addition, other chaperones have been identified in

different subcellular compartments, which do not fit into this classification, some of which are under the control of stress response systems other than the general heat shock response (Allen et al., 2009; Kim et al., 2013; Quan and Bardwell, 2012). The various classes of chaperones have been found to work synergistically in the proteostasis network to ensure proper folding of synthesized or stress-denatured proteins inside cellular compartments (Balchin et al., 2016; Goemans et al., 2014). Hence, many chaperones are not only produced under stress conditions but are also essential for housekeeping functions. Although, as postulated by Anfinsen, any protein has the potential to fold into its native state on its own, it is now known that in the complex, crowded environment of the cell, the majority of proteins require the network of molecular chaperones in order to fold quantitatively on a biologically relevant time scale (Hartl, 1996; Kerner et al., 2005). In view of their importance for cellular health, it is important to understand the mechanism by which chaperones assist in protein folding.

1.3.1 Molecular chaperones and kinetic partitioning of folding proteins

Even though the various classes of chaperones are structurally very different, they all follow the same basic mechanistic principles of target recognition. By definition, a molecular chaperone is a protein or molecule that binds to and assists in the folding and/or assembly of another protein or protein complex without being part of its final structure (Figure 6) (Hartl, 1996). Chaperones target non-natively folded proteins by binding to hydrophobic segments and unstructured portions of the polypeptide backbone exposed by the substrate or client protein. Most proteins expose hydrophobic surfaces and regions of unstructured backbone during their synthesis, or because of cellular stress-induced unfolding, misfolding, or damage through oxidation. Not surprisingly, most chaperones have low client specificity (Balchin et al., 2016). As a result, the client binding site of most chaperones exhibits an inherent structural flexibility to accommodate client proteins of different size, shape, and folded state (Bardwell and Jakob, 2012; Hoffmann et al., 2010; Joachimiak et al., 2014; Mayer and Breton, 2015). In addition to their general affinity for hydrophobic surfaces, in many cases, complementary electrostatic interactions are an important factor for chaperone-client binding (Clerico et al., 2015; Joachimiak et al., 2014; Karagöz et al., 2014; Perrett et al., 1997). Apart from the chemical properties of client binding, different classes of chaperones recognize structurally different binding motifs on their client proteins, implying that the functional space of chaperones does not completely overlap. Generally, two different binding modes have been described for chaperones: (1) the chaperone recognizes defined short structural or sequence motifs on its client protein with high affinity, or (2) the client protein is bound as a structurally dynamic ensemble. In the first case, bound client segments experience low conformational flexibility, whereas in the latter case, the client protein interacts through many weak non-covalent bonds over a broad binding surface with the chaperone. This allows the client protein to undergo conformational changes while bound to the chaperone (Figure 5) (Burmann and Hiller, 2015). For example, Hsp70s bind to short unfolded stretches of a polypeptide chain through a small client binding cleft (Clerico et al., 2015), whereas Hsp90s usually target an

already partially folded state via a broad binding surface (Karagöz and Rüdiger, 2015; Mayer and Breton, 2015). This indicates that Hsp70 may operate earlier during de-novo protein folding than does Hsp90 (Balchin et al., 2016). Indeed, some Hsp70 homologues have been found to directly associate with the ribosome to aid in folding (Peisker et al., 2010).

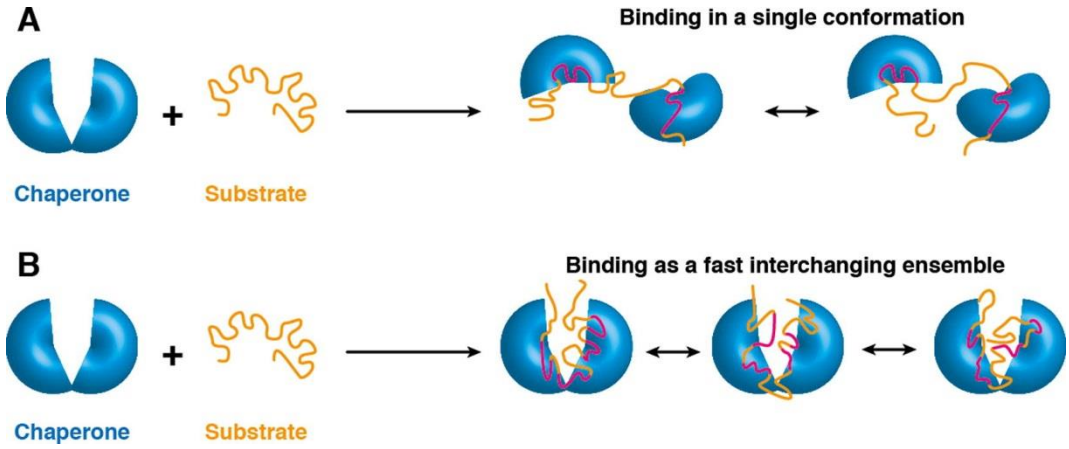

Figure 5: Two modes of chaperone-client interaction. (A) The chaperone (blue) recognizes distinct sites (magenta) on the client or substrate protein (orange). While unbound parts of the client protein are free to move and can possibly adopt a more folded state, bound segments exhibit low conformational flexibility defined by the client binding site on the chaperone. (B) Client protein interacts with the chaperone through many weak non-specific interactions through a broad surface, which allows the client protein to interchange conformations while bound to the chaperone (image taken from (Burmann and Hiller, 2015) with permission from the journal).

By shielding hydrophobic surfaces, chaperones prevent cell-toxic protein aggregation by lowering the concentration of aggregation-prone folding intermediates. Generally, it is thought that complete protein folding is only achieved after release of the client from the chaperone as hydrophobic surfaces bound by the chaperone presumably need to be buried inside the core for folding to be complete (Kim et al., 2013; Mattoo and Goloubinoff, 2014). In contrast, unbound parts of the client's polypeptide chain can presumably fold in the presence of the chaperone (Rüßmann et al., 2012; Walton et al., 2009). Optimal folding is achieved when client binding to the chaperone is faster than aggregation, and folding after chaperone release is faster than re-binding to the chaperone (Figure 6). Consistently, many chaperones have been documented to bind their non-native client proteins very rapidly (Fekkes et al., 1995; Joachimiak et al., 2014; Maier et al., 2001; Perrett et al., 1997). Such a kinetic portioning mechanism implies that chaperones prevent misfolding and aggregation of their client proteins, but do not catalyze the folding reaction themselves. Hence, to assist slow-folding or frequently misfolding proteins, chaperones either must undergo several binding and release cycles or, if folding fails, transfer these clients to other chaperones with different mechanistic properties (e.g., from Hsp70 to Hsp60). Client proteins that fail to fold to their native state entirely, even after multiple such cycles, are transferred to the degradation machinery, highlighting the interplay of different components of the PN (Kim et al., 2013). As a consequence, protein stability *in vivo* may be kinetically controlled, which counterbalances the inherently marginal thermodynamic stability of most proteins. This mechanism of

chaperone action implies that client release is a regulated process, which will be discussed in the next section.

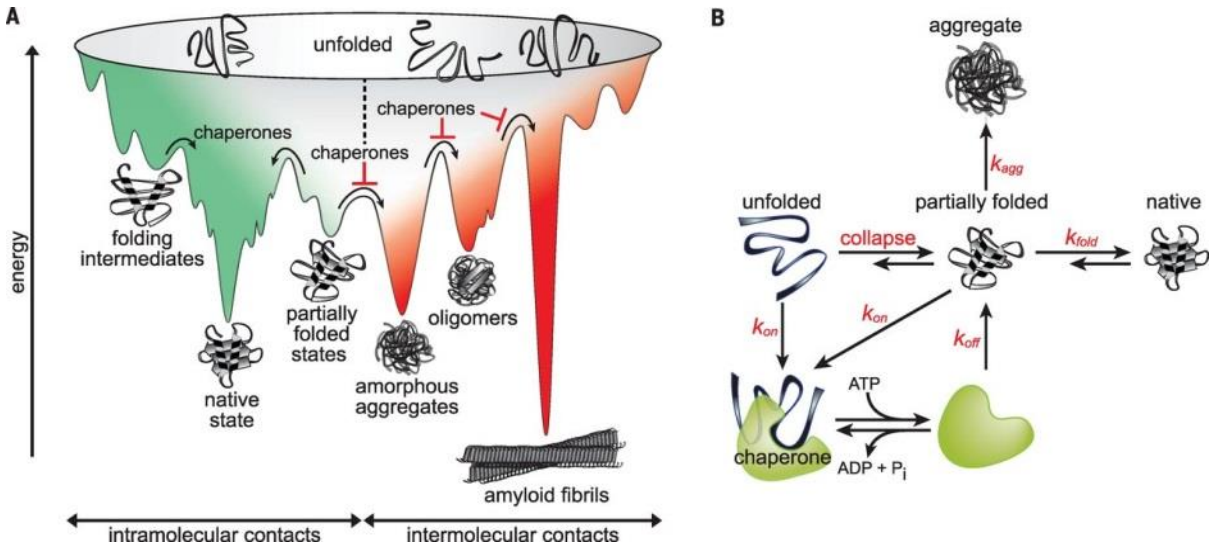

Figure 6: Folding helpers, such as molecular chaperones, prevent off-pathway misfolding and aggregation. Proteins are inherently vulnerable to misfolding and aggregation in the crowded environment of the cell. To ensure proteome integrity, cells invest a tremendous amount of energy in a network of folding helpers called the proteostasis network (PN). (A) A more realistic protein folding energy landscape without the assumption of total isolation (compare to Figure 1). While proteins may sample partially folded states along their folding pathways, or partially unfold due to a change in environmental parameters (green), they are prone to aggregation (red). Although most aggregates are amorphous, the misfolded state of some proteins tends to form thermodynamically very stable β-sheet-rich fibrillar aggregates called amyloid fibrils. Proteins folding via a multi-state transition that populate partially folded intermediate states are usually more aggregation-prone than two-state folders and thus are particularly vulnerable in the crowded environment of the cell. Molecular chaperones constitute the central part of the PN, as they prevent misfolding and assist in folding potentially by lowering energy barriers that trap folding proteins in a partially folded or misfolded state. In addition, they prevent unwanted intermolecular interactions, which result in cell-toxic aggregation (red). (B) Molecular chaperones generally assist in protein folding by kinetically sequestering non-natively folded proteins, which otherwise aggregate, also referred to as kinetic partitioning. Hence, efficient folding is achieved if the rate constant for folding (k_{fold}) is faster than binding to the chaperone (k_{on}) and binding is faster than aggregation (k_{agg}). The affinity of a chaperone is often regulated allosterically through conformational changes mediated either by ATP binding and hydrolysis or other external, often stress related factors. Hence, client protein release is regulated and is thought to be necessary for complete folding. (Figure adapted from (Balchin et al., 2016) with permission from the Journal).

1.3.2 Holding or folding? Regulation of chaperone function and client release

The fate of the bound client, whether it is assisted in folding or not, is determined by the class of chaperone it is bound to and the environmental conditions. Molecular chaperones can be divided into two categories based on whether they directly facilitate the folding of their client or not; they are thus

sometimes referred to as “foldases” and “holdases.” Since the “ase” suffix implies enzyme activity, a more appropriate terminology may be folding and holding chaperones (Mattoo and Goloubinoff, 2014; Winter and Jakob, 2004a). Although both types of chaperones recognize exposed hydrophobic surfaces on non-native proteins, folding chaperones reversibly bind their client proteins and hence guide their client proteins to the native state, sometimes through iterative release and rebinding cycles. Holding chaperones, on the other hand, bind very tightly to their client proteins. In doing so, they not only prevent aggregation but also prevent (re)folding. Hence holding chaperones exhibit strong anti-folding activity (Huang et al., 2016). To promote folding, holding chaperones necessarily have to transfer their bound client protein to a folding chaperone (Kim et al., 2013). Thus, client binding and release occurs in a coordinated fashion in most cases of folding and holding chaperones, namely through conformational changes that switch the chaperone between low and high client affinity states. In many cases, this switch is achieved in the high affinity state by the exposure of surface enriched in essential hydrophobic residues that recognizes and binds to the client protein. These residues are partially buried in the low affinity state, resulting in a decrease in the chaperone’s affinity for the client. Examples of folding chaperones that follow this mechanism are the Hsp60 chaperonin GroEL as well as the conditionally activated chaperone HdeA of *E. coli*. GroEL mediates this high to low affinity switch through ATP-driven conformational changes (Figure 7) (Saibil et al., 2013). In contrast, HdeA exposes a client binding site upon pH shift to acidic conditions (Foit et al., 2013). Similarly, some folding chaperones, especially in ATP-devoid cellular compartments, and many holding chaperones sense protein aggregation-promoting changes in the environment such as through shifts in temperature, pH, or redox potential, which increase their client binding affinity (Franzmann et al., 2008; Tapley et al., 2010; Teixeira et al., 2015; Winter and Jakob, 2004a). In addition, structural rearrangements can lock a client protein in the chaperone’s client binding site, as was observed for Hsp70 or bacterial Hsp90. Like Hsp60, such structural rearrangements are coordinated allosterically by ATP binding and hydrolysis (Figure 8 and Figure 9), which allows these chaperones to switch quickly between the high and low affinity states that are required to promote protein folding (Kim et al., 2013; Mayer, 2013; Mayer and Breton, 2015). These ATP-driven chaperone cycles are further regulated by other co-factors, many of which are holding chaperones (also often termed co-chaperones). For example, Hsp90 in the cytosol of mammalian cells cooperates with over 20 known co-chaperones (Figure 9) (Röhl et al., 2013). Differences in client specificity and regulation of expression of co-chaperones allow the cell to fine-tune client specificity and the capacity of their folding machines through co-chaperones. In addition, co-chaperones often alter the conformational equilibrium of their concomitant folding chaperone upon interaction either by stimulating the ATP-driven conformational cycles or arresting the chaperone in a specific conformation custom tailored to the needs of the particular client protein (Clerico et al., 2015; Hayer-Hartl et al., 2016; Röhl et al., 2013). Hence, post-translational regulation of folding chaperones is an essential feature that not only

allows a controlled client release, but also provides the cell with a tool to quickly respond to and fine-tune the chaperone response to different environmental conditions (Balchin et al., 2016).

Under conditions that cause excessive protein unfolding and aggregation, such as heat or oxidative stress, chaperone expression is upregulated through stress response pathways. Under these conditions, when proteome integrity is in danger and ATP levels are low, holding chaperones act as molecular vacuums that suck up unfolding proteins to buffer aggregation of the cellular proteome. Thus, chaperones keep their client proteins in a folding-competent state until the environmental conditions shift back to normal, at which point bound client proteins are slowly turned over to folding chaperones. Examples of holding chaperones are Hsp40s, which function as co-chaperones of Hsp70 foldases (Clerico et al., 2015), SecB, which is a member of the Sec pathway in *E. coli* that transfers unfolded protein to SecA (Huang et al., 2016), *E. coli* Hsp33 and eukaryotic Get3, which are redox regulated holding chaperones (Groitl et al., 2016; Voth et al., 2014), as well as the small heat shock proteins (sHsps) and peroxiredoxins, both of which become chaperone-active at elevated temperatures and have been shown to transfer their client proteins to Hsp70 foldases (Teixeira et al., 2015; Treweek et al., 2015).

The two categories of folding and holding chaperones are the extremes; in fact, there is a whole spectrum of chaperones that lie in between. Depending on the client protein and the environmental conditions, some chaperones switch from folding to holding or vice versa (Huang et al., 2016; Winter and Jakob, 2004a). For example, under heat stress, ATP levels are low, causing Hsp70 to switch from folding to holding. Similar observations were made for GroEL (Mattoo and Goloubinoff, 2014) In addition, though most folding chaperones are ATP-dependent, some ATP-independent chaperones have been found to exhibit folding activity as well. These include the bacterial chaperone Tigger Factor, which associates with the ribosome and hence functions very early on in de novo protein folding (Hoffmann et al., 2010). Other examples are the bacterial chaperones HdeA and Spy (Quan et al., 2011; Tapley et al., 2010). Spy is the major focus of this thesis and will therefore be discussed in great detail below.

1.3.3 Do folding chaperones accelerate client folding rates?

The extent to which folding chaperones such as the bacterial chaperonin GroEL or Hsp70 alter the folding energy landscape upon interaction with their client proteins and therefore affect folding is still a matter of debate (Ambrose et al., 2015; Clerico et al., 2015; Hayer-Hartl et al., 2016; Park et al., 2005; Tyagi et al., 2011). According to the kinetic partitioning mechanism, chaperones may be seen to act solely in a passive manner. This implies that they keep their client proteins in a folding-competent state by preventing misfolding and aggregation, and allowing sequential domain folding of larger proteins, but do not directly influence the folding energy landscape and therefore folding kinetics or pathways (Apetri and Horwich, 2008; Clare et al., 2012; Horst et al., 2007; Horwich and Fenton, 2009;

Horwich et al., 2009; Saibil et al., 2013). ATP consumption may then only act to regulate client release or transfer to another chaperone (Saibil et al., 2013). A prominent example is the Anfinsen's cage model for GroEL (Horwich et al., 2009). On the other hand, most ATP-consuming folding chaperones have been shown to pull proteins out of a thermodynamically stable misfolded state (Clerico et al., 2015; Lin et al., 2008; Priya et al., 2013a; Sharma et al., 2008, 2010; Shtilerman et al., 1999), for which they are sometimes also termed "unfoldases" (Mattoo and Goloubinoff, 2014). The unfoldase concept depicts chaperones as motors that unfold misfolded client proteins by iterative binding and release cycles that may require mechanical movements driven by energy consumption in the form of ATP. Binding to the chaperone is thought to catalytically lower the energy barrier for unfolding and thereby drive the misfolded client protein out of local energy minimum in which it is trapped. The partially unfolded client protein is then released to fold on its own in solution or to rebind to the chaperone, as outlined above (Mattoo and Goloubinoff, 2014). If binding of a client protein to a chaperone catalyzes unfolding, then it may also be conceivable that chaperones generally smooth the rugged folding energy landscape of proteins that fold through one or many on-pathway intermediate states. Chaperones may therefore act as catalysts that generally lower energy barriers and enhance folding kinetics/pathways (Figure 6). For example, there is evidence that the chaperone GroEL remodels the folding energy landscape of TIM barrel proteins, yielding increased folding rates (Georgescauld et al., 2014). However, it is not clear if the increased folding rate observed is due to a direct catalysis of TIM barrel folding by GroEL or due to catalyzed unfolding or even a simple prevention of the formation of kinetically stable misfolded states (Ambrose et al., 2015). The latter possibility suggests either the unfoldase or passive model. Similarly, Hsp90 is speculated to lower energy barriers and hence help its client proteins to interconvert between active and inactive conformations (Park et al., 2011a; Verba et al., 2016). Thus far, there is no evidence that Hsp70 affects the folding rates of its client proteins (Clerico et al., 2015).

In the context of this controversy, the next sections focus on the mechanistic and structural aspects of chaperone assisted protein folding exemplified by the classic folding chaperones Hsp60, Hsp70, and Hsp90, and by the molecular chaperone Spy, which is the focus of this thesis. Unlike the classic cytosolic folding chaperones, which are complex, large folding machineries regulated allosterically by ATP consumption and co-chaperones, Spy mediates client folding independently of co-factors (Quan et al., 2011). The following sections therefore highlight the mechanisms of client recognition and effects the chaperones have on protein folding, while the structural and mechanistic aspects of ATP-driven conformational changes are only mentioned briefly.

1.3.3.1 Hsp60

The Hsp60 family of chaperones, also called chaperonins, is found in all three branches of the tree of life (Kim et al., 2013) and can be divided into two groups based on sequence homology. Type I chaperonins are encoded in the genome of most bacteria and in the endosymbiotic organelles of

eukaryotes, whereas type II chaperonins facilitate protein folding in most archaea and in the eukaryotic cytosol. Although they show significant differences in primary sequence, both types are structurally conserved and share many mechanistic similarities. Both form oligomeric back-to-back stacked double-ring structures that provide two chambers with unique properties that allow the client protein to fold in isolation, thereby avoiding unwanted intermolecular interactions with the cellular proteome (Figure 7). In both types, each monomer consists of three domains. The N-terminal apical domain mediates recognition and binding of the client protein and exhibits the most structural differences between the two types. The C-terminal ATPase domain is connected to the apical domain via an intermediate or hinge domain (Hayer-Hartl et al., 2016; Lopez et al., 2015). The ATPase domain, also called the equatorial domain due to its position at the interface of both rings, harbors most of the interaction surfaces that stabilize the oligomeric double ring structure and mediate the allosterically-controlled conformational changes of the chaperone cycle driven by ATP binding and hydrolysis (Hayer-Hartl et al., 2016; Lopez et al., 2015). The bacterial type I chaperonins (e.g., *E. coli* GroEL) and some archaeal type II chaperonins (also called thermosomes) are built of 7 and 8 or 9 identical monomers, respectively (Hayer-Hartl et al., 2016; Lopez et al., 2015). On the other hand, each ring of the eukaryotic chaperonin TRiC (TCP-1 Ring Complex, also called CCT (Chaperonin Containing TCP-1 complex)) contains 8 paralogous monomers that provide asymmetry in all steps of the chaperone cycle, crucial for the folding of increasingly complex eukaryotic genomes (Figure 7 D) (Joachimiak et al., 2014).

The ATP-driven chaperone cycle of both types of chaperonins shares certain similarities, as shown in Figure 7. Binding of non-native client proteins is mediated through patches of hydrophobic residues exposed in the apical domains of an empty ring. Hydrophobic interactions between chaperonin and client are accompanied by electrostatic interactions in both types, as both exhibit a negative net charge (Joachimiak et al., 2014; Perrett et al., 1997). After client binding, ATP binding and hydrolysis in the equatorial domains triggers a conformational change in the apical domain mediated through the intermediate domain, which lowers client affinity and therefore releases the client protein into the chamber. At the same time, the chamber closes. Relatively slow ATP hydrolysis gives the client protein time to fold inside the chamber. After completion of ATP hydrolysis, ADP and client are released. Client rebinding may occur if folding is incomplete (Hayer-Hartl et al., 2016; Joachimiak et al., 2014). The chamber of both chaperonin types is large enough to encapsulate client proteins up to 60 kDa (Hayer-Hartl et al., 2016; Lopez et al., 2015). Larger client proteins may still use the chaperonin system by binding to the apical client binding sites, which has been shown to facilitate folding outside the chamber (Chaudhuri et al., 2009; Priya et al., 2013a). In addition, partial encapsulation of larger multidomain proteins has been reported for TRiC, which allows isolated folding of domains separately (Roh et al., 2016a; Rüßmann et al., 2012). Apart from hydrophobicity-mediated client interaction, other commonalities of client binding and release have been reported for both chaperonin types. Several hydrophobic segments of the client may bind to several or all apical

domains simultaneously, which has been reported to partially unfold the client protein, potentially by binding to less structured states selected out of the flexible client ensemble (Farr et al., 2000; Priya et al., 2013a; Villebeck et al., 2007). Client unfolding may be further fostered through conformational changes triggered by ATP binding (Lin et al., 2008). This may pull the client protein out of a kinetically trapped misfolded state. In addition, client release into the chamber may occur sequentially. In the case of GroEL, bound segments with increased hydrophobicity may get released later than segments with less hydrophobicity. This controlled release may delay the hydrophobic collapse and hence the formation of non-native hydrophobic interactions and therefore misfolding (Chen et al., 2013; Motojima, 2015). As mentioned above, TRiC adds another layer to control client release as each monomer exhibits client binding sites with distinct sets of charged and hydrophilic residues surrounding the hydrophobic binding patch. This allows selective binding of distinct client segments and hence a pre-orientation of the bound client protein, potentially mediating a productive sequential folding upon release (Joachimiak et al., 2014; Roh et al., 2016b; Tam et al., 2009). Differences in ATP binding affinity of each of the 8 TRiC monomers (4 with high affinity and 4 with low affinity, see Figure 7 D) may lead to timely delayed conformational changes and hence may facilitate an ordered release of client segments into the chamber, again potentially avoiding misfolding of topologically complex client proteins (Jiang et al., 2011; Reissmann et al., 2012). Similarly to Hsp90 (see below), TRiC has also been shown to fulfill regulatory functions by interacting with natively folded proteins, thereby inhibiting protein function (Svanström and Grantham, 2016).

While ATP-driven conformational cycles and mechanisms of client recognition are rather well established, what happens inside the chamber is less clear. The interior lining of the cavity of both GroEL and TRiC is hydrophilic in the closed state. Whereas the GroEL cavity exhibits an overall negative net charge, the interior wall of TRiC forms a gradient of positive to negative net charge from one side of the chamber to the other (Chaudhry et al., 2003; Leitner et al., 2012). Interestingly, conformational changes that lead to the closure of the GroEL cavity and ejection of the client protein into the chamber, bury most of the hydrophobic residues in the apical domain involved in initial client binding (Clare et al., 2012). In contrast, the conformational changes inducing cage closure of TRiC only minorly affect the apical client binding site, making it available for client binding even in the closed conformation (Dekker et al., 2011; Joachimiak et al., 2014). Not much is known about the interactions of the encapsulated client protein with the chaperonin wall; however, its charged nature is thought to drive the formation of a hydrophobic core and minimize the interaction of the encapsulated polypeptide chain with the cage wall (Gupta et al., 2014). Experimental evidence from electron

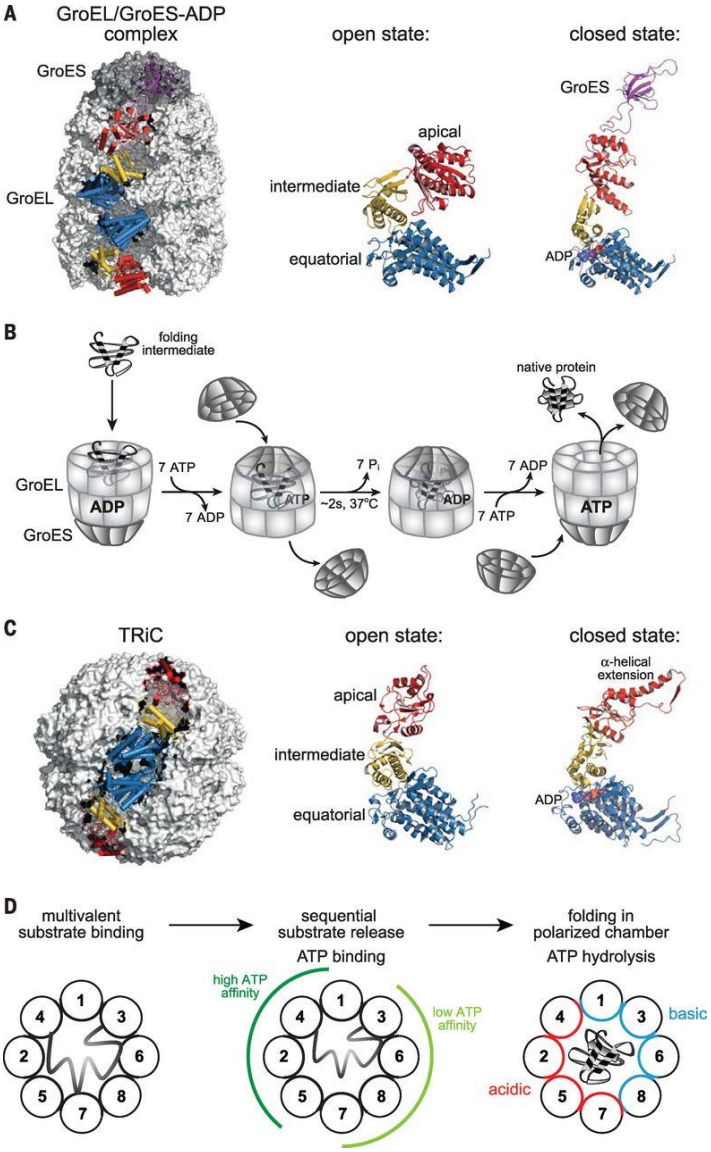

Figure 7: The two classes of Hsp60 form folding cages. (A, C) Both the bacterial chaperonin type I tetradecameric GroEL (A, PDB ID: 1PFQ) and the eukaryotic chaperonin type II hexadecameric TRiC (C, PDB ID: 4V94) form double-ring structures that undergo conformational changes driven by ATP binding and hydrolysis (B, D) in order to bind and encapsulate non-natively folded client proteins. The ADP-bound state closed conformation is shown on the left in the grey surface representation. One monomer of each is shown in the schematic representation, with domains colored in red (apical), yellow (intermediate), and blue (equatorial) from N to C terminus. Cartoons of the isolated monomers in both the open and closed conformations are shown in the middle and right to illustrate the conformational changes mediated by ATP binding and hydrolysis. ADP bound to the equatorial domain in the closed state is shown in blue (adenosine) and red (phosphates) spheres. Note that the bacterial chaperonin GroEL forms a complex with its co-chaperone GroES (purple), which functions as a lid to encapsulate the client protein in the closed state. The eukaryotic TRiC, on the other hand, contains a built in lid in the form of an α-helical extension in the apical domain. (B) ATP-driven chaperone cycle of GroEL. Binding of non-natively folded client to the apical domains of an empty heptameric ring triggers ATP and GroES binding to the same ring (cis ring), which releases the client protein into the chamber. Client folding occurs inside the chamber until all ATP molecules in the equatorial domains are hydrolyzed. This allows ATP binding to the opposite ring (trans ring), which as a result triggers ADP, GroES, and client release in the cis ring. Client rebinding may occur if folding was unsuccessful. (D) ATP-driven chaperone cycle of TRiC. Each octameric ring of the hexadecamer consists of eight sequentially different but homologous monomers, offering client binding sites in the apical domains with distinct chemical properties. Depending on the chemical properties of the client protein, a non-native client protein may interact with a distinct subset of apical domains (left). Following client binding, ATP binds to the equatorial domains in a sequential fashion due to differences in ATP binding affinity. This triggers sequential conformational changes that result in a sequential client release into the chamber and closing of the chamber (middle). Folding inside the chamber may occur during ATP hydrolysis. The polarized environment of the chamber may aid in client folding (right). (Figure adapted from (Balchin et al., 2016) with permission from the Journal).

microscope and X-ray structures as well as single molecule spectroscopy conducted with GroEL and TRiC in the closed state suggests interactions of the chaperonin wall with folding intermediates of certain client proteins (Chen et al., 2013; Dekker et al., 2011; Hofmann et al., 2010). These interactions may improve client folding yields but sacrifice folding rates. In addition, the inherent asymmetry of TRiC has been shown to be essential for successful folding of specific eukaryotic proteins such as actin, which can't be folded by GroEL (Tian et al., 1995).

About 10% of the *E. coli* and yeast proteomes have been reported to utilize the chaperonins for folding (Balchin et al., 2016). Although there are no clear binding motifs, these client proteins share some very broad overall structural similarities. For instance, many of them have complex topologies that are stabilized by many long range contacts. As a result, many chaperonin substrates have rugged folding landscapes in which kinetically trapped folding intermediates and misfolded states are frequently populated (Azia et al., 2012; Kerner et al., 2005; Narayanan et al., 2016; Yam et al., 2008). It is postulated that encapsulation may enhance folding rates for these complex folders. However, whether or not client encapsulation affects the folding landscape of a chaperone client and therefore enhances folding rates is not yet clear. Three models of chaperone action for GroEL have been proposed. In the first model, GroEL may act passively by completely isolating the folding client protein and thereby preventing aggregation. This model is supported by data collected for a number of proteins, showing that folding kinetics are independent of whether folding takes place or not while encapsulated or while free in solution (Apetri and Horwich, 2008; Brinker et al., 2001; Horwich et al., 2009; Tyagi et al., 2011). The second model proposes that spatial confinement and electrostatic repulsion of GroEL's interior may enhance client folding rates. Such an enhancement in rate has been shown experimentally for proteins that populate entropically stabilized (i.e., flexible) intermediate states with higher probability. In these cases, spatial confinement may reduce the entropic penalty associated with folding by destabilizing flexible intermediates (see Table 1). In addition, the flexible C termini of each subunit pointing inside the chamber may entropically support the folding process via entropic transfer (Brinker et al., 2001; Chakraborty et al., 2010; Georgescauld et al., 2014; Gupta et al., 2014; Weaver and Rye, 2014; Zhang and Kelly, 2014). In contrast, the third model proposes that iterative binding and release of the client protein to the apical domains may enhance folding rates by pulling client proteins out of an enthalpically stable misfolded state. This mechanism may especially apply to proteins that are too large to enter the folding chamber (Priya et al., 2013a). In a fourth potential model, protein folding yields may be improved via interaction of the folding client protein with the interior wall of the chaperonin in the closed conformation, slowing down folding (Hofmann et al., 2010; Sirur and Best, 2013).

1.3.3.2 Hsp70

Hsp70 is conserved both in prokaryotes and eukaryotes. It is the most versatile chaperone, involved not only in protein (re)folding of the major part of the proteome, but also in cellular trafficking,

protein aggregate disassembly, protein degradation, etc. Hence, depending on the organism, the genome may encode multiple paralogues of Hsp70. Hsp70s function in conjunction with a set of co-chaperones, collectively called Hsp40s or J-proteins, since they all contain a J-domain that is required for interaction with Hsp70 (Clerico et al., 2015). As each Hsp40 exhibits a somewhat distinct client specificity, Hsp70s client affinity is fine-tuned by its co-chaperones. In addition, Hsp70 requires nucleotide exchange factors, which trigger ADP to ATP exchange. The number of Hsp70 paralogues, J-proteins, and nucleotide exchange factors increases from prokaryotes to higher eukaryotes, reflecting the increased complexity of the proteome. In human cells, for example, about 50 J-proteins have been identified so far (Kampinga and Craig, 2010). Despite the apparent versatility of Hsp70s and co-factors, all Hsp70s are structurally conserved and follow basically the same mechanistic principles (Figure 8). Hsp70s usually consist of two domains, an ATPase domain and a substrate binding domain (SBD). The SBD contains a β-sheet-rich N-terminal subdomain that recognizes short stretches of unfolded backbone of 5 to 7 amino acids exposed in non-native client proteins. This binding motif is enriched with hydrophobic residues often flanked by basic residues, indicating that electrostatic interactions are important for binding as well (Rüdiger et al., 1997). In addition, the SBD contains a C-terminal α-helical subdomain. Opening and closing of this α-helical subdomain is mediated through ATP binding and hydrolysis (Figure 8) in the ATPase domain (Zhuravleva and Gierasch, 2015). It was previously thought that Hsp70 predominantly recognizes short unfolded polypeptide stretches and that its affinities for these binding motifs is regulated through the opening and closing motions of this α-helical lid (Figure 8) (Clerico et al., 2015; Mayer, 2013; Zhuravleva and Gierasch, 2015). However, recent work suggests that the α-helical subdomain may not only close upon an unfolded polypeptide stretch, but rather also mediate contacts to more folded portions of the client protein, hence allowing Hsp70 to bind folding intermediates and even near-native conformations (Mashaghi et al., 2016) Under cellular stress such as heat shock, ATP levels are low and hence Hsp70 stays bound to its client protein in an ADP-mediated closed conformation, thereby preventing aggregation (Clerico et al., 2015; Mashaghi et al., 2016; Mayer, 2013; Zhuravleva and Gierasch, 2015). Alternatively to nucleotide exchange factors, attenuator proteins exist in higher eukaryotes that can arrest Hsp70 in a ADP-bound state, delaying client release for further downstream folding and transfer to another chaperone or degradation (Li et al., 2013). In addition, some Hsp70s have been identified that contain an inactive ATPase domain. These Hsp70s don’t switch between high and low affinity states and hence exhibit fast client binding and release rates. An example is Ssz1, which directly associates with the ribosome in yeast (Leidig et al., 2013).

Most high-resolution structural work on Hsp70 has been conducted with short model peptides from which general sequence preferences could be derived. These may vary somewhat from Hsp70 homologue to homologue, but most of the client specificity is dictated by the accompanying Hsp40, as described above. Hsp70 binding motifs are found in most proteins and are usually buried in the hydrophobic core of natively folded proteins. However, they are exposed during protein synthesis and

under denaturing conditions such as heat or oxidative damage, and therefore are an indicator of protein folding stress (Clerico et al., 2015).

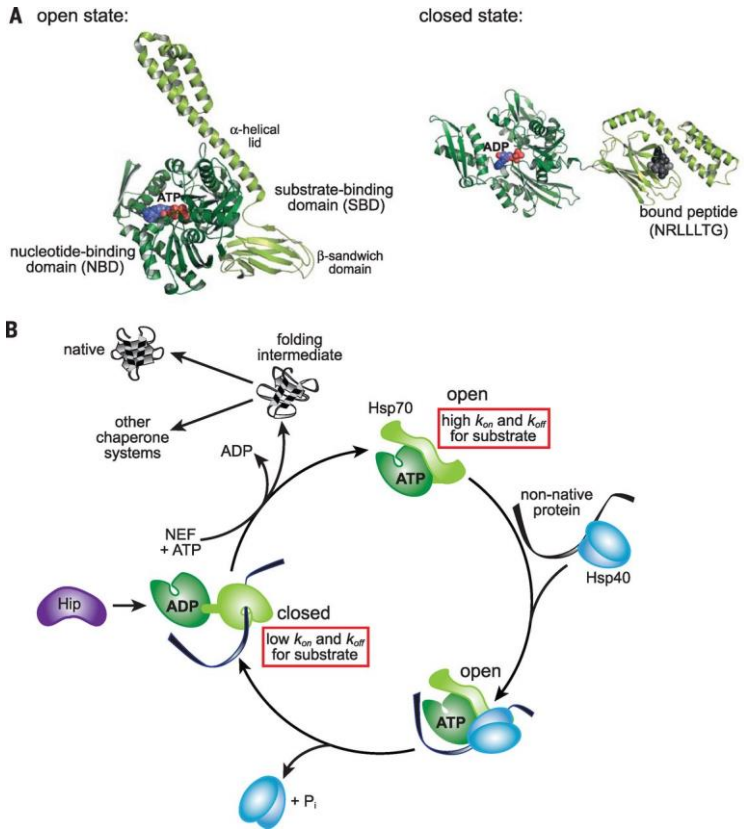

Figure 8: Conformational cycle of Hsp70. (A) Open (left, PDB ID: 4B9Q) and closed (right, PDB ID: 2KHO) conformations of Hsp70 are shown as cartoons with bound nucleotide (ATP, left; ADP, right) depicted as blue (adenosine) and red (phosphates) spheres. The ATPase domain is colored in dark green, and the substrate binding domain (SBD) is colored in light green. In the closed state, a model peptide with the sequence NRLLTG (grew surface representation) is bound to the SBD. (B) Conformational cycle of Hsp70 that mediates client binding and release, controlled by co-chaperones Hsp40 and nucleotide exchange factors (NEF). The open conformation allows fast binding and release of the client protein, resulting in a relatively low client affinity. Once the Hsp40-client complex is bound to the ATP-bound state of Hsp70, ATP hydrolysis is triggered, which causes the α-helical subdomain to close over the bound polypeptide stretch, leading to the closed conformation of Hsp70. In this conformation, client binding and release is slow, yielding high client affinity. The ATPase activity has been reported to be higher with increased affinity to the bound polypeptide stretch and is further enhanced in the presence of an Hsp40 co-chaperone. Image taken from (Balchin et al., 2016). (Figure adapted from (Balchin et al., 2016) with permission from the Journal).

A high-resolution structure of a complex of Hsp70 with a bound client protein has not yet been obtained. Although there is evidence that Hsp70 recognizes the same binding motifs in full-length proteins as was determined in peptide studies, how binding affects the client protein is not well understood. Hsp70s are thought to undergo consecutive bind and release cycles with their client proteins, in which binding and release kinetics are custom tailored to the needs of the particular folding client protein through ATP and co-factors (Figure 8) (Clerico et al., 2015; Mashaghi et al.,

2016). The unstructured C terminus of the SBD may function as a second client binding site to keep the client in close proximity, increasing the chance of client rebinding if necessary. Binding of Hsp70 is thought to keep the client protein in a less structured state, thereby inhibiting misfolding and allowing the client protein to explore and form native-like secondary structure before a global hydrophobic collapse, either during protein synthesis or post stress-induced denaturation (Clerico et al., 2015). Very recently however, evidence has emerged that Hsp70 also recognizes folding intermediates and even native-like conformations of its client proteins. In addition, client folding may at least partially while bound to Hsp70 (Mashaghi et al., 2016). There is also evidence that Hsp70s can unfold misfolded or even folded proteins through selectively binding to conformations that transiently expose hydrophobic binding motives, shifting the folding equilibrium to more unfolded conformations and thereby remodeling the folding energy landscape (Clerico et al., 2015; Marcinowski et al., 2013; Mattoo and Goloubinoff, 2014).

1.3.3.3 Hsp90

Hsp90 is a homodimeric chaperone conserved in many bacteria and eukaryotes (Balchin et al., 2016). Each monomer consists of three domains: and N-terminal ATPase domain (NTD, green), a middle domain (MD, orange), and a C-terminal dimerization domain (CTD, blue). Hsp90 provides a large surface for client binding, mainly provided by the MD, but depending on the client, the binding site may stretch throughout all three domains. The client binding site is enriched in dispersed hydrophobic residues and exhibits an overall net negative charge, which is important for client binding (Karagöz and Rüdiger, 2015).

Bacteria contain a single Hsp90 homologue, which functions independently of known co-chaperones. Hence, for bacterial Hsp90, ATP binding and hydrolysis and client binding determine the conformational cycle (Mayer and Breton, 2015). In contrast, the eukaryote yeast has two cytosolic Hsp90 paralogues, one of which is heat inducible. Eukaryotic Hsp90 assists in client folding through its interaction with many co-chaperones and with the folding chaperone Hsp70, which regulate Hsp90's conformational equilibrium rather than solely affecting ATP binding and hydrolysis (Figure 9) (Balchin et al., 2016; Mayer and Breton, 2015; Röhl et al., 2013). The co-chaperones act sequentially along the conformational cycle of Hsp90 by stabilizing certain conformations and enhancing Hsp90's client specificity (Figure 9 B) (Röhl et al., 2013). Both bacterial and eukaryotic Hsp90 appear to preferentially bind partially folded proteins, but also exhibit some affinity for

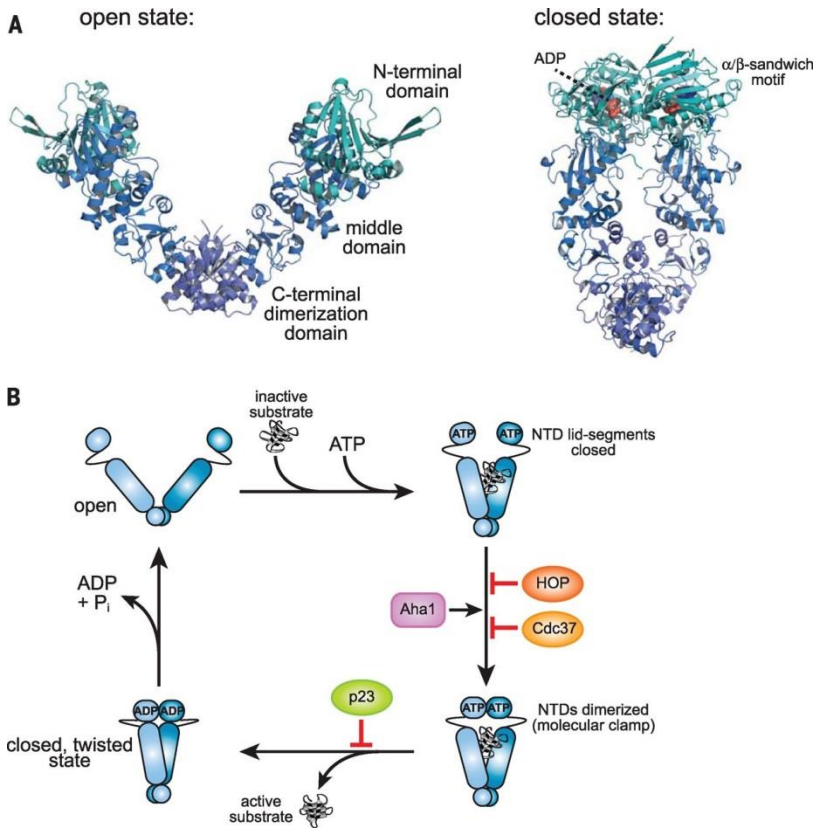

Figure 9: Hsp90 is a flexible homodimer. (A) Open (left, bacterial, PDB ID: 2IOQ) and closed (right, yeast PDB ID: 2CG9) conformations of Hsp90 are shown in cartoon representation with bound nucleotide (ADP, right) depicted as blue (adenosine) and red (phosphates) spheres. The N-terminal ATPase domain (NTD) is colored in cyan, the middle domain (MD) in light blue, and the C-terminal dimerization domain (CTD) is colored in dark blue. In the ATP-free state, both monomers of Hsp90 exhibit a high degree of conformational flexibility. This apparent flexibility allows Hsp90 to accommodate clients of different sizes and conformational states. (B) The ATP-dependent conformational cycle of eukaryotic Hsp90 is regulated by co-chaperones and client binding. Client and ATP binding to the open conformation trigger Hsp90 to switch to a closed conformation, resulting in dimerization of the NTDs and increased ATPase activity. This step can be slowed down by the co-chaperones HOP and Cdc37. Although functionally different, both co-chaperones stabilize the open conformation of Hsp90, thereby delaying ATP hydrolysis, which is important for client binding (assisted through Cdc37) or transfer from Hsp70 (mediated through HOP). In contrast, Aha1 mediates transitions to the closed conformation and thereby fosters ATP hydrolysis. P23 acts late in the conformational cycle. It stabilizes the closed conformation after client binding, but inhibits ATPase activity. This keeps the Hsp90-client complex in the closed state and delays client release, which is important for the maturation of some client proteins. Note that there are many more co-factors than illustrated here. (Figure adapted from (Balchin et al., 2016) with permission from the Journal).

unfolded and near-natively folded states. Hence, there is no common predictable binding motif known for Hsp90 to date (Karagöz and Rüdiger, 2015; Mayer and Breton, 2015). Although eukaryotic Hsp90 assists in the folding of many proteins, it is especially important for the folding and conformational maturation and maintenance of proteins involved in signal transduction and metabolism. Many of these client proteins exhibit low intrinsic thermodynamic stability, and therefore unfold easily. This

tendency to unfold easily may be one criterion determining Hsp90's client specificity (Balchin et al., 2016; Taipale et al., 2012). For example, binding of Hsp90 to folded p53 shifts p53's folding equilibrium towards a less folded state. Chaperone-client contacts are mediated through multiple weak non-specific and transient interactions along the broad binding surface on Hsp90, and p53 appears to be rather dynamic while bound (Park et al., 2011a, 2011b). In addition, a recent high-resolution cryo-EM structure illustrates the interaction of human Hsp90 assisted by the co-chaperone Cdc37 with the thermodynamically marginally stable client protein kinase Cdk4. The complex reveals interactions of unfolded and partially folded segments of the kinase with Hsp90. Both studies demonstrate Hsp90's capacity to accommodate both unfolded and partially folded states. Cdc37 recruits Cdk4 to Hsp90 and stabilizes it in the complex with Hsp90 (Verba et al., 2016). Hsp90 in complex with the intrinsically disordered protein Tau, recently solved via NMR, reveals that Hsp90 recognizes a distinct stretch on Tau that is enriched in hydrophobic residues and positive charge (Karagöz et al., 2014). These examples demonstrate how the large amphiphilic client binding surface allows the accommodation of client proteins with different structures and dynamics.

1.4 The molecular chaperone Spy

Despite three decades since the discovery of molecular chaperones, whether chaperones actually catalyze protein folding or not, is still debated. This lack of knowledge is in part due to the structural and mechanistic complexity of the classic folding chaperones Hsp60, Hsp70, or Hsp90, as well as their dependence on co-factors (Figure 7, Figure 8, and Figure 9). Together these complicating factors have hampered detailed, atomic-level analyses of the specific interactions that occur between a client protein and its chaperone during the process of folding (Clerico et al., 2015; Hayer-Hartl et al., 2016; Karagöz and Rüdiger, 2015). Atomic-resolution data would provide insights into the interactions between a chaperone and a client protein as well as the effects of chaperones on the client's folding landscape. Contrarily to the large folding machineries of the cytosol, some ATP-independent chaperones are structurally simpler and yet still exhibit folding activity. These properties render these chaperones valuable for conducting detailed, high-resolution studies of the effect that folding chaperones have on protein folding.

One of these ATP-independent folding chaperones is the stress response protein Spy. Spy is part of the proteostasis network of the *E. coli* periplasm (Allen et al., 2009; Goemans et al., 2014). Its chaperone activity was first discovered via a genetic selection system developed by Bardwell and co-workers that directly links the *in vivo* folding stability of a protein to antibiotic resistance (Foit et al., 2009; Quan et al., 2011). By giving *E. coli* the choice between death and stabilizing an unstable mutant of the protein Im7 (L53A I54A), which is trapped in its folding intermediate state (Figure 3) (Capaldi et al., 2002), the cells responded by dramatically enhancing the expression of the periplasmic protein Spy (Quan et al., 2011). Spy is a 16 kDa soluble periplasmic protein that is highly overexpressed in response to spheroblast formation, which led to its name, Spheroblast Protein Y (Hagenmaier et al., 1997). In *E.*

coli, Spy's expression is controlled by three of the five envelope stress response pathways. It is strongly controlled by the Cpx, Bae envelope stress response pathways, and to a more minor extent by the Rcs regulatory system. All of these regulatory systems sense protein unfolding and membrane or cell wall related stresses (Bury-Moné et al., 2009; Raffa and Raivio, 2002; Srivastava et al., 2014; Vogt and Raivio, 2012). Spy is induced under folding stress conditions, for instance by disruption of membrane integrity by antimicrobial substances that target the envelope (Fränzel et al., 2012; Raivio et al., 2000). In addition, externally added compounds that cause protein precipitation or unfolding in the periplasm induce Spy expression up to 500 fold. These compounds include tannins, which are anti-microbial plant-derived secondary metabolites that are known to cause protein aggregation (D. Zanchi et al, 2008; Scalbert, 1991; Zoetendal et al., 2008), as well as other naturally occurring compounds including zinc and protein denaturants such as butanol, or ethanol (Bury-Moné et al., 2009; Rutherford et al., 2010; Srivastava et al., 2014; Wang and Fierke, 2013; Yamamoto et al., 2008). Not surprisingly, *E. coli* strains lacking Spy are sensitive to tannin exposure (Quan et al., 2011). Spy expression is also upregulated upon over-expression of certain envelope proteins which easily misfold (Raffa and Raivio, 2002). Interestingly, Spy over-expression has also been shown to compensate for lethal protein folding defects found in cells that lack the periplasmic chaperones SurA and Skp (Schwalm et al., 2013). Deletion of the *spy* gene, on the other hand, upregulates the expression of DegP, which is a periplasmic chaperone and protease, as well as RpoH, the general heat shock response transcription factor (Raivio et al., 2000). Proteins homologous to Spy appear in a wide range of different enterobacteria, protobacteria, and some cyanobacteria, underlining its importance in the envelope stress response of gram negative bacteria (Quan et al., 2011; Srivastava et al., 2014). In addition, *E. coli* expresses at least two Spy paralogues, namely CpxP and Zrap. CpxP is the periplasmic repressor of the Cpx stress response pathway, which also regulates Spy expression as mentioned above. Under ideal growth conditions, CpxP is bound to the periplasmic domain of the transmembrane protein CpxA in order to suppress the Cpx stress response. Under conditions which induce unfolding, CpxP binds to unfolded proteins which compete with CpxA for CpxP binding. Release of CpxP from CpxA activates the Cpx stress response (Vogt and Raivio, 2012). Consistent with its capacity to bind unfolded proteins, CpxP has been shown to have chaperone activity (Quan et al., 2011). Zrap, like Spy and CpxP, is a periplasmic protein. Zrap has been shown to form zinc binding decameric barrels and is an important resistance factor against antimicrobial peptides (Appia-Ayme et al., 2012). Similar to Spy, Zrap exhibits chaperone function (Appia-Ayme et al., 2012).

1.4.1 Spy functions as an ATP- and cofactor-independent folding chaperone *in vitro*

Unlike the folding chaperones described above that require ATP and large networks of co-chaperones to facilitate protein folding (Figure 7, Figure 8, and Figure 9) (Clerico et al., 2015; Hayer-Hartl et al., 2016; Mayer and Breton, 2015), Spy acts independently of ATP- and co-factors, at least *in vitro*. This

independence is consistent with Spy's physiological function as a folding chaperone in the ATP-devoid periplasm of *E. coli.* Like other folding chaperones, Spy has broad substrate specificity. It not only increases the stability of an instable Im7 mutant *in vivo*, but unlike most other ATP independent chaperones it also prevents the aggregation and promotes proper refolding of a diverse set of proteins in the presence of a variety of protein denaturants *in vitro*. Spy prevents the inactivation of the periplasmic proteins DsbB, alkaline phosphatase as well as the enzyme aldolase in presence of the naturally occurring protein denaturant tannic acid (Quan et al., 2011). In addition, Spy prevents the aggregation and promotes refolding of malate dehydrogenase, aldolase, lactate dehydrogenase, and alpha-lactalbumin, denatured either thermally or chemically via urea, guanidine, or dithiothreitol (in case of alpha-lactalbumin), as well as the naturally occurring denaturant, ethanol (Appia-Ayme et al., 2012; Quan et al., 2011, 2014). The fact that all these chemicals follow a different mechanism to unfold proteins suggests that Spy has a general affinity for non-natively folded proteins. Consistently, Spy also forms stable complexes with bovine α-casein, a disordered protein commonly used to study chaperone-client interactions (Lin et al., 1995; Nam and Walsh, 2003; Quan et al., 2011). α-casein contains residual secondary and tertiary structure and constitutively exposes hydrophobic surfaces, and thus is similar to an intermediate folding state (Creamer et al., 1981; Hoagland et al., 2001). The apparent affinity for non-native states in combination with the structural diversity of protected client proteins is strong evidence that Spy has broad substrate specificity. The fact that Spy overexpression alone is sufficient to stabilize an instable Im7 mutant *in vivo* in the periplasm, which is an ATP devoid environment suggests that Spy's chaperone function is intrinsic, i.e., independent of regulatory cofactors such as ATP, *in vitro* as well as *in vivo* (Quan et al., 2011). However, these results do not exclude the possibility that Spy interacts with other folding helpers or cofactors in the periplasm *in vivo* to assist in protein folding. In addition to its ability to stabilize folding intermediates *in vivo*, Spy has also been shown to inhibit the formation of cross-β-sheet containing functional amyloids *in vitro* and *in vivo* (Evans et al., 2011).

1.4.2 Spy serves as a molecular cradle for folding proteins

Spy is a rather effective but ATP-independent folding chaperone, but how does a protein as small as 16 kDa protein assist such a broad clientele in folding? Some clues may come from examining its crystal structure. Spy is an α-helical rich homodimer with a unique, thin cradle-like shape with an average thickness of 9.2 Å (Figure 10 A)(Kwon et al., 2010; Quan et al., 2011). Each Spy monomer contains 4 helices as well as unstructured N- and C-termini (which are not visible in the crystal structure). The two monomers are oriented in an antiparallel fashion forming a coiled coil interaction with helix 3 which buries a surface of 1850 $Å^2$ at the dimer interface. This large interaction surface results in a high affinity for dimerization and stabilizes the rather unusual cradle-like shape of Spy (Kwon et al., 2010; Quan et al., 2011). Due to the lack of a globular core, Spy exposes a rather large number of residues compared to globular folded proteins. Both sides of the cradle are chemically

distinct. While the convex side of the cradle exposes an even distribution of positive and negative charge as well as polar but little hydrophobic residues, the convex side contains two hydrophobic patches per half of the symmetric homodimer and in addition an excess of positive charge surrounding the hydrophobic patches (Figure 10). The hydrophobic patch P1 is formed by residues between helix 1 and helix 2 while the hydrophobic patch P2 is located at the interface of the coiled coil forming helix 3 of each monomer (Kwon et al., 2010; Quan et al., 2011). The amphiphilic nature of the convex site of Spy seems to be ideally poised to bind non-natively folded proteins that expose hydrophobic residues. Indeed, investigations of the client binding site on Spy via cross-linking using a truncated version of Im7 (amino acid 6 to 45, which contains the entire Spy binding site) suggest that Im7 binding occurs on the concave side of Spy as none of the amino acid residues exposed on the convex side of Spy were found to be cross-linked with this truncated version of Im7 (Quan et al., 2014). Zero-length cross-linker reagents crosslinked Im7 to K39 and Y104 on Spy, implying that these Spy residues are directly involved in the interaction with the Im7 peptide. Both residues are near Spy's hydrophobic patch P1 (Figure 10 B). Hydrogen-deuterium exchange, limited proteolysis experiments and fluorescence labeling approaches indicate that Im7 and casein binding occurs over the entire surface of the convex side including the rim as well as the unstructured termini as all of these areas are more protected against hydrogen exchange or proteolysis and reveal changes in fluorescence in presence of the Im7 or casein (Quan et al., 2011, 2014). This suggests that in addition to the very broad binding surface of the cradle, the unstructured termini may play a role in accommodating client proteins of different sizes and folded status perhaps by flexibly expanding the surface available for interaction. These experiments also revealed that in addition to the flexible termini, the rim of the Spy cradle exhibits high flexibility. This includes helices 1 and 4 and the unstructured linker between helix 1 and 2. These regions are also possibly involved in generating a broad client specificity through adaptive client binding (Quan et al., 2011, 2014). The Spy dimer forms stoichiometric 1 to 1 complexes with Im7, consistent with the notion of a broad and flexible binding site (Quan et al., 2014).

As outlined above, most ATP-independent chaperones are holding chaperones, implying that they need to partner with a folding chaperone in order to release their bound client protein and facilitate its folding (Balchin et al., 2016). However, Spy assists in protein refolding in absence of cofactors. This demands that Spy releases its client proteins with a reasonable rate. Kinetic measurements using bio-layer interferometry (BLI) revealed that Spy binds reversibly to partially folded Im7 L53A I54A, which was immobilized on a chip (Quan et al., 2014). The absolute numbers for the binding (k_{on}) and release (k_{off}) rate constants measured with BLI need to be taken with caution, since immobilization can artificially slow binding reactions (Schreiber et al., 2009). However, these experiments do reveal that the Spy:Im7 complex is dynamic in nature since Im7 binding and release occurs relatively rapidly, as measured with this approach (Quan et al., 2014). Similarly loose/dynamic complexes were described for several other chaperones including eukaryotic Hsp90 in absence of co-factors as well as the ATP-

independent chaperone Trigger Factor of *E. coli* (Hoffmann et al., 2010; Karagöz and Rüdiger, 2015; Saio et al., 2014).

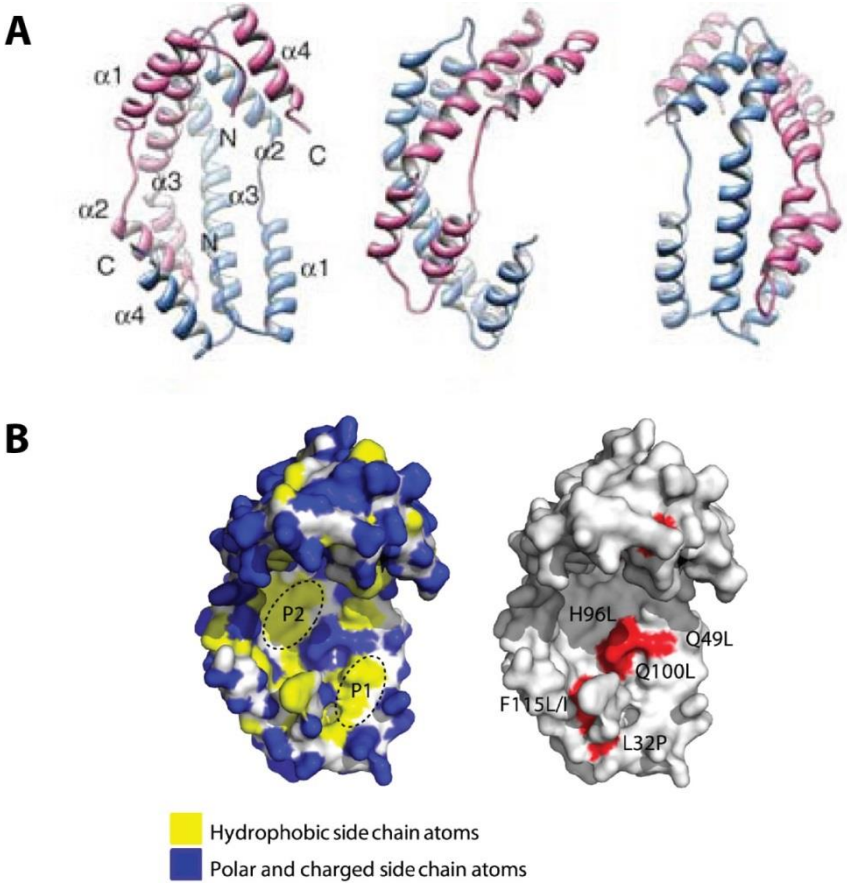

Figure 10: The Spy dimer forms a molecular cradle that provides shelter for folding proteins. (A) Crystal structure of the Spy homodimer (PDB ID: 3O39) shown in cartoon representation (monomers are colored in blue and magenta respectively). Each monomer consists of four helices (α1-α4). Helix 3 forms most of the dimer interface via a long coiled-coil, accompanied by contacts between helix 4 and helix 2. Helix 1 and 2 are separated by a rather flexible linker. Note that the 28 N-terminal and the 14 C-terminal residues are disordered and therefore invisible in the crystal structure (Image was adapted from (Quan et al., 2011)). (B) Surface representation of the Spy dimer (PDB ID: 3O39) to visualize the chemical properties. Left: The concave site of the Spy dimer is of amphiphilic nature, as two hydrophobic patches (P1 and P2, yellow) are surrounded by polar and charged side chains (blue). Right: Mutations which increase chaperone activity (red) either expand the existing two hydrophobic patches (Q49L, H96L, Q100L) or increase the flexibility of the N-terminus (L32P and F115L/I). Most mutations have been shown to increase the overall flexibility of the Spy dimer (Image was adapted from (Quan et al., 2014)).

A recent genetic approach to enhance the chaperone activity of Spy has shed light on the importance of the exposed hydrophobic surface and flexibility for the chaperone function of Spy (Quan et al., 2014). Following a directed evolution approach, in which Spy was randomly mutagenized and then genetically selected for improved chaperone activity in the *E. coli* periplasm, Spy mutants were isolated that not only yielded in improved ability to stabilize Im7 L53A I54A *in vivo*, but remarkably

also revealed generally improved chaperone function. Compared to the wild type of Spy, these mutants more efficiently protected chemically denatured alpha-lactalbumin and aldolase against aggregation and improved aldolase refolding *in vitro*. This result was surprising as previous attempts to improve chaperone function of other chaperones such as GroEL or Hsp70 using a single bait protein resulted mostly in variants that worked better on just the selected protein but were generally worse on other test proteins (Aponte et al., 2010; Schweizer et al., 2011; Wang et al., 2002). This work also provided valuable insights into client binding and folding by Spy (Quan et al., 2014). Strikingly, most of the gain of function mutations increased the hydrophobic surfaces on the concave site of Spy (Figure 10 B). Leucine-substitutions of Q100, H96, and Q49, all of which are in direct proximity to and therefore extended either of the two hydrophobic patches, increased the affinity of Spy variants for Im7-L53AI54A and yielded in improved aggregation suppression and protein refolding. In addition, three further mutations were selected that improved the overall chaperone function, namely F115L/I and L32P. While these mutations did not increase the size of the hydrophobic surface, they increased the flexibility at the tip-region of Spy and also thermodynamically destabilized the dimer, as revealed by hydrogen-deuterium exchange experiments. Similarly, the Q100L mutation increased the structural flexibility as well (Quan et al., 2014). Interestingly, one mutation, Q25R, increased the positive charge of the unstructured N-terminus, implying that positive charge may be important for chaperone function as well. These mutations demonstrate important aspects of how Spy assist in client folding, namely through extensive hydrophobic interactions and adaptive binding mediated through increased flexibility, characteristics that many chaperones share (Bardwell and Jakob, 2012). However, high-resolution details on the interaction between Spy and Im7 and how these affect the folding landscape of Im7 (Figure 3) were not determined.

2 Objective

Despite its small size and apparent independence from ATP and co-chaperones, Spy assists in folding of numerous proteins of different size and conformation. This makes Spy a very attractive model to study the effect of a chaperone on folding client proteins. Spy is amenable to a number of biophysical approaches, including X-ray crystallography, that are difficult to apply on large energy-dependent chaperone complexes such as GroEL or Hsp90. Hence, the aim of this PhD thesis was to dissect the mechanism of chaperone action of Spy kinetically, thermodynamically, as well as structurally, using the client proteins Im7 and α-casein with the goal of obtaining detailed mechanistic insights on how client proteins bind, fold, and are released from the chaperone *in vitro*. Im7 is ideally suited as a model client protein for this study, as it is a) an *in vivo* client protein of Spy (Quan et al., 2011), and b) its *in vitro* folding pathway is well established (Bartlett and Radford, 2010; Capaldi et al., 2002; Friel et al., 2009; Gsponer et al., 2006). Mutants of Im7 are available which are trapped in either the unfolded or intermediate state (Figure 3). The mutations L18A L19A L37A prevent the formation of native-like hydrophobic contacts early on in folding and hence trap Im7 in the unfolded state (Pashley et al., 2012). In addition, the mutations L53A I54A prevent the formation and docking of the third helix to the hydrophobic core and thus destabilize the native state of Im7 (Capaldi et al., 2002). This mutant populates conformations that closely resemble intermediate folding state. Both mutants are soluble at pH 7, which allows a detailed investigation of the interactions between Spy and the unfolded and intermediate folding state of Im7 at equilibrium in the absence of protein aggregation. Investigating the folding kinetics of Im7 in presence of Spy in combination with equilibrium binding experiments with the individual folding states should reveal how Spy influences the energy folding landscape of its client proteins. As outlined above, many known folding chaperones either utilize the cell's primary energy source ATP and other co-factors to drive the chaperone cycle or are conditionally activated (Bardwell and Jakob, 2012; Winter and Jakob, 2004b). However, Spy does assist in the refolding of chemically or thermally unfolded proteins, without a known external factor that regulates its activity. Examining the kinetic and thermodynamic factors that control the binding and release of Im7 may enable the determination of how Spy assists client folding in an energy-independent manner, and if multiple, consecutive binding and release cycles are necessary for Spy to successfully (re)fold proteins.

In addition, Im7 folds through an obligate on-pathway intermediate, and therefore represents a typical chaperone client as the majority of proteins populate folding intermediates. It has been proposed that chaperones, such as GroEL, may smooth the energy folding landscape either by unfolding kinetically trapped misfolded states or by destabilizing folding intermediates, either of which is expected to enhance folding rates (Balchin et al., 2016; Mattoo and Goloubinoff, 2014). Understanding how Spy affects Im7 folding might allow one to determine if Spy catalyzes folding of Im7 by smoothing the Im7's folding energy landscape, or if it changes the folding pathway of Im7.

High-resolution structural information about how chaperones affect their client proteins is still scarce. The fact that Spy prevents the aggregation and promotes the folding of structurally different proteins suggests that, as characteristic for chaperones, Spy has a low client specificity (Quan et al., 2011). High flexibility in conjunction with a broad amphiphilic binding surface may provide Spy with a broad client specificity and suggests that Spy may bind its client proteins rather loosely as conformational ensembles (Figure 5) (Burmann and Hiller, 2015; Quan et al., 2014). To test this hypothesis, a characterization of the structural and conformational properties of Spy in complex with a client protein was conducted by X-ray crystallography. Apo-Spy (without client bound) had been crystallized in the past (Kwon et al., 2010; Quan et al., 2011). We reasoned that this might offer a good starting point for further structural investigations of Spy in complex with various client proteins including the Im7 variants mentioned above and α-casein. Most notably, a truncated but soluble version of Im7 (amino acid 6 to 45) is available that is known to bind to Spy with similar affinities as the full-length protein (Quan et al., 2014). This peptide is unstructured, but exhibits a propensity to transiently form α-helices present in the native state (Quan et al., 2014). As it is shorter than intact Im7 it may impact crystal-packing to a lesser degree and thus might represent an ideal model client to start to investigate the structural features of Spy in complex with a partially folded client protein via X-ray crystallography. This investigation might also allow one to assess if Spy influences the peptides conformational equilibrium, potentially shifting it towards a more folded state. Bovine α-casein is a disordered protein classically used to study chaperone-client interaction and it has been shown to form stable complexes with Spy (Lin et al., 1995; Nam and Walsh, 2003; Quan et al., 2011). α-casein contains residual secondary and tertiary structure and constitutively exposes hydrophobic surfaces. Hence α-casein exhibits similarities to an intermediate folding state of many protein, but with the convenience of it being completely soluble which allows the conduction of equilibrium binding and X-ray crystallography experiments with Spy (Creamer et al., 1981; Hoagland et al., 2001). X-ray crystallography conducted in this thesis allowed to assess the side chains and distinctive portions of Spy involved in client binding, as well as conformational rearrangements in Spy necessary for client binding and folding. In the same time, influences of Spy on the conformational equilibrium of the bound client protein were determined.

In the next section, I give a summary and discussion of my 3 first-author/co-first author publications. These provide a very detailed understanding of the molecular basis of Spy's chaperone mechanism (Horowitz et al., 2016a; Koldewey et al., 2016; Stull et al., 2016). Additional six publications were published during my PhD thesis, in which I am listed as an author (but not co-first or first author). Four of these include *in vitro* studies on the molecular chaperones HdeA from *E. coli* (Dahl et al., 2015, 2016), YPL067C from *Saccharomyces cerevisiae* (Horowitz et al., 2016b), and mitochondrial peroxiredoxin from *Leishmania infantum* (Teixeira et al., 2015). In another publication, I determined the oligomeric states of different *de novo* engineered protein cages *in vitro* (Sciore et al., 2016). Finally, together with Scott Horowitz, I designed and supervised an X-ray crystallography lab class, in

which undergraduate students had to solve a protein structure using a published electron density map. A subsequent statistical analysis revealed, that surprisingly 14 % of the protein structures built by the students exceeded the overall quality of the structure originally deposited on the Protein Data Bank (Horowitz et al., 2014). As all of these six publications are only tangentially related to my main project on the molecular chaperone Spy, they are not further discussed in this thesis. A complete list of publications can be found in section 8.

I will start off with a summary of the results I obtained, followed by a very brief overview of the methods I used. I will end with a discussion of my results.

3 Summary of results obtained in each first author publication

The complete individual publications can be found online.

3.1 Substrate protein folds while it is bound to the ATP-independent chaperone Spy.

Frederick Stull*, **Philipp Koldewey***, Julia R. Humes, Sheena E. Radford, James C. A. Bardwell. **Nat Struct Mol Biol.** 2016 Jan;23(1):53-8. doi: 10.1038/nsmb.3133.

* These authors contributed equally to this work

First, we set out to determine how Spy facilitates protein folding without the means of a regulated client release, such as ATP-driven conformational changes. For that, we investigated the effects of Spy on the folding pathway of Im7. We determined if Im7 release is necessary for complete folding and if Spy-binding smooths the folding energy landscape, or in other words accelerates client folding, as postulated for some chaperones (Hayer-Hartl et al., 2016). Im7 folds through an obligatory intermediate state (Capaldi et al., 2002). To obtain an idea of how Spy binding affects the folding pathway of Im7, we first conducted binding experiments to determine the binding affinity as well as kinetic rate constants of Spy binding to the individual folding states of Im7, namely unfolded, intermediate, and native. For that, we utilized mutants of Im7 that trap the protein either in unfolded state (Im7 L18A L19A L37A) or intermediate state (Im7 I53A V54A) but are soluble, hence avoiding protein aggregation that would interfere with our measurements. We started out by determining the affinity of Spy to the three folding states of Im7 by isothermal titration calorimetry (ITC), and found that Spy binds to all three folding states of Im7, in a 1:1 stoichiometric complex with micromolar affinities, although the native state was bound with the weakest affinity. This was surprising considering the fact that chaperones are generally thought to exhibit low affinity to natively folded proteins due to the lack of exposed hydrophobic surface (Kim et al., 2013). Next, we investigated the kinetics of complex formation between Spy and the three folding states using a stopped-flow fluorimeter by monitoring the tryptophan fluorescence of Im7. A fluorescence change could be observed for all 3 variants of Im7 when mixed with Spy, confirming our equilibrium ITC titrations demonstrating that Spy binds to all three states. However, while this fluorescence change directly allowed us to calculate binding and release rate constants for Spy binding to the unfolded state, it only reported on a conformational change induced by Spy-binding for Im7 I53A V54A and wild type Im7. We found that complex formation between Spy and Im7 L18A L19A L37A is a diffusion controlled process with a binding rate constant of $1.3*10^7$ M^{-1} s^{-1}. Such rapid association rates are likely fast enough to effectively prevent the aggregation of unfolded Im7 *in vivo*. While the binding and release rate constants could not be directly determined for Im7 I53A V54A and wild type Im7, the conformational changes recorded indicated that Spy binding resulted in partially unfolding of wild type Im7 confirming the higher affinity of Spy for the intermediate state, as observed via ITC. The

data further indicated that Im7 unfolding and refolding occurred while bound to Spy. This and the fact that Spy binds all three folding states of Im7 suggested that Im7 folding may occur while continuously bound to Spy.

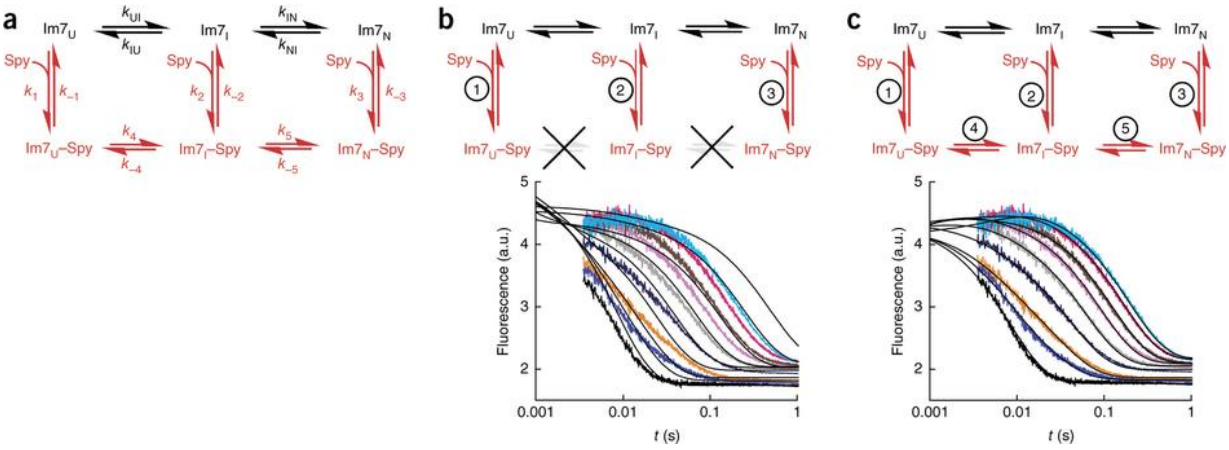

Figure 11: Global fitting of Im7–Spy kinetic data. (a) Global fitting built upon the well-characterized mechanism (Capaldi et al., 2001, 2002; Ferguson et al., 1999; Friel et al., 2003, 2009; Knowling et al., 2009) for Im7 folding in the absence of Spy (black path). The experimental fluorescence traces for Spy binding to Im7 WT and folding of Im7 WT in the presence and absence of Spy were globally fitted to different mechanisms containing various combinations of the steps in red. (b, c) Top: Attempted global fitting to the mechanism that omits (b) or allows (c) the folding of Im7 while it is bound to Spy. For clarity, only the traces for Im7 folding in the presence of Spy are shown (bottom). The black lines in the plots (bottom) are the best fit to the data. The mechanism that completely omits folding of Im7 while it is bound to Spy (b) fails to fit the data, whereas the mechanism that allows folding of Im7 while it is bound (c) can successfully fit the data. Global fitting to additional mechanisms and the best fit for the Spy–Im7 WT binding data are shown in the Supplementary Information of (Stull et al., 2016) (Figure and legend were adapted from (Stull et al., 2016)).

To test this hypothesis, we measured refolding rate constants of urea denatured wild type Im7 in the absence and presence of Spy. While Im7 folding slowed down when mixed with increasing concentrations of Spy, folding was not completely inhibited. A global fit analysis to determine the simplest kinetic model that could satisfyingly fit the data revealed that indeed Spy allows Im7 to fully fold into its native state while it remains bound to the surface of the chaperone (Figure 11). Hence, client release from the binding surface of Spy, as postulated for other chaperones, is not necessary in order for Im7 folding to occur. The kinetic model also suggested that Spy binding did not alter the folding pathway of Im7. As in solution, Im7 folded through the obligatory intermediate state while bound to Spy. In short, this data suggests that Spy provides a binding surface that can accommodate different conformational states of the client protein Im7 without substantially biasing their relative thermodynamic stabilities and, more importantly, allows free interconversion between these states. Folding in the protective environment of the chaperone, however occurs with a kinetic penalty, as folding is slowed down 30 fold, potentially by the contacts with Spy that stabilize both unfolded and intermediate state relative to the native state. The kinetic analysis presented in this publication

explains how Spy assists in the folding of client proteins without the need of an external factor that regulates client release. Under stress conditions that cause protein unfolding in the periplasm, Spy rapidly associates with denatured proteins, thus preventing their aggregation. The denatured proteins stay bound until stress is relieved, which thermodynamically favors protein folding. Client proteins can then fold while bound to Spy, eliminating the need for a conformational change that ejects the client protein into bulk solution in order to fold. A lower affinity to the native state may result in a release of the client protein once folded.

Contributions: For this publication I conducted ITC as well as stopped-flow kinetic experiments of Im7 L18A L19A L37A, Im7 L53A I54A, and Im7 wild type with Spy (as reported in Figure 2, 3, 4 a, Supplementary Figure 2) and cooperated with Frederic Stull and Julia Humes in the recording and global fitting of the Im7 folding kinetics in presence of Spy as reported in Figure 5, Table 1, and Supplementary Figure 3 and 6 (see the Methods Section for more details).

3.2 Forces driving chaperone action

Philipp Koldewey, Frederick Stull, Scott Horowitz, Raoul Martin, James C. A. Bardwell. **Cell**. 2016 Jul 14;166(2):369-79. doi: 10.1016/j.cell.2016.05.054.

In the previous paper we showed that Im7 binding occurs very rapidly but with a moderate micromolar affinity. Folding of Im7 occurs while bound to the chaperone without substantial changes in the folding pathway. Contrary to many chaperones, client release is not allosterically regulated but rather may occur upon completion of folding, as the native state of Im7 binds to Spy with the least affinity (Stull et al., 2016). This suggests that the binding surface of Spy offers enough contacts to keep different conformational states Im7 flexibly bound during folding, but doesn't bind any of the folding states too tightly to inhibit folding. In this paper, the nature of the molecular forces was investigated, that drive Im7 into complex with Spy, help it fold, and determine its release from the chaperone. To achieve this goal, we utilized mainly the same techniques used in the first paper described above, namely isothermal titration calorimetry and fluorescence stopped-flow, and investigated the chemistry of molecular contacts that govern the interaction between the chaperone Spy and the unfolded state compared to the folded state of Im7. To conduct these experiments, we used a modified version of the unfolded mutant Im7, that yielded an increased signal sensitivity upon Spy binding (L18A L19A L37A H40W), as well as wild type Im7. Contrary to the common notion that chaperones recognize their client proteins by exposed hydrophobic contacts (Balchin et al., 2016), we found that initial contacts between Spy and the unfolded state of Im7 are facilitated through long-range electrostatic interactions. These are responsible for the high binding rate constants determined by our previous kinetic analysis (Stull et al., 2016). Under stress conditions at which the periplasmic concentration of Spy can be 2 mM (Quan et al., 2011; Rutherford et al., 2010; Srivastava et al., 2014), these electrostatic interactions drive Im7 in complex with Spy in less than 8 µsec half-time. This is likely fast enough to prevent protein aggregation. The fact that the Im7 mutant we used is constitutively

unfolded, allowed us to also determine the chemistry of contacts that need to be broken in order for the unfolded state of Im7 to be released from Spy. We found that the unfolded state is bound to Spy via a mixture electrostatic and hydrophobic interactions, consistent with the amphiphilic nature of the Spy client binding site, which exposes hydrophobic patches surrounded by an excess of positive charge (Kwon et al., 2010; Quan et al., 2011, 2014). Hence, after the electrostatically mediated encounter between Spy and the unfolded state of Im7, hydrophobic contacts form between the two proteins that stabilize the complex. As determined by ITC, the binding enthalpy ΔH however is positive, implying that complex formation is entropically driven and the formed contacts are weak and non-specific.

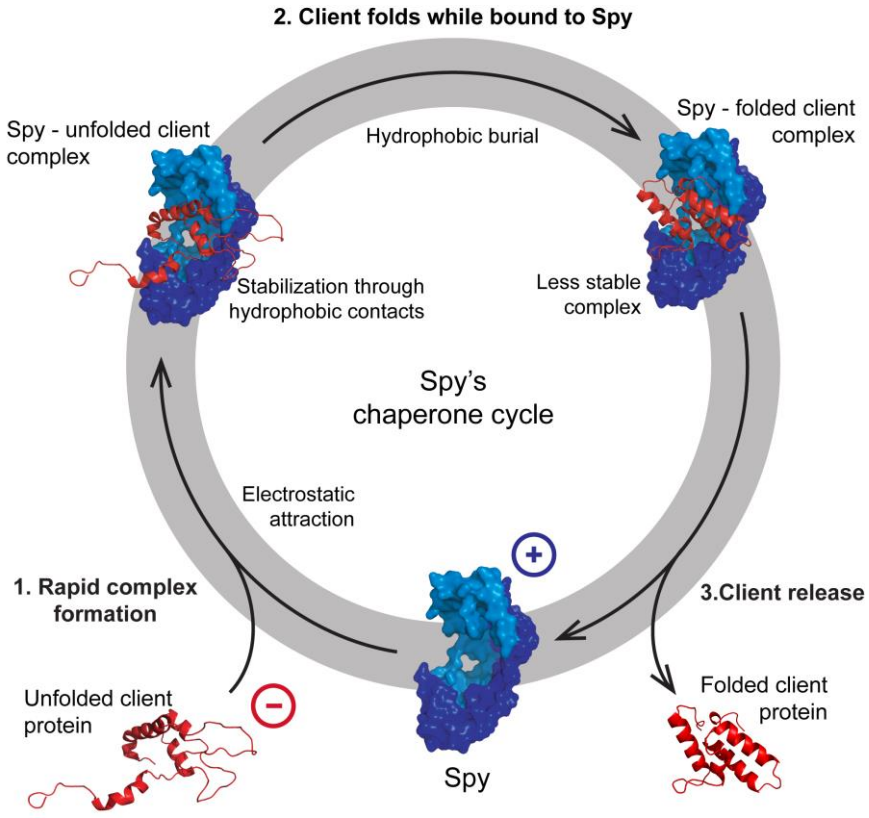

Figure 12: Mechanistic scheme of Spy-client interaction. (1) Client binding rates are maximized through long-range electrostatic attraction, which allows Spy (blue) to effectively compete with aggregation of the unfolded client protein (red). Client release, on the other hand, is energetically disfavored mainly by the solvation of hydrophobic surface area on the client and Spy, which are buried in the complex. (2) Folding of the client results in the burial of hydrophobic residues in the client's core, which decreases its affinity for Spy and therefore (3) favors release of the client protein. The electrostatic interactions, however, allow the client to stay bound to Spy while it folds (Figure and legend were adapted from (Koldewey et al., 2016)).

What causes the reduced affinity of the native state of Im7 compared to the unfolded state, which may be responsible for the release of Im7 once folded? Contrary to the unfolded state of Im7, the native state is mostly tethered to Spy via hydrophilic interactions, suggesting that the folding induced burial of hydrophobic residues in the core of Im7 while bound to Spy decreases the number of hydrophobic contacts between native Im7 and Spy. Consistent with the loss of stabilizing hydrophobic contacts between the chaperone and client, and in line with our previous kinetic analysis (Stull et al., 2016), the

release rate constant of natively folded Im7 is 13 fold higher than the unfolded state. In conclusion, the examination of the chemistry of molecular forces between Spy and unfolded and native states of Im7 allowed the construction of a more detailed picture of the mechanism of chaperone function (Figure 12). The aggregation prone unfolded state of Im7 rapidly binds to Spy via long-range guiding electrostatic interactions, kinetically competing aggregation. These are complemented by hydrophobic contacts in the complex. The interplay of electrostatic and hydrophobic contacts presumably allows Im7 to bind Spy loosely enough to undergo conformational transitions to the native state without the need for a pre-native release. Release of the native state from Spy is then mediated due to the loss of hydrophobic contacts.

Contributions: All experiments presented in this paper were conducted by me, except the following: Frederick Stull measured the salt dependent binding kinetics of casein$_{133\text{-}193}$ peptide to Spy as reported in Figure S4 and Raoul Martin determined the dissociation constants and the stoichiometry of Im7 L18A L19A L37A H40W with wild type Spy via equilibrium fluorescence titration experiments, as reported in Figure S1 A-D (see the Methods Section for more details).

3.3 Visualizing chaperone-assisted protein folding

Scott Horowitz*, Loïc Salmon*, **Philipp Koldewey***, Logan S. Ahlstrom, Raoul Martin, Shu Quan, Pavel V. Afonine, Henry van den Bedem, Lili Wang, Qingping Xu, Raymond C. Trievel, Charles L. Brooks 3rd, James C. A. Bardwell. **Nat Struct Mol Biol.** 2016 Jul;23(7):691-7. doi: 10.1038/nsmb.3237.

* These authors contributed equally to this work

The previous two publications drew a detailed picture of the mechanism of chaperone function of Spy, that allows Im7 to fold while bound to Spy mediated through amphiphilic contacts formed in the complex (Koldewey et al., 2016; Stull et al., 2016). Since both Spy and Im7 are amenable to X-ray crystallography, we sought to solve the structure of Spy-client complex. The goal was to gain high-resolution structural insights into interactions of a client bound to the chaperone Spy at different stages of folding. To improve crystal packing, we used a truncated version of Spy, missing the flexible N- and C-terminus. Further, to increase our chances to obtain a crystal of Spy in complex with a client protein, we attempted to co-crystalize several variants of Im7 and α-casein with Spy, namely wild type Im7, the unfolded Im7 mutant L18A L19A L37A, the Im7$_{6\text{-}45}$ peptide, and a soluble 30 amino acid peptide derived from bovine α-S1-casein that contains a Spy binding site. Binding affinities and binding sites to the truncated version of Spy were determined for all client proteins via isothermal titration calorimetry and fluorescence spectroscopy and were found to be in line with those observed for the wild type of Spy, implying that the terminal truncations do not interfere with the client binding capacity of Spy. We obtained well diffracting crystals for Spy in complex with all client proteins used, which allowed us to solve the structure of Spy to a resolution of 1.8 Å. However, only rather scattered residual electron density was visible for the client proteins. Our previous investigations suggested that

Im7 can flexibly interconvert between conformations while bound to Spy (Koldewey et al., 2016; Stull et al., 2016). Hence, we postulated that the lack of continuous electron density was due to structural heterogeneity of the client protein bound to Spy in the crystal. Solving the structure of heterogeneous and dynamic protein complexes has proven difficult via X-ray crystallography.

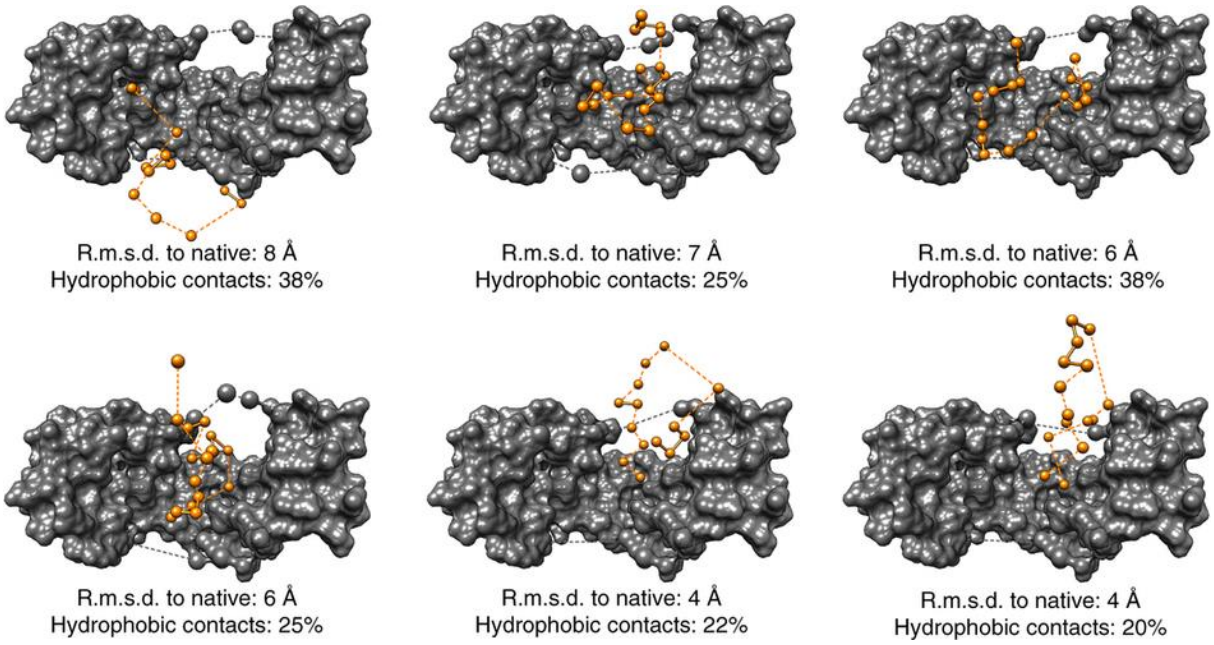

Figure 13: Spy–Im7$_{6–45}$ ensemble, arranged by r.m.s. deviation (r.m.s.d.) from the native state of Im7$_{6–45}$. Although the six-membered ensemble from the READ selection should be considered only as an ensemble, for clarity, the individual conformers are shown separately. Spy is depicted as a gray surface, and the Im7$_{6–45}$ conformer is shown as orange balls. Atoms that were not directly selected in the READ procedure or whose position could not be justified on the basis of agreement with the residual electron density were removed, thus leading to noncontiguous sections. Dashed lines connect noncontiguous segments of the Im7$_{6–45}$ client. Residues of the Spy flexible linker region that fit the residual electron density are shown as large gray spheres. Shown below each ensemble member is the r.m.s. deviation of each conformer from the native state of Im7$_{6–45}$, as well as the percentage of contacts between Im7$_{6–45}$ and Spy that are hydrophobic (Figure and legend were adapted from (Horowitz et al., 2016a)).

Therefore, to test our hypothesis, we developed a new method, called READ (Residual Electron and Anomalous Density), which allowed us to map the position of specific residues of the client protein labeled with the anomalous scatterer iodine within the residual electron density. We labeled eight residues of Im7$_{6to45}$ with iodine, one at a time, and identified multiple positions for seven out of the eight substituted residue, implying that Im7$_{6to45}$ was bound to Spy in multiple conformations. We then applied a sample-and-select procedure commonly used in NMR spectroscopy (Salmon and Blackledge, 2015) that allowed us to fit conformers of Im7$_{6to45}$ to the iodine positions and residual electron density. These conformers were selected from a pool of thermodynamically reasonable conformations of Im7$_{6to45}$ bound to Spy, generated via molecular dynamics simulations (Karanicolas and Brooks III, 2002). The best fit was achieved using a set of 6 conformers. This ensemble of Im7$_{6to45}$

bound to Spy yielded the first near-residue-level resolution information of a client protein bound to a chaperone. It provided six snapshots of $Im7_{6to45}$ exploring its folding landscape while bound to the concave side of Spy, sampling different folding states from unfolded to partially folded to near-native (Figure 13). Contrarily, free in solution, this Im7 peptide is mostly unfolded (Quan et al., 2014). While all conformers made hydrophobic as well as hydrophilic and charged contacts with Spy, strikingly, the content of hydrophobic contacts decreased from unfolded to near native states, confirming our previous kinetic and thermodynamic analysis demonstrating that folding reduces intermolecular hydrophobic contacts between chaperone and client (Koldewey et al., 2016). Residues on the surface of Spy interacting with Im7 involved large portions of the concave side, including the hydrophobic patches with surrounding positively charged residues and the more flexible rim. In addition, the N-terminal helix as well as the adjacent linker region became more flexible in the client-bound state compared to the apo-structure (Figure 10 and Figure 13), implying that flexibility is important to accommodate different conformations of Im7. Summarized, the structural ensemble of Im7 solved via X-ray crystallography confirms our previous kinetic and thermodynamic analyses and suggests a mechanism by which Spy provides a shelter for folding proteins such as Im7 by offering an amphiphilic, flexible binding site that allows the client protein to explore conformational space along its folding energy landscape without constrains imposed by the chaperone.

Contributions: In this paper, I collaborated with Scott Horowitz and Raoul Martin in screening of crystallization conditions and optimization, crystal harvest, data collection from over a thousand Spy-Im7 or -casein co-crystals at the synchrotron (Argonne National Laboratories), model building and mapping of iodine positions (as reported in Figure 1 a and b; Table 1; Figure 5 a; Supplementary Table 1 and 2) as well as qualitative assessment of chaperone bound client protein in the crystals (as reported in Supplementary Figure 2). In addition I conducted all ITC, analytical ultracentrifugation, as well as equilibrium fluorescence titration experiments as reported in Figure 5 b and Supplementary Figure 8 (see the Methods Section for more details).

3.4 Methods summary

This chapter provides a brief overview of the methods I used to accomplish the goals of my main project outlined above. Methodologies contributed solely by collaborators are not described here. For detailed procedures, see the materials and methods sections of each publication that can be found online (Horowitz et al., 2016a; Koldewey et al., 2016; Stull et al., 2016).

3.4.1 Proteins and peptides

Protein expression and purification. All variants of Spy, Im7, and bovine α-casein (except Im7 and casein peptides used for X-ray crystallography, which were purchased from New England Peptide) were expressed in the *E. coli* strain BL21 (DE3), using a modified pET28b expression vector containing a N-terminal His_6-SUMO-tag (Horowitz et al., 2016a; Koldewey et al., 2016; Quan et al., 2011, 2014; Stull et al., 2016). Different variants of Spy and Im7 were generated using site directed mutagenesis following the QuikChange protocol (Agilent Technologies). α-lactalbumin was purchased from Sigma and carboxymethylated as previously described to obtain a fully unfolded and soluble version of α-lactalbumin (Schechter et al., 1973). A complete list of proteins used in this work can be found in Table 2.

Protein purification. In brief, after cell lysis and centrifugation, proteins of interest, contained in the soluble fraction, were separated from the majority of cell proteins by nickel-affinity chromatography, followed by tag-cleavage using the SUMO-specific protease ULP1 (Sun et al., 2011). High protein purity was achieved by adding additional steps of size-exclusion and/or ion-exchange chromatography, depending on the chemical properties of the individual protein. Detailed purification protocols can be found in the respective publications (Horowitz et al., 2016a; Koldewey et al., 2016; Quan et al., 2011, 2014; Stull et al., 2016).

Table 2: Summary of proteins and peptides used in this work.

Protein	Description
Spy wild type[1,2,3]	The chaperone studied in this thesis
Spy Q100L[2]; Spy H96L[2]	Variants of the chaperone Spy with increased chaperone activity (see section 1.4.2 in the introduction)
$Spy_{29\text{-}124}$[3]	Variant of Spy with truncated termini used for X-ray crystallography
$Spy_{29\text{-}124}$ SeMet[3]	Variant of Spy with truncated termini used for X-ray crystallography, labeled with selenomethionine
$Spy_{29\text{-}124}$ H96L[3]	Variant of Spy with truncated termini used for X-ray
Im7 wild type[1,2,3]	Colicin E7 immunity protein and *in vivo* client protein of Spy
Im7 L18A L19A L37A[1,2,3]	Variant of Im7 that is constitutively unfolded

Im7 L18A L19A L37A H40W[2,3]	Variant of Im7 that is constitutively unfolded, H40W substitution increases signal-sensitivity for Spy-binding
Im7 L18A L19A L37A W75F[2]	Variant of Im7 that is constitutively unfolded and lacks the endogenous tryptophan; used as a competitor for binding competition experiments
Im7 L53A I54A[1]	Variant of Im7 that populates conformations closely resembling the intermediate folding state of Im7
$Im7_{6-45}$[3]	Minimal Spy-binding segment that encompasses and partially populates the first two helices of Im7
$Im7_{6-45}$ pI-Phe[3]	$Im7_{6-45}$ peptide containing one residue substituted by 4-iodo-phenylalanine (8 peptides, each of which containing a unique substitution).
α-S1-$casein_{133-193}$[2]	Fragment of bovine α-S1-casein, which is a naturally disordered protein with little secondary structure (Koczan et al., 1991).
α-S1-$casein_{148-177}$[3]	Fragment of bovine α-S1-casein, which is a naturally disordered protein with little secondary structure (Koczan et al., 1991).
Carboxymethylated α-lactalbumin[2]	fully unfolded and soluble version of bovine α-lactalbumin

[1]Protein used in publication (Stull et al., 2016)

[2]Protein used in publication (Koldewey et al., 2016)

[3]Protein used in publication (Horowitz et al., 2016a)

3.4.2 Kinetic and thermodynamic aspects of Spy-client interaction and folding

Sedimentation velocity analytical ultracentrifugation. Sedimentation velocity (SV) experiments performed in an analytical ultracentrifuge are used to determine the molecular weight as well as hydrodynamic properties of solutes, such as proteins, by continuously monitoring their sedimentation in a gravitational field generated by centrifugation. In addition, this technique provides information about attractive and repulsive interaction of the sedimenting solutes. Hence, SV analytical ultracentrifugation is a versatile tool used to study the dynamics of protein complexes at equilibrium (Lebowitz et al., 2002). Here, SV experiments were performed to determine the stoichiometry of Spy-client complexes as well as the oligomeric state of each protein free in solution. Information about the complex-stoichiometry was vital to establish a correct model for the fitting of kinetic as well as structural data (described below). In brief, each protein individually or client protein in combination with Spy mixed in various ratios were sedimented using a Beckman Proteome Lab XL-I analytical ultracentrifuge (Beckman Coulter). All proteins were dialyzed prior to performing the experiments to ensure uniformity of the buffer. Samples were loaded into cells containing standard sector shaped two-channel Epon centerpieces with 1.2 cm path-length (Beckman Coulter) and temperature-equilibrated in the centrifuge for at least 1 hr prior to sedimentation. Samples were spun at 42,000 to 48,000 rpm,

and the sedimentation of the protein was monitored continuously either by absorbance at 280 nm or via interference. Data analysis was conducted with the program SEDFIT using the continuous c(s) distribution model (Lebowitz et al., 2002; Schuck, 2000). Buffer density, viscosity as well as partial specific volume of the proteins (based on the protein sequence) were calculated with SEDNTERP (http://sednterp.unh.edu). The sedimentation distribution plots, c(s) as a function of s, obtained by SEDFIT were integrated to determine the relative distribution of complexed and freely sedimenting client protein. The integrated area was then plotted as a function of total client concentration in order to obtain binding isotherms. The stoichiometry was calculated by fitting the binding isotherms with a quadratic equation (Horowitz et al., 2016a; Koldewey et al., 2016).

Isothermal titration calorimetry. ITC is a very sensitive method to determine the stoichiometry, affinity, as well as thermodynamic parameters of protein-protein interactions by measuring the heat release or absorption of the binding reaction (Jelesarov and Bosshard, 1999). Here, ITC was used to determine the affinities, stoichiometries as well as the binding enthalpy (ΔH) and entropy (ΔS) of Spy-client complexes. The thermodynamic parameters ΔH as well as ΔS provided detailed insights into the chemistry of Spy-client interaction (as detailed in the introduction). In brief, ITC experiments were performed using a MicroCal iTC200 (Malvern Instruments) with client protein in the cell and a 10 fold molar excess of Spy in the titration syringe. All samples were dialyzed against the respective buffer overnight prior to running the experiment to ensure uniformity of the buffer. Concentrations of Spy dimer and client protein were varied depending on the dissociation constant (K_d) of the binding reaction at the respective condition as well as the detection limit of the instrument. For any Spy-client complex investigated in this thesis (Table 2), a stoichiometry of 1:1 (Spy dimer:client) was determined by analytical ultracentrifugation as well as ITC. Hence, ITC thermograms were fit to a one-site model using the software Origin (OriginLab). Salt concentrations of the buffer were varied to investigate the involvement of electrostatic interactions in the binding reaction. The environmental temperature was varied in order to determine the chemistry of contacts between Spy and different folding states of Im7 by determining the heat capacity change (ΔC_p) of the binding reaction. The change in heat capacity (ΔC_p) upon Spy-client complex formation was derived from the slope of a linear fit of the enthalpy change (ΔH) as a function of temperature (Baldwin, 1986; Horowitz et al., 2016a; Koldewey et al., 2016; Stull et al., 2016).

Equilibrium fluorescence spectroscopy and fluorescence anisotropy. Fluorescence spectroscopy was used as an additional method to determine the affinity and stoichiometry of diverse client proteins to the chaperone Spy. While Spy is tryptophan-free, Im7, casein, and alpha-lactalbumin each contain endogenous tryptophan residues. The fluorescence emission of tryptophan is sensitive to changes in the direct protein and solvent environment of the tryptophan residue induced for example by protein-protein interaction or protein conformational changes (Lakowicz, 2006). Hence, tryptophan fluorescence was used as a readout for Spy-client complex formation. Alternatively, a change in client

tryptophan fluorescence anisotropy, that resulted upon binding of the client to Spy, provided a measure for complex formation. Fluorescence anisotropy reports on the tumbling time of the fluorophore, in this case the endogenous tryptophan of the client protein. Spy-binding increases the tumbling time of the client protein which results in an increase in observed fluorescence anisotropy (Lakowicz, 2006).

Fluorescence spectroscopy. In brief, client protein was titrated with increasing concentrations of Spy and binding was detected by monitoring the change in tryptophan fluorescence of the client protein in a fluorimeter. The fluorescence was then plotted as a function of Spy dimer concentration and the resulting binding isotherms were fitted either to a square hyperbola function or to a quadratic equation, in order to determine the binding affinity or stoichiometry, respectively (Horowitz et al., 2016a; Koldewey et al., 2016).

Tryptophan fluorescence based binding competition experiments were performed to determine the binding affinity of the tryptophan-free $Im7_{6–45}$ peptide for Spy, used for X-ray crystallography (Table 2). For that, a pre-mixed complex of Spy with Im7 L18A L19A L37A H40W (which contains two tryptophans), was titrated with tryptophan-free $Im7_{6–45}$ to compete with Im7 L18A L19A L37A H40W for Spy binding. The observed fluorescence change was plotted as a function of the logarithm of the $Im7_{6–45}$ concentration. Based on the stoichiometry obtained by ITC, and SV experiments, the competition data was then fit for a one-site-binding competition model to obtain an apparent dissociation constant (K_d) for $Im7_{6–45}$ on the basis of its ability to compete with Im7 L18A L19A L37A H40W. The actual K_d for $Im7_{6–45}$ binding to Spy was calculated with the Cheng–Prusoff equation (Horowitz et al., 2016a).

Fluorescence anisotropy. Tryptophan fluorescence anisotropy was used to determine the stoichiometry of Spy-client complexes. Different variants of Im7 were titrated with increasing concentrations of Spy and the fluorescence anisotropy was measured in a fluorimeter. Respective fluorescence anisotropy values were plotted as a function of Spy dimer concentration and fitted to a quadratic equation to obtain the complex stoichiometry (Koldewey et al., 2016).

Stopped-flow fluorescence and data fitting. Stopped-flow is a tool used to record fast reaction kinetics by monitoring a signal change that occurs upon rapidly mixing the reactants in a mixing chamber (Zheng et al., 2015). Here, stopped-flow was used to determine the binding (k_{on}) and release (k_{off}) rate constants of complex formation between Spy and the different variants of Im7, casein, or alpha-lactalbumin. It was also used to monitor the folding kinetics of Im7 in absence and presence of Spy. As for the equilibrium titrations described above, tryptophan fluorescence of the endogenous tryptophan(s) of the client proteins was used as a readout to monitor the kinetic of Spy-client complex formation. In addition, the single tryptophan of Im7 reports on the folding state and hence was used to

monitor Im7 folding in absence and presence of Spy (principles are detailed in the introduction section 1.2.1.3).

Determination of binding (k_{on}) and release (k_{off}) rate constants of Spy-client complexes: Briefly, the transient kinetics of Spy-client complex formation were recorded using a SF-2004 stopped-flow spectrofluorometer (KinTek) by monitoring the change in tryptophan fluorescence of the client protein upon rapid mixing with different concentrations of Spy. Recorded transients were fitted to a sum of exponential functions to derive observed rate constants (k_{obs}). The final concentration after mixing of client and Spy was chosen such that a pseudo-first-order approximation could be used for the data analysis and the observed rate constants did not exceed the limits of the instrument (the dead time of our instrument was determined to be 1.3 msec). To determine the binding rate constant (k_{on}), the observed rate constants (k_{obs}) derived from the exponential fit were plotted as a function of Spy concentration. Phases that showed an increasing linear dependence of k_{obs} on Spy concentration were fitted with a linear function to obtain k_{on} from the slope (Kozlov and Lohman, 2002). The dissociation constant (k_{off}) was calculated either from the y-intercept of linearly fitted observed rate constants (Kozlov and Lohman, 2002), and alternatively determined via binding competition with a tryptophan free variant of Im7 (Im7 L18A L19A L37A W75F, see Table 2). In the latter case, pre-formed complex of Spy and client were loaded into the stopped-flow instrument and rapidly mixed with tryptophan-free Im7 L18A L19A L37A W75F as a competitor for Spy's client binding site. The concentration of competitor was increased until the point where no additional change in the observed rate constant was detected. The recorded transient at the maximum concentration of competitor, at which no additional change was observed, was then fitted with an exponential function to derive the k_{off}. In case, the transients derived from the competition experiment revealed at least two exponentials, double-mixing binding experiments were performed in order to determine which of the exponentials corresponded to client release, and which to a conformational change. In this case, complex formation between Spy and client was allowed for 6 to 10 msec. The reaction was then chased with tryptophan free competitor Im7 L18A L19A L37A W75F. The recorded transients were then fit with a single exponential function to derive the k_{off} (Koldewey et al., 2016; Stull et al., 2016).

Im7 folding in absence and presence of Spy: The kinetic mechanism of Im7 folding free in solution has been established previously (Capaldi et al., 2001; Ferguson et al., 1999; Friel et al., 2009). Upon dilution, urea-denatured Im7 folds to its native state through an obligatory intermediate state, which is stabilized in presence of sodium sulfate (folding mechanism detailed in the introduction section 1.2.1.3). However, sodium sulfate could not be used in this work, as it also screens electrostatic interactions between Spy and Im7 and hence dramatically reduces their binding affinity. Thus, the microscopic rate constants of each step of the three-state folding pathway of Im7 (k_{ui} and k_{iu} for the transition between unfolded and intermediate state, as well as k_{in} an k_{ni} for the transition between intermediate and folded state) in absence of Spy had to be determined for the experimental conditions

used. This was done by measuring unfolding and refolding kinetics of Im7 at increasing urea concentrations as described previously in the absence of Spy (Capaldi et al., 2001; Ferguson et al., 1999; Friel et al., 2009; Stull et al., 2016). To measure folding of Im7 in the presence Spy, wild type Im7, denatured in 8 M urea, was rapidly diluted into buffer containing increasing concentrations of Spy and the tryptophan fluorescence of Im7 was recorded while folding (Stull et al., 2016).

Global fitting of kinetic data. In order to identify the correct kinetic mechanism that describes the folding of Im7 in presence of Spy, stopped-flow fluorescence traces for Spy binding to Im7 under non-denaturing conditions and the traces for urea-denatured Im7 refolding in the absence and the presence of different Spy concentrations were globally fitted to various kinetic mechanisms using the KinTek Explorer (KinTek Corporation) (Johnson et al., 2009). This program fits kinetic data by direct numerical integration of the differential rate equations that define the kinetic model, to directly obtain the microscopic rate constants and the fluorescence signals of the individual species of the kinetic mechanism. Potential kinetic mechanisms for Spy–assisted folding of Im7 used for the data fitting were built using the already known three-state mechanism for Im7 folding (Capaldi et al., 2001; Ferguson et al., 1999; Friel et al., 2009). In order to determine the simplest mechanism that can adequately describe the data , sequentially steps of Spy binding to the individual folding states of Im7 as well as folding transitions of Im7 while bound to Spy were added to the Im7 folding mechanism, and used to simultaneously fit the kinetic data sets (Stull et al., 2016).

Circular dichroism (CD) spectroscopy. CD spectroscopy can be used to estimate the content of secondary structure of proteins based on the property of secondary structural elements such as α-helices, β-sheets, as well as random coil to absorb circular polarized light differently (Micsonai et al., 2015; Whitmore and Wallace, 2008). Here, CD spectroscopy was used to determine content of secondary structure of different Im7 mutants used for the kinetic as well as thermodynamic analyzes of Spy-client interaction as described above (Koldewey et al., 2016).

3.4.3 Structural characterization of Spy-client complexes

Protein crystallization. In order to solve the structure of a Spy-client complex, a Spy mutant that lacks the unstructured termini (amino acid 29 to 124, termed $Spy_{29\text{-}124}$) was engineered. Since flexible regions in proteins are hard to crystallize, this truncation improved crystal packing. Sparse matrix crystallization screens (Wooh et al., 2003) were then set up with various concentrations of $Spy_{29\text{-}124}$, in absence or presence of different Im7 or casein variants (Table 2), in order to screen for conditions that yielded well diffracting crystals in presence of client protein, but not in the absence. After optimizing for suitable conditions, crystals were harvested and flash frozen in the presence of a cryoprotectant. To verify that client protein was indeed bound to Spy, the crystals were washed in crystallization solution, in order to remove all surface-bound and precipitated client protein, and then dissolved for visualization by SDS–PAGE (Horowitz et al., 2016a).

X-ray crystallography. Data collection took place at the Advanced Photon Source at the Argonne National Laboratory, Illinois, USA. Harvested crystals diffracted up to 1.8 Å resolution. Data integration and scaling were performed with iMosflm (Battye et al., 2011) and AIMLESS (Winn et al., 2011), respectively. Despite the high quality of the diffraction data, the structure could not be solved using the apo-structure of Spy as a molecular-replacement search model, indicating substantial differences between the apo-structure and the client –bound structure of Spy. Therefore, seleno-methionine (SeMet)-labeled Spy in complex with the $Im7_{6\text{-}45}$ peptide was crystallized and the structure was solved via Se-SAD phasing (Terwilliger et al., 2016), followed by density modification and initial model building using AutoSol, which is part of the software suite Phenix (Adams et al., 2010). While there was continuous electron density for Spy, only residual scattered density was visible for the Im7 peptide, which therefore could not be built in. The obtained model of Spy from this data set was then used as a molecular-replacement search model to solve the structures of the remaining Spy-client complexes via Phenix. Model building and refinements for all Spy-client complexes were conducted using the programs COOT (Emsley et al., 2010) and Phenix (Afonine et al., 2012).

Localizing client protein bound to Spy in the crystal. Similar to $Im7_{6to45}$, the electron density of any client protein, co-crystallized with Spy, was fragmented and hence neither of the client proteins could be built into the electron density map. The lack of continuous density for the client proteins presumably was due to structural heterogeneity and/or low occupancy of the client protein bound to Spy in the crystal. In order to map the positions of amino acids residues of the client within the residual electron density, $Im7_{6–45}$ was labeled with 4-Iodo-Phenylalanie (pI-Phe). Iodine is a strong anomalous scatterer, hence allowing its localization, even at low occupancy. We decided to label $Im7_{6–45}$ because of its unique structural properties (see section 2 for details) and because it yielded the most robust crystals in conjunction with Spy. In addition, pI-Phe modified versions were commercially available (New England Peptide). In total, 8 peptides were ordered, each with a unique pI-Phe substitution site. To identify the atomic positions of pI-Phe residues of the bound Im7 peptides, anomalous difference maps of the complexes were generated with phases from a molecular-replacement search using the non-labeled Spy–$Im7_{6–45}$ (with no $Im7_{6–45}$ built in) complex as the search model, as described above. Placed iodine atoms were then refined in Phenix (Afonine et al., 2012) with anomalous group refinement (Horowitz et al., 2016a).

4 Discussion

Understanding how chaperones work is an important question in basic biology. Given their function as guardians of protein folding and homeostasis, chaperones not only play roles in protein folding diseases such as Alzheimer's and cystic fibrosis but, due to their involvement in signal transduction pathways, are also important in the pathogenesis of cancer (Whitesell and Lindquist, 2005; Xu and Neckers, 2007). In addition, chaperones are thought to be major players in the process of aging, as chaperone levels drop dramatically in an aging organism, likely causing collapse of protein homeostasis (Hipp et al., 2014; Labbadia and Morimoto, 2015; Morawe et al., 2012). However, how protein folding is affected by chaperones is still a matter of debate. While much is known to date about structural, thermodynamic, as well as kinetic features of the sometimes complex conformational changes that drive chaperone cycles of chaperone foldases (Hayer-Hartl et al., 2016; Mayer and Breton, 2015; Zhuravleva and Gierasch, 2015), so far no equivalent detailed information is available on the effects of chaperones on the folding energy landscape of their client proteins (Balchin et al., 2016). This is in part due to the structural and mechanistic complexity of the major chaperone foldases and, in addition, the heterogeneity and flexibility of chaperone-client complexes (Burmann and Hiller, 2015). The most prominent example of chaperone-client interaction is the *E. coli* chaperone GroEL, but even in this case there is not yet good agreement concerning the fundamentals of this interaction and the effects on the client's folding landscape, in part due to the large and complex structure of this chaperone and its catalytic cycle (Apetri and Horwich, 2008; Brinker et al., 2001; Chakraborty et al., 2010; Georgescauld et al., 2014; Gupta et al., 2014; Horwich et al., 2009; Priya et al., 2013a; Tyagi et al., 2011; Weaver and Rye, 2014; Zhang and Kelly, 2014) .

I have closely examined the molecular chaperone Spy of *E. coli* as a much simpler model system with the hope of being able of identifying kinetic, thermodynamic, and structural aspects that allow molecular chaperones to promote client folding and how a chaperone affects the folding landscape of client proteins. The three publications summarized in this thesis allow us to draw a picture of the mechanism by which Spy facilitates folding of its client protein Im7 in unprecedented detail (Horowitz et al., 2016a; Koldewey et al., 2016; Stull et al., 2016). I found that the fundamental basis of Spy's chaperone activity is the "folding-friendly" amphiphilic and flexible nature of Spy's client binding site. The client binding site encompasses a large part of the concave surface of Spy's cradle shaped structure. It consists of 4 hydrophobic patches surrounded by hydrophilic residues, that are enriched in positive charges. The N-terminal helix and the linker between helix 1 and helix 2 are flexible. This combination of flexibility and amphiphilic binding surface allows Spy to flexibly bind the many conformational states that are populated along the folding trajectory of Im7 and hence mediate folding of Im7 while continuously but loosely bound to Spy (Horowitz et al., 2016a; Stull et al., 2016). Binding of the unfolded state of Im7 is mediated through a mixture of hydrophobic and hydrophilic interactions (Koldewey et al., 2016). This allows Im7 to explore its folding landscape while bound.

We confirmed this result by X-ray crystallography, showing that the $Im7_{6-45}$ peptide is bound to Spy as a conformationally heterogeneous ensemble, sampling conformations from unfolded, to intermediately and near-natively folded (Horowitz et al., 2016a). Though Im7 experiences a relatively high degree of conformational freedom while bound to Spy, conformational transitions are slowed down (Stull et al., 2016). This slowing may provide a mechanism whereby the pre-mature formation of tertiary contacts and hence misfolding are avoided. Although the kinetics of Im7 folding are slowed, our kinetic model predicts that the Im7's folding pathway remains relatively unaffected by the chaperone environment (Stull et al., 2016). The folding intermediate, found while folding of Im7 in the absence of the chaperone, is predicted to still populated in the chaperone bound state. Folding of Im7 results in the formation of a hydrophobic core within the interior of Im7 itself and hence reduces stabilizing hydrophobic contacts between Spy and Im7 (Horowitz et al., 2016a; Koldewey et al., 2016). This destabilizes the complex helping to trigger Im7 release (Koldewey et al., 2016). Such an amphiphilic binding and folding mechanism may explain how Spy can facilitate folding of the various unrelated proteins (Quan et al., 2011, 2014), as it avoids the need for client specific folding instructions. More importantly, rather than being dictated by the chaperone, client folding itself regulates client binding and release. Hence, we consider that Spy acts rather passively, by providing a sanctuary for folding proteins that prevents protein aggregation and misfolding, while the folding pathway remains dictated by the primary sequence of the client protein. The electrostatic interactions formed between Spy and Im7 are a central component of this mechanism. Not only do these enhance the client binding rate and therefore kinetically prevent protein aggregation, they also help keep Im7 bound while it folds and hence eliminate the need for pre-native client release, which has often been described as being essential for the successful folding of clients by chaperones (Kim et al., 2013; Mattoo and Goloubinoff, 2014).

Spy provides an amphiphilic and flexible client binding site to which different non-native conformational states of Im7 can loosely bind with about equal affinity (Horowitz et al., 2016a; Koldewey et al., 2016; Stull et al., 2016). This is mediated through many weak, non-specific contacts that allow Im7 to explore its conformational space without restrictions imposed by Spy. Amphiphilic client binding surfaces are a common feature of chaperones (Hoffmann et al., 2010; Karagöz and Rüdiger, 2015; Lopez et al., 2015). Thus it is possible that such a protective, but passive mode of chaperone action, that promotes self-guided client folding, might be the underlying mechanism for several folding chaperones. Indeed, loose client binding has been postulated as essential for chaperone-binding mediated client folding (Burmann and Hiller, 2015; Jewett and Shea, 2006). For example Hsp90 as well as the chaperonins GroEL, and CCT/TRiC provide broad and heterogeneous client binding sites (Joachimiak et al., 2014; Karagöz and Rüdiger, 2015; Perrett et al., 1997). The client binding site of these classic folding machines reveals similarities to Spy, in that they expose small hydrophobic patches that are surrounded by electrostatic residues. I note however, that the net charge of Spy and the cytosolic chaperones is inverted, probably due to the difference in proteome

charge propensities between the cytosol and periplasm (Joachimiak et al., 2014; Karagöz and Rüdiger, 2015; Kwon et al., 2010; Perrett et al., 1997; Quan et al., 2011). Hps90's large amphiphilic client binding surface allows this chaperone to potentially accommodate client proteins of different size and folding state, from completely unfolded to native, raising the possibility that Hsp90, like Spy may bind its clients as dynamic ensembles (Hagn et al., 2011; Karagöz and Rüdiger, 2015; Park et al., 2011a, 2011b; Verba et al., 2016). For example, some studies have shown that p53 forms a conformationally heterogeneous relatively unfolded ensemble that interacts with Hsp90 (Park et al., 2011a, 2011b) though other studies indicate that p53 may also bind Hsp90 in a native-like conformation (Hagn et al., 2011). While the less folded states interact with Hsp90 through multiple transient hydrophilic as well as hydrophobic contacts (Park et al., 2011a, 2011b), the native-like state in contrast preferentially binds through electrostatic interactions (Hagn et al., 2011). It is conceivable that like Spy, Hsp90 also allows its client proteins to explore their conformational space and fold while bound The finding that ATP binding or hydrolysis does not affect affinity of Hsp90 to different client proteins in the absence of co-chaperones strengthens this hypothesis, as client release may occur at any time of the chaperone cycle (Karagöz and Rüdiger, 2015; Park et al., 2011a). Further, Hsp90's secretion into the ATP-devoid extracellular space in mammals indicates that the ATPase-activity may not be absolutely required for the essential chaperone function (Wang et al., 2009). So why did evolution then implement an ATP-dependent conformational cycle for Hsp90, regulated in addition by a set of co-factors in the case of eukaryotic Hsp90? Perhaps, the addition of ATP- and co-chaperone-regulation may have been evolutionary refinements that allow the cell to fine-tune the client specificity and enhance the efficiency of their chaperones. The structures and folding pathways of eukaryotic proteomes are far more versatile and complex compared to the relatively simple *E. coli* periplasmic proteome (Balchin et al., 2016). Hence, in the case of eukaryotic Hsp90 there thus may be a need for customization and fine-tuning of client affinity as well as binding and release kinetics through co-factors (Röhl et al., 2013).

The interior wall of the hetero-octameric eukaryotic chaperonin CCT/TRiC, in the closed conformation, is chemically heterogeneous like that of the concave surface of Spy. For CTT/TRiC, hydrophobic patches, mostly located in the apical domains, are surrounded by an abundance of charge (Dekker et al., 2011; Joachimiak et al., 2014). One could speculate that this heterogeneity may provide CCT/TRiC with a similar "folding-friendly" surface as found for Spy. Bolstering this consideration is the observation that during folding in the enclosed interior of CCT/TRiC, the client protein α-actin binds in an extended, partially folded state to a broad surface of CCT/TRiC that encompasses several subunits (Dekker et al., 2011). The fact that CCT/TRiC contains eight different subunits enhances the heterogeneity of this binding surface, potentially allowing it to accommodate an even broader clientele that differs in structure and conformation (Dekker et al., 2011; Joachimiak et al., 2014; Roh et al., 2016b). Hence, if one was to apply similar kinetic and structural methods, which helped me uncover the chaperone mechanism for Spy, in conjunction with the utilization of client proteins with well-

defined folding pathways, one might gain further insights into the chaperone mechanism of Hsp90 or TRiC.

Formation of a hydrophobic core drives Im7 folding while bound to Spy. Desolvation of hydrophobic residues is generally the major entropic force that drives folding of globular proteins (Baldwin, 2007; Chandler, 2005; Dill, 1990, 1985; Kellis et al., 1988; Pace et al., 2011; Southall et al., 2002). This applies to Im7 as well, where the formation of the intermediate state of Im7 is driven by the burial of most of the hydrophobic residues (Bartlett and Radford, 2010). What then drives folding of Im7 in complex with Spy, if desolvation of hydrophobic contacts has already occurred through binding to the chaperone (Koldewey et al., 2016)? It is generally thought that chaperones have to release the bound hydrophobic surfaces of the client protein so that the client can complete folding (Kim et al., 2013; Mattoo and Goloubinoff, 2014). However, Spy allows Im7 to fold all the way to the native state while it remains bound to the chaperone (Stull et al., 2016). This implies that the formation of a hydrophobic core within Im7 is thermodynamically more favorable than Im7's hydrophobic interactions with the chaperone Spy. The data presented in this thesis uncovers two possible solutions to this quandary. A) not all hydrophobic residues in Im7 may be shielded by Spy. Most contacts of Im7 with Spy occur in the region of amino acid 6 to 45 (Quan et al., 2014), thus, C-terminal hydrophobic resides are not protected by Spy to the same degree as those present in the 6-45 region and hence are freely available for desolvation upon folding of Im7. B) The binding enthalpy of Im7 to Spy is positive for both the unfolded and intermediate states (Koldewey et al., 2016; Stull et al., 2016). However, folding enthalpies for globular proteins are usually negative due to the tight packing of the hydrophobic core, which releases more heat than consumed by desolvation of hydrophobic residues upon folding (Table 1). Our X-ray crystallography data suggests that Im7 binds Spy as a dynamic ensemble with moderate affinity, allowing Im7 to explore considerable conformational space (Horowitz et al., 2016a). Hence, the Spy-Im7 complex is stabilized through many weak non-specific interactions, therefore lacks the tight packing that is usually found in globularly folded proteins. Heat released by the formation of weak non-specific interactions likely cannot compensate for the desolvation enthalpy, which may be why the binding enthalpy for Spy and non-native Im7 is positive. This could explain the thermodynamic basis of why Im7 still folds even though its hydrophobic residues are already partially desolvated by Spy: while contacts between Spy and Im7 are non-specific and weak, the core of Im7 is tightly packed, resulting in a difference of binding vs folding enthalpy. This enthalpic difference may favor folding over hydrophobic interactions with the chaperone. Interactions of exposed charged residues keep Im7 bound to Spy, even when hydrophobic burial reduces stabilizing hydrophobic contacts. The reduction in entropic folding energy by chaperone-binding, however, comes with a kinetic penalty. Bound to Spy, Im7 transits from the unfolded to the intermediate state 300-fold slower compared to folding free in solution. In addition, the transition from the intermediate state to the native state is slowed down 30-fold. Since the formation of the intermediate state is driven mostly by the entropic effect of the hydrophobic collapse, which buries hydrophobic surface preferentially bound by

Spy, the hydrophobic contacts between Spy and Im7 may slow down this first folding step more than the second. To my knowledge, folding enthalpies and entropies have not been determined experimentally for Im7 so far. Hence, recording folding kinetics of Im7 in absence and presence of Spy, at a series of temperatures would allow one to dissect of the thermodynamics of chaperone bound client-folding.

A chaperone mediated reduction in the rates of client folding kinetics has been reported, for example, for both GroEL and trigger factor (Corrales and Fersht, 1995; Gary, T. E.; Fersht, 1993; Hoffmann et al., 2010; Hofmann et al., 2010; Sirur and Best, 2013). The broad interactions of the client protein rhodanase with the interior wall of the chaperonin GroEL/ES complex for instance slows down the folding kinetics of rhodanase but leaves the folding pathway rather unaltered (Hofmann et al., 2010; Sirur and Best, 2013). Similarly, the folding of barnase is slowed 100-fold when bound to the apical domains of GroEL in the absence of GroES and ATP (Corrales and Fersht, 1995; Gary, T. E.; Fersht, 1993). While encapsulation of the folding client protein is often seen as essential for the mechanism of GroEL (Hayer-Hartl et al., 2016), apical binding of the client protein may still be a relevant mechanism for GroEL to assist in folding either at low ATP concentrations or for client proteins that are too large to fit into the GroEL chamber (Priya et al., 2013b). Slowing protein folding kinetics through interactions with the chaperone may improve folding yields, especially for proteins which fold through many intermediate states, by preventing the formation of non-native tertiary hydrophobic contacts that kinetically trap the folding protein in a misfolded state.

Recently, folding rate enhancements have been shown for GroEL assisting folding of the TIM Barrel protein, DapA (Georgescauld et al., 2014). However, the proposed underlying mechanism is rather different than the one shown here for Spy. DapA has the same net charge as the GroEL cavity, and hence after release into the cavity, repulsive rather than attractive interactions drive formation of the hydrophobic core in DapA. In addition, it is postulated that spatial confinement destabilizes entropically stabilized folding intermediates, hence remodeling the folding pathway and enhancing folding rates of DapA. In contrast to GroEL mediated DapA folding, Spy doesn't affect the folding pathway of Im7. These fundamental mechanistic differences make a direct comparison between these two different folding processes difficult.

Electrostatic interactions may be a generally essential feature of chaperone function. Spy binds Im7 as well as casein and alpha-lactalbumin with rate constants close to the theoretical diffusion limit, thus client binding by Spy appears to be a diffusion-controlled process. This very rapid binding to its client is probably one of the main features of Spy that allows it to effectively out-compete protein aggregation through kinetic partitioning. Spy drives rapid binding by utilizing electrostatic attraction. Spy has considerable positive charge on its client-binding site which allows for long-range guiding interactions to occur with the client protein thereby driving client-chaperone interaction and increasing the chance of successful complex formation with the non-native aggregation-prone protein. Similarly

fast association rates as well as the involvement of electrostatics in client recognition have been reported for other chaperones (Clerico et al., 2015; Hoffmann et al., 2010; Joachimiak et al., 2014; Karagöz et al., 2014; Perrett et al., 1997). Hence it is conceivable that fast, electrostatically guided client binding is a common and essential feature of various chaperones. However, electrostatic interactions may have roles in this reaction cycle beyond their ability to enhance Spy-Im7 binding rates. They may also have implications regarding Im7 folding while it is bound to Spy. Compared to the effect of hydrophobic interactions between Im7 and Spy that slow folding kinetics, electrostatic interactions may favor helix formation of Im7, as suggested by our crystallography data: The $Im7_{6-45}$ peptide used in our crystallography experiments has an overall net negative charge, which may interfere with secondary and tertiary structure formation in solution through repulsive effects. Such repulsive effects may be screened by interaction with compensatory charged residues on Spy. In agreement with this hypothesis, more natively folded conformations of our Im7 peptide ensemble were found to form more electrostatic interactions and less hydrophobic interactions with Spy than less folded conformations did (Horowitz et al., 2016a). For example, Asp 32 and Asp 35 of $Im7_{6-45}$ are in close proximity and hence likely hamper the formation of helix 2 in the peptide. Both residues have been found to form ionic bonds with positively charged residues on Spy in the unfolded as well as native-like conformations of our Im7 ensemble, indicating that these interactions likely foster the formation of α-helices in Im7 (Horowitz et al., 2016a). This observation is supported by a recent NMR based study, conducted in our lab, which showed that the $Im7_{6-45}$ peptide bound to Spy has an increased helical propensity compared to what it exhibits free in solution (Salmon et al., 2016). Hence, it appears that Spy binding shifts the folding equilibrium of the $Im7_{6-45}$ peptide towards more folded conformations. Very similar effects may be expected for chaperones that assist the folding of oppositely charged client proteins, for example in the case of CCT/TRiC and the overall negatively charged client protein α-actin (Dekker et al., 2011).

Partial client unfolding induced by chaperone binding may be an important factor of the chaperone function of Spy. At the conditions used in our study, binding of Spy to Im7 results in a 7-fold destabilization of the native state. As a result, Im7 partially unfolds as a consequence of binding to Spy (Stull et al., 2016). A similar or more dramatic destabilization of client proteins upon binding to chaperones has been observed for different chaperones and client proteins, and has been postulated to be one mechanism whereby client proteins are pulled out of kinetically trapped misfolded states (Clerico et al., 2015; Farr et al., 2000; Libich et al., 2015; Lin et al., 2008; Marcinowski et al., 2013; Park et al., 2011a; Priya et al., 2013a; Villebeck et al., 2007). For example, the folding equilibrium of a thermodynamically destabilized mutant of the protein SH3 has been shown to be shifted towards the intermediate folding state upon binding to the apical domain of GroEL. Similar to Spy and Im7, SH3 can convert in between native and intermediate folding states while bound to GroEL. In this case, the intermediate state also forms more hydrophobic contacts with GroEL than the native state (Libich et al., 2015). Similarly, GroEL and CCT/TRiC have been demonstrated to unfold kinetically trapped

misfolded states of rhodanase and luciferase (Priya et al., 2013a). In another example, the bacterial Hsp70 homolog, DnaK, has been shown to interact with folded σ^{32}, the heat shock response transcription factor. The interaction with DnaK partially unfolds σ^{32}, rendering it susceptible to degradation (Clerico et al., 2015). Hsp90 has also been shown to interact with mostly folded but unstable client proteins in order to modulate their activity and to stabilize them in a ligand binding competent state. In agreement, it has been shown that Hsp90 binding to its client proteins p53 and CDK4 may lead to partially unfolding (Park et al., 2011b; Verba et al., 2016). While Hsp70 may selectively bind transiently populated, less structured states that expose Hsp70 binding sites (Marcinowski et al., 2013), Hsp90, TRiC as well as GroEL may accomplish this feat by offering broad and heterogeneous binding sites to its clients (Joachimiak et al., 2014; Karagöz and Rüdiger, 2015; Perrett et al., 1997), that induce partial client unfolding upon binding. Hence, it is conceivable that Spy's foldase activity results partially from its ability to unfold misfolded client proteins, such as Im7. Consistent with this idea, Spy variants, selected in our lab, that showed a higher capacity for unfolding native Im7 *in vitro,* also had an increased ability to suppress aggregation and facilitate refolding of a diverse variety of client proteins ((Quan et al., 2014) and data not shown). These mutations increased the hydrophobicity and flexibility of Spy's client binding site. As these mutants were selected for improved ability to protect the intermediate state mimic Im7 L53A I54A *in vivo*, we speculated that these mutations might have the effect of elevating the affinity for the flexible and relatively hydrophobic intermediate folding state of Im7. Hence, Spy's ability to fold or unfold, and hold or release clients may be modulated by changing the content of hydrophobicity of its surface and its overall structural flexibility. The variants our lab selected in Spy that yielded an increased aggregation suppression and client refolding were found at different positions of the Spy monomer. It would be interesting to determine, if combinations of these mutations would yield an even stronger unfolding activity and potential improvement of general chaperone function, or if these combinations have the reciprocal effect, namely converting Spy from a foldase to a holdase, that perhaps inhibits client folding entirely.

Implications of the proposed mechanism on the physiological role of Spy. The *E. coli* periplasmic proteome is particularly vulnerable to naturally occurring protein denaturants such as tannins, ethanol or butanol, because these chemicals can diffuse freely into the periplasmic space, due to the permeability of the outer membrane (Cama et al., 2015). Upon exposure of *E. coli* to chemically-induced protein unfolding in the periplasm, Spy expression is upregulated through the Cpx and Bae envelope stress response pathways that are responsive to protein folding stress. As a consequence, periplasmic concentrations of Spy go from near-zero to > 2 mM, so that Spy can comprise up to 50 % of the periplasmic proteome (Quan et al., 2011; Rutherford et al., 2010; Srivastava et al., 2014). The majority of the proteins, that make up the periplasmic proteome, is negatively charged around neutral pH (Heidary et al., 2014). This complementarity in charge, together with our observation that Spy binds three model clients at rates near the diffusion limit (Koldewey et al., 2016) opens up the

possibility that Spy may bind to a wide range of non-native, aggregation prone periplasmic proteins near-instantaneously through long-range electrostatic interactions. We have observed that Spy binds to three model clients, Im7, α-casein, or α-lactalbumin, with relatively modest binding affinities, which are on the order of 0.2-30 µM (Koldewey et al., 2016; Stull et al., 2016). We do not know exactly with what affinity it binds other client proteins, but given its broad client specificity, we speculate that Spy may exhibit similar micromolar or sub-micromolar binding affinities for other periplasmic proteins. If this is the case, since Spy's physiological concentration when fully induced is > 2 mM following chemical stress, then the vast majority of each of its client proteins will be bound to Spy, effectively preventing their aggregation. Spy's amphiphilic and flexible binding surface not only allows Spy to accommodate denatured proteins of different size and sequence, but extrapolating our results with Im7 to other clients, likely also allows these clients to fold while bound to Spy, once protein denaturant concentrations drop (Stull et al., 2016). This should effectively reduce the concentration of aggregation prone folding intermediates even after stress relief, while allowing its clients to fold to their aggregation-insensitive native state. Folding lowers the affinity of the client protein to Spy and hence may trigger release (Koldewey et al., 2016). This stands in stark contrast to many other stress activated chaperones, such as HdeA, that contain an intrinsic switch that mediates client release and chaperone inactivation upon stress relief (Franzmann et al., 2008; Tapley et al., 2010; Teixeira et al., 2015; Winter and Jakob, 2004a). As Spy-client binding and release is regulated mostly by the client's folding state, Spy may be seen as a folding buffer induced under folding stress conditions that ensures secure folding of periplasmic proteins at the expense of the folding rate (Stull et al., 2016). How and to what extent the concentration of Spy drops after folding stress is unknown. It is conceivable that this may occur slowly through dilution due to cell-growth and -division or it may be accelerated by proteolysis.

4.1 Perspective

While this study provides many details as to how the client protein Im7 folds in presence of the chaperone Spy and how Spy accomplishes this feat (Horowitz et al., 2016a; Koldewey et al., 2016; Stull et al., 2016), there are still a number of unknowns. For example, does the mechanism of self-guided folding while bound to Spy apply to other client proteins that exhibit differences in structure and folding pathways compared to Im7? Applying similar kinetic and structural methods, which helped me uncover chaperone assisted folding of Im7, to other client proteins that also have well-defined folding pathways, may in the future help to see how broadly applicable our model of chaperone function is. Im7 is a 10 kDa protein that forms 1 to 1 stoichiometric complexes with Spy (Horowitz et al., 2016a; Koldewey et al., 2016; Stull et al., 2016). It is conceivable that larger proteins may be bound by multiple Spy dimers in their unfolded state. It would be interesting to determine how such oligomeric Spy-client complexes affect the folding landscape of the bound client protein.

While it appears that the overall folding pathway of Im7 is not substantially affected by Spy (Stull et al., 2016), we do have evidence that on a microscopic level, the chaperone environment may alter folding energetics. Non-native states of Im7 are stabilized through hydrophobic contacts with Spy which slows the folding kinetics(Koldewey et al., 2016; Stull et al., 2016). On the other hand, the chaperone environment seems to stabilize secondary structure of the $Im7_{6to45}$ peptide, potentially through electrostatic interactions that screen repulsive effects hampering folding free in solution (Horowitz et al., 2016a; Salmon et al., 2016). Decades of thorough kinetic analyzes in conjunction with protein engineering, hydrogen-deuterium exchange, and molecular dynamics simulations provided residue-resolution information about every step of the folding pathway of Im7 in solution (Bartlett and Radford, 2010; Capaldi et al., 2001, 2002; Friel et al., 2009; Gsponer et al., 2006; Pashley et al., 2012). Conducting similar types of experiments (phi-value analysis (Jackson et al., 1993)) in the presence of Spy may provide insights as to how Spy affects the folding pathway of Im7 on a microscopic level.

The folding pathway of Im9 has been characterized in the laboratory of Sheena Radford who also characterized Im7 folding (Ferguson et al., 1999; Gorski et al., 2001). Since recent research conducted in our lab indicates that Im9 and Spy interact (data not shown), characterizing how Spy affects Im9 folding is a very promising direction of research, since interestingly, Im9 only folds through an intermediate state at low pH while at neutral pH Im9 folding follows a two-step folding mechanism (Ferguson et al., 1999; Gorski et al., 2001). The intermediate state is most populated at pH 5, which is close to Im9's isoelectric point (pI = 4.53), indicating that a reduction of net negative charge stabilizes the intermediate folding state. A similar compensation of negative charge may be achieved by Spy-binding, which offers an excess of compensatory positive charge. Hence, one could speculate that binding to Spy may induce the formation of an intermediate folding state, which would be interesting to test.

In addition, it has been shown that Spy efficiently inhibits the formation of functional amyloids formed by the *E. coli* protein CsgA (Evans et al., 2011). Functional amyloids are produced by *E. coli* to form extracellular matrices vital for biofilm formation, and interestingly these fibrils are structurally similar to amyloid fibrils that are associated with numerous human diseases (Dueholm et al., 2011). Examples are Alzheimer's or Parkinson's disease, which are caused in part by the intrinsically disordered protein alpha-synuclein (Kim et al., 2004). Studying the structural and kinetic aspects that underlie the mechanism of amyloid suppression by Spy may provide insights as to how chaperones in general affect amyloid biogenesis.

And finally, as already mentioned above, testing our proposed mechanism on other folding chaperones, such as TRiC, Hsp90, Hsp70 or Trigger Factor would be informative. In this regard, single molecule experiments using Hsp70 have very recently been conducted that suggest that Hsp70 may also allow folding of client proteins while bound to the chaperone (Mashaghi et al., 2016).

5 References

Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 213–221.

Afonine, P. V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect. D Biol. Crystallogr. *68*, 352–367.

Akerfelt, M., Morimoto, R.I., and Sistonen, L. (2010). Heat shock factors: integrators of cell stress, development and lifespan. Nat. Rev. Mol. Cell Biol. *11*, 545–555.

Allen, W.J., Phan, G., and Waksman, G. (2009). Structural Biology of Periplasmic Chaperones. Adv. PROTEIN Chem. Struct. Biol. - Vol. 78 51–97.

Ambrose, A.J., Fenton, W., Mason, D.J., Chapman, E., and Horwich, A.L. (2015). Unfolded DapA forms aggregates when diluted into free solution, confounding comparison with folding by the GroEL/GroES chaperonin system. FEBS Lett. *589*, 497–499.

Anfinsen, C.B. (1973). Principles that Govern the Folding of Protein Chains. Science (80-.). *181*, 223–230.

Anfinsen, C.B., Haber, E., Sela, M., and White, F.H. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Natl. Acad. Sci. U. S. A. *47*, 1309–1314.

Apetri, A.C., and Horwich, A.L. (2008). Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. Proc. Natl. Acad. Sci. U. S. A. *105*, 17351–17355.

Aponte, R.A., Zimmermann, S., and Reinstein, J. (2010). Directed Evolution of the DnaK Chaperone: Mutations in the Lid Domain Result in Enhanced Chaperone Activity. J. Mol. Biol. *399*, 154–167.

Appia-Ayme, C., Hall, A., Patrick, E., Rajadurai, S., Clarke, T. a, and Rowley, G. (2012). ZraP is a periplasmic molecular chaperone and a repressor of the zinc-responsive two-component regulator ZraSR. Biochem. J. *442*, 85–93.

Arai, M., and Kuwajima, K. (2000). Role of the molten globule state in protein folding. Adv. Protein Chem. *53*, 209–282.

Assimon, V.A., Gillies, A.T., Rauch, J.N., and Gestwicki, J.E. (2013). Hsp70 protein complexes as drug targets. Curr. Pharm. Des. *19*, 404–417.

Azia, A., Unger, R., and Horovitz, A. (2012). What distinguishes GroEL substrates from other Escherichia coli proteins? FEBS J. *279*, 543–550.

Bai, Y., and Englander, S.W. (1996). Future Directions in Folding: The Multi-State Nature of Protein Structure. PROTEINS Struct. Funct. Genet. *24*, 145–151.

Bai, Y., Sosnick, T., Mayne, L., and Englander, S. (1995). Protein folding intermediates: native-state hydrogen exchange. Science (80-.). *269*, 192–197.

Baker, D. (1998). Metastable states and folding free energy barriers. Nat. Struct. Biol. *5*, 1021–1024.

Balchin, D., Hayer-Hartl, M., and Hartl, F.U. (2016). In vivo aspects of protein folding and quality control. Science (80-.). *353*, aac4354-aac4354.

Baldwin, R.L. (1986). Temperature Dependence of the Hydrophobic Interaction in Protein Folding. Proc. Natl. Acad. Sci. *83*, 8069–8072.

Baldwin, R.L. (1995). The nature of protein folding pathways: The classical versus the new view. J. Mol. Biol. *5*, 103–109.

Baldwin, R.L. (2007). Energetics of Protein Folding. J Mol Biol *371*, 283–301.

Baldwin, R.L. (2008). The search for folding intermediates and the mechanism of protein folding. Annu Rev Biophys *37*, 1–21.

Baldwin, R.L., and Rose, G.D. (1999). Is protein folding hierarchic? II. Folding intermediates and transition states. Trends Biochem. Sci. *24*, 77–83.

Bardwell, J.C. a, and Jakob, U. (2012). Conditional disorder in chaperone action. Trends Biochem. Sci. *37*, 517–525.

Barrick, D. (2009). What have we learned from the studies of two-state folders, and what are the unanswered questions about two-state protein folding? Phys. Biol. *6*, 15001.

Bartlett, A.I., and Radford, S.E. (2009). An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. Nat. Struct. Mol. Biol. *16*, 582–588.

Bartlett, A.I., and Radford, S.E. (2010). Desolvation and development of specific hydrophobic core packing during Im7 folding. J. Mol. Biol. *396*, 1329–1345.

Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011). iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. Sect. D Biol. Crystallogr. *67*, 271–281.

Beauchamp, K. a., McGibbon, R., Lin, Y.-S., and Pande, V.S. (2012). Simple few-state models reveal hidden complexity in protein folding. Proc. Natl. Acad. Sci. *109*, 17807–17813.

Bédard, S., Krishna, M.M.G., Mayne, L., and Englander, S.W. (2008). Protein folding: independent unrelated pathways or predetermined pathway with optional errors. Proc. Natl. Acad. Sci. U. S. A. *105*, 7182–7187.

Bloom, J.D., Raval, A., and Wilke, C.O. (2007). Thermodynamics of neutral protein evolution. Genetics *175*, 255–266.

Bolognesi, B., Kumita, J.R., Barros, T.P., Esbjorner, E.K., Luheshi, L.M., Crowther, D.C., Wilson, M.R., Dobson, C.M., Favrin, G., and Yerbury, J.J. (2010). ANS Binding Reveals Common Features of Cytotoxic Amyloid Species. ACS Chem Biol *5*, 735–740.

Bowler, B.E. (2012). Residual structure in unfolded proteins. Curr. Opin. Struct. Biol. *22*, 4–13.

Brandts, J.F., Halvorson, H.R., and Brennan, M. (1975). Consideration of the Possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. Biochemistry *14*, 4953–4963.

Braselmann, E., Chaney, J.L., and Clark, P.L. (2013). Folding the proteome. Trends Biochem. Sci. *38*, 337–344.

Brinker, A., Pfeifer, G., Kerner, M.J., Naylor, D.J., Hartl, F.U., and Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. Cell *107*, 223–233.

Brockwell, D.J., and Radford, S.E. (2007). Intermediates: ubiquitous species on folding energy landscapes? Curr. Opin. Struct. Biol. *17*, 30–37.

Bryngelson, J.D., Onuchic, J.N., Socci, N.D., and Wolynes, P.G. (1995). Funnels, pathways, and the

energy landscape of protein folding: A synthesis. Proteins Struct. Funct. Genet. *21*, 167–195.

Buchberger, A. (2013). Regulated Proteolysis in Microorganisms.

Burger, V.M., Nolasco, D.O., and Stultz, C.M. (2016). Expanding the Range of Protein Function at the Far End of the Order-Structure Continuum. J. Biol. Chem. *291*, 6706–6713.

Burmann, B.M., and Hiller, S. (2015). Chaperones and chaperone-substrate complexes: Dynamic playgrounds for NMR spectroscopists. Prog. Nucl. Magn. Reson. Spectrosc. *86–87*, 41–64.

Bury-Moné, S., Nomane, Y., Reymond, N., Barbet, R., Jacquet, E., Imbeaud, S., Jacq, A., and Bouloc, P. (2009). Global analysis of extracytoplasmic stress signaling in Escherichia coli. PLoS Genet. *5*, e1000651.

Cama, J., Bajaj, H., Pagliara, S., Maier, T., Braun, Y., Winterhalter, M., and Keyser, U.F. (2015). Quantification of Fluoroquinolone Uptake through the Outer Membrane Channel OmpF of Escherichia coli. J. Am. Chem. Soc. *137*, 13836–13843.

Capaldi, A.P., Shastry, M.C.R., Kleanthous, C., Roder, H., and Radford, S.E. (2001). Ultrarapid mixing experiments reveal that Im7 folds via an on-pathway intermediate. Nat. Struct. Biol. *8*, 68–72.

Capaldi, A.P., Kleanthous, C., and Radford, S.E. (2002). Im7 folding mechanism: misfolding on a path to the native state. Nat. Struct. Biol. *9*, 209–216.

Chak, K.F., Safo, M.K., Ku, W.Y., Hsieh, S.Y., and Yuan, H.S. (1996). The crystal structure of the immunity protein of colicin E7 suggests a possible colicin-interacting surface. Proc. Natl. Acad. Sci. U. S. A. *93*, 6437–6442.

Chakraborty, K., Chatila, M., Sinha, J., Shi, Q., Poschner, B.C., Sikor, M., Jiang, G., Lamb, D.C., Hartl, F.U., and Hayer-Hartl, M. (2010). Chaperonin-Catalyzed Rescue of Kinetically Trapped States in Protein Folding. Cell *142*, 112–122.

Chandler, D. (2005). Interfaces and the driving force of hydrophobic assembly. Nature *437*, 640–647.

Chaudhry, C., Farr, G.W., Todd, M.J., Rye, H.S., Brunger, A.T., Adams, P.D., Horwich, A.L., and Sigler, P.B. (2003). Role of the g -phosphate of ATP in triggering protein folding by GroEL ± GroES : function , structure and energetics. *22*, 4877–4887.

Chaudhuri, T.K., Verma, V.K., and Maheshwari, A. (2009). GroEL assisted folding of large polypeptide substrates in Escherichia coli: Present scenario and assignments for the future. Prog. Biophys. Mol. Biol. *99*, 42–50.

Chen, T., and Chan, H.S. (2015). Native contact density and nonnative hydrophobic effects in the folding of bacterial immunity proteins. PLoS Comput. Biol. *11*, e1004260.

Chen, D.H., Madan, D., Weaver, J., Lin, Z., Schröder, G.F., Chiu, W., and Rye, H.S. (2013). Visualizing GroEL/ES in the act of encapsulating a folding protein. Cell *153*.

Chiti, F., and Dobson, C.M. (2006). Protein Misfolding, Functional Amyloid, and Human Disease. Annu. Rev. Biochem. *75*, 333–366.

Chodera, J.D., and Mobley, D.L. (2013). Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. Annu. Rev. Biophys. *42*, 121–142.

Chung, H.S., McHale, K., Louis, J.M., and Eaton, W.A. (2012). Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. Science (80-.). *335*, 981–984.

Ciryam, P., Tartaglia, G., Morimoto, R.I., Dobson, C.M., and Vendruscolo, M. (2013a). Widespread Aggregation and Neurodegenerative Diseases Are Associated with Supersaturated Proteins. Cell Rep.

5, 781–790.

Ciryam, P., Morimoto, R.I., Vendruscolo, M., Dobson, C.M., O'Brien, E.P., and O'Brien, E.P. (2013b). In vivo translation rates can substantially delay the cotranslational folding of the Escherichia coli cytosolic proteome. Proc. Natl. Acad. Sci. U. S. A. *110*, E132-40.

Clare, D.K., Vasishtan, D., Stagg, S., Quispe, J., Farr, G.W., Topf, M., Horwich, A.L., and Saibil, H.R. (2012). ATP-triggered conformational changes delineate substrate-binding and -folding mechanics of the GroEL chaperonin. Cell *149*, 113–123.

Clerico, E.M., Tilitsky, J.M., Meng, W., and Gierasch, L.M. (2015). How Hsp70 Molecular Machines Interact with Their Substrates to Mediate Diverse Physiological Functions. J. Mol. Biol. *427*, 1575–1588.

Cohen-kaplan, V., Livneh, I., Avni, N., Cohen-rosenzweig, C., and Ciechanover, A. (2016). The ubiquitin-proteasome system and autophagy: Coordinated and independent activities. Int. J. Biochem. Cell Biol. *In press*, 1–16.

Corrales, F.J., and Fersht, A.R. (1995). The folding of GroEL-bound barnase chaperonin-mediated protein folding model for chaperonin-mediated protein folding. *92*, 5326–5330.

Creamer, L.K., Richardson, T., and Parry, D. a (1981). Secondary structure of bovine alpha s1- and beta-casein in solution. Arch. Biochem. Biophys. *211*, 689–696.

D. Zanchi et al (2008). Tannin-assisted aggregation of natively unfolded proteins. EPL (Europhysics Lett. *82*, 58001.

Dahiya, V., and Chaudhuri, T.K. (2013). Functional intermediate in the refolding pathway of a large and multidomain protein malate synthase G. Biochemistry *52*, 4517–4530.

Dahl, J.U., Koldewey, P., Salmon, L., Horowitz, S., Bardwell, J.C.A., and Jakob, U. (2015). HdeB functions as an acid-protective chaperone in bacteria. J. Biol. Chem. *290*, 65–75.

Dahl, J.U., Koldewey, P., Bardwell, J.C.A., and Jakob, U. (2016). Detection of the pH-dependent Activity of Escherichia coli Chaperone HdeB In Vitro and In Vivo. J Vis Exp.

Dekker, C., Roe, S.M., McCormack, E.A., Beuron, F., Pearl, L.H., and Willison, K.R. (2011). The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. EMBO J. *30*, 3078–3090.

Dellus-Gur, E., Toth-Petroczy, A., Elias, M., and Tawfik, D.S. (2013). What makes a protein fold amenable to functional innovation? fold polarity and stability trade-offs. J. Mol. Biol. *425*, 2609–2621.

Dennis, C.A., Videler, H., Pauptit, R.A., Wallis, R., James, R., Moore, G.R., and Kleanthous, C. (1998). A structural comparison of the colicin immunity proteins Im7 and Im9 gives new insights into the molecular determinants of immunity-protein specificity. Biochem. J. *333*, 183–191.

DePristo, M. a, Weinreich, D.M., and Hartl, D.L. (2005). Missense meanderings in sequence space: a biophysical view of protein evolution. Nat. Rev. Genet. *6*, 678–687.

Dill, K.A. (1990). Dominant forces in protein folding. Biochemistry *29*, 7133–7155.

Dill, K. a (1985). Theory for the folding and stability of globular proteins. Biochemistry *24*, 1501–1509.

Dill, K.A., and MacCallum, J.L. (2012). The Protein-Folding Problem, 50 Years On. Science (80-.). *338*, 1042–1046.

Dill, K.A., Ozkan, S.B., Shell, M.S., and Weikl, T.R. (2008). The Protein Folding Problem. Annu. Rev. Biophys *37*, 289–316.

Dinner, a R., Sali, a, Smith, L.J., Dobson, C.M., and Karplus, M. (2000). Understanding protein folding via free-energy surfaces from theory and experiment. Trends Biochem. Sci. *25*, 331–339.

Docter, B.E., Horowitz, S., Gray, M.J., Jakob, U., and Bardwell, J.C.A. (2016). Do nucleic acids moonlight as molecular chaperones? Nucleic Acids Res. gkw291.

Dueholm, M.S., Nielsen, S.B., Hein, K.L., Nissen, P., Chapman, M., Christiansen, G., Nielsen, P.H., and Otzen, D.E. (2011). Fibrillation of the major curli subunit CsgA under a wide range of conditions implies a robust design of aggregation. Biochemistry *50*, 8281–8290.

Le Duff, C.S., Whittaker, S.B.M., Radford, S.E., and Moore, G.R. (2006). Characterisation of the Conformational Properties of Urea-unfolded Im7: Implications for the Early Stages of Protein Folding. J. Mol. Biol. *364*, 824–835.

Dyson, H.J., and Wright, P.E. (2002). Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. Adv. Protein Chem. *62*, 311–340.

Ellis, R.J., and Minton, A.P. (2006). Protein aggregation in crowded environments. Biol. Chem. *387*, 485–497.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 486–501.

Englander, S.W., and Mayne, L. (2014). The nature of protein folding pathways. Proc. Natl. Acad. Sci. *111*, 15873–15880.

Englander, S.W., Mayne, L., and Krishna, M.M.G. (2007). Protein folding and misfolding: mechanism and principles. Q. Rev. Biophys. *40*, 287–326.

Espargaró, A., Castillo, V., de Groot, N.S., and Ventura, S. (2008). The in Vivo and in Vitro Aggregation Properties of Globular Proteins Correlate With Their Conformational Stability: The SH3 Case. J. Mol. Biol. *378*, 1116–1131.

Estes, B., Hsu, Y.R., Tam, L.T., Sheng, J., Stevens, J., and Haldankar, R. (2015). Uncovering methods for the prevention of protein aggregation and improvement of product quality in a transient expression system. Biotechnol. Prog. *31*, 258–267.

Evans, M.L., Schmidt, J.C., Ilbert, M., Doyle, S.M., Quan, S., Bardwell, J.C. a., Jakob, U., Wickner, S., and Chapman, M.R. (2011). E. coli chaperones DnaK, Hsp33 and Spy inhibit bacterial functional amyloid assembly. Prion *5*, 323–334.

Farr, G.W., Furtak, K., Rowland, M.B., Ranson, N.A., Saibil, H.R., Kirchhausen, T., and Horwich, A.L. (2000). Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. Cell *100*, 561–573.

Fass, D. (2012). Disulfide Bonding in Protein Biophysics. Annu. Rev. Biophys. *41*, 63–79.

Fekkes, P., Den Blaauwen, T., and Driessen, A.J.M. (1995). Diffusion-limited interaction between unfolded polypeptides and the Escherichia coli chaperone SecB. Biochemistry *34*, 10078–10085.

Felitsky, D.J., Lietzow, M. a, Dyson, H.J., and Wright, P.E. (2008). Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. Proc. Natl. Acad. Sci. U. S. A. *105*, 6278–6283.

Ferguson, N., Capaldi, A.P., James, R., Kleanthous, C., and Radford, S.E. (1999). Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. J. Mol. Biol. *286*, 1597–1608.

Fersht, A.R. (1993). Protein folding and stability: the pathway of folding of barnase. FEBS Lett. *325*,

5–16.

Fersht, a R. (1997). Nucleation mechanisms in protein folding. Curr. Opin. Struct. Biol. *7*, 3–9.

Fersht, A.R., and Sato, S. (2004). Phi-value analysis and the nature of protein-folding transition states. Proc. Natl. Acad. Sci. U. S. A. *101*, 7976–7981.

Figueiredo, A.M., Whittaker, S.B.M., Knowling, S.E., Radford, S.E., and Moore, G.R. (2013). Conformational dynamics is more important than helical propensity for the folding of the all α-helical protein Im7. Protein Sci. *22*, 1722–1738.

Foit, L., Morgan, G.J., Kern, M.J., Steimer, L.R., von Hacht, A. a, Titchmarsh, J., Warriner, S.L., Radford, S.E., and Bardwell, J.C. a (2009). Optimizing protein stability in vivo. Mol. Cell *36*, 861–871.

Foit, L., George, J.S., Zhang, B.W., Brooks, C.L., and Bardwell, J.C.A. (2013). Chaperone activation by unfolding. Proc. Natl. Acad. Sci. *110*, E1254-62.

Fränzel, B., Penkova, M., Frese, C., Metzler-Nolte, N., and Andreas Wolters, D. (2012). Escherichia coli exhibits a membrane-related response to a small arginine- and tryptophan-rich antimicrobial peptide. Proteomics *12*, 2319–2330.

Franzmann, T.M., Menhorn, P., Walter, S., and Buchner, J. (2008). Activation of the Chaperone Hsp26 Is Controlled by the Rearrangement of Its Thermosensor Domain. Mol. Cell *29*, 207–216.

Friel, C.T., Capaldi, A.P., and Radford, S.E. (2003). Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: Similarities and differences in the folding of homologous proteins. J. Mol. Biol. *326*, 293–305.

Friel, C.T., Smith, D.A., Vendruscolo, M., Gsponer, J., and Radford, S.E. (2009). The mechanism of folding of Im7 reveals competition between functional and kinetic evolutionary constraints. Nat. Struct. Mol. Biol. *16*, 318–324.

Gary, T. E.; Fersht, A.R. (1993). Refolding of Barnase in the Presence of GroE. J. Mol. Biol. *232*, 1197–1207.

Georgescauld, F., Popova, K., Gupta, A.J., Bracher, A., Engen, J.R., Hayer-Hartl, M., and Hartl, F.U. (2014). GroEL/ES chaperonin modulates the mechanism and accelerates the rate of TIM-barrel domain folding. Cell *157*, 922–934.

Gershenson, A., Gierasch, L.M., Pastore, A., and Radford, S.E. (2014). Energy landscapes of functional proteins are inherently risky. Nat. Chem. Biol. *10*, 884–891.

Gianni, S., Dogan, J., and Jemth, P. (2016). Coupled binding and folding of intrinsically disordered proteins: What can we learn from kinetics? Curr. Opin. Struct. Biol. *36*, 18–24.

Goemans, C., Denoncin, K., and Collet, J.F. (2014). Folding mechanisms of periplasmic proteins. Biochim. Biophys. Acta - Mol. Cell Res. *1843*, 1517–1528.

Goncearenco, A., and Berezovsky, I.N. (2015). Protein function from its emergence to diversity in contemporary proteins. Phys. Biol. *12*, 45002.

Gorski, S.A., Capaldi, A.P., Kleanthous, C., and Radford, S.E. (2001). Acidic conditions stabilise intermediates populated during the folding of Im7 and Im9. J. Mol. Biol. *312*, 849–863.

Gorski, S.A., Le Duff, C.S., Capaldi, A.P., Kalverda, A.P., Beddard, G.S., Moore, G.R., and Radford, S.E. (2004). Equilibrium Hydrogen Exchange Reveals Extensive Hydrogen Bonded Secondary Structure in the On-pathway Intermediate of Im7. J. Mol. Biol. *337*, 183–193.

Gosavi, S. (2013). Understanding the Folding-Function Tradeoff in Proteins. PLoS One *8*.

Gosavi, S., Whitford, P.C., Jennings, P.A., and Onuchic, J.N. (2008). Extracting function from a beta - trefoil folding motif. Proc. Natl. Acad. Sci. *105*, 10384–10389.

Grahnen, J.A., Nandakumar, P., Kubelka, J., and Liberles, D.A. (2011). Biophysical and structural considerations for protein sequence evolution. BMC Evol. Biol. *11*, 361.

Gray, M., Wholey, W.Y., Wagner, N., Cremers, C., Mueller-Schickert, A., Hock, N., Krieger, A., Smith, E., Bender, R., Bardwell, J.A., et al. (2014). Polyphosphate Is a Primordial Chaperone. Mol. Cell *53*, 689–699.

Groitl, B., Horowitz, S., Makepeace, K.A.T., Petrotchenko, E. V., Borchers, C.H., Reichmann, D., Bardwell, J.C.A., and Jakob, U. (2016). Protein unfolding as a switch from self-recognition to high-affinity client binding. Nat. Commun. *7*, 10357.

Gruebele, M. (1999). The Fas Protein Folding Problem. Annu. Rev. Biophys. Chem. *50*, 485–516.

Gruebele, M. (2005). Downhill protein folding: Evolution meets physics. Comptes Rendus - Biol. *328*, 701–712.

Gruebele, M., Dave, K., and Sukenik, S. (2016). Globular Protein Folding In Vitro and In Vivo. Annu. Rev. Biophys *45*, 233–251.

Gsponer, J., Hopearuoho, H., Whittaker, S.B.-M., Spence, G.R., Moore, G.R., Paci, E., Radford, S.E., and Vendruscolo, M. (2006). Determination of an ensemble of structures representing the intermediate state of the bacterial immunity protein Im7. Proc. Natl. Acad. Sci. U. S. A. *103*, 99–104.

Gupta, A.J., Haldar, S., Miličić, G., Hartl, F.U., and Hayer-Hartl, M. (2014). Active cage mechanism of chaperonin-assisted protein folding demonstrated at single-molecule level. J. Mol. Biol. *426*, 2739–2754.

Hagenmaier, S., Stierhof, Y., and Henning, U.L.F. (1997). A New Periplasmic Protein of Escherichia coli Which Is Synthesized in Spheroplasts but Not in Intact Cells. Microbiology *179*, 2073–2076.

Hagn, F., Lagleder, S., Retzlaff, M., Rohrberg, J., Demmer, O., Richter, K., Buchner, J., and Kessler, H. (2011). Structural analysis of the interaction between Hsp90 and the tumor suppressor protein p53. Nat. Struct. Mol. Biol. *18*, 1086–1093.

Haran, G. (2012). How, when and why proteins collapse: The relation to folding. Curr. Opin. Struct. Biol. *22*, 14–20.

Hartl, F.U. (1996). Molecular chaperones in cellular protein folding. Nature *381*, 571–579.

Hayer-Hartl, M., Bracher, A., and Hartl, F.U. (2016). The GroEL-GroES Chaperonin Machine: A Nano-Cage for Protein Folding. Trends Biochem. Sci. *41*, 62–76.

Heidary, S., Rahim, R.A., Eissazadeh, S., Moeini, H., Chor, A.L.T., and Abdullah, M.P. (2014). Proteome analysis of Escherichia coli periplasmic proteins in response to over-expression of recombinant human interferon α2b. Biotechnol. Lett. *36*, 1479–1484.

Hetz, C., Chevet, E., and Oakes, S.A. (2015). Proteostasis control by the unfolded protein response. Nat. Cell Biol. *17*, 829–838.

Hipp, M.S., Park, S.H., and Hartl, U.U. (2014). Proteostasis impairment in protein-misfolding and - aggregation diseases. Trends Cell Biol. *24*, 506–514.

Hoagland, P.D., Unruh, J.J., Wickham, E.D., and Farrell, H.M. (2001). Secondary structure of bovine alphaS2-casein: theoretical and experimental approaches. J. Dairy Sci. *84*, 1944–1949.

Hoffmann, A., Bukau, B., and Kramer, G. (2010). Structure and function of the molecular chaperone

Trigger Factor. Biochim. Biophys. Acta *1803*, 650–661.

Hofmann, H., Hillger, F., Pfeil, S.H., Hoffmann, A., Streich, D., Haenni, D., Nettels, D., Lipman, E. a, and Schuler, B. (2010). Single-molecule spectroscopy of protein folding in a chaperonin cage. Proc. Natl. Acad. Sci. U. S. A. *107*, 11793–11798.

Horovitz, a., and Fersht, a. R. (1992). Co-operative interactions during protein folding. J. Mol. Biol. *224*, 733–740.

Horowitz, S., Koldewey, P., and Bardwell, J.C. (2014). Undergraduates improve upon published crystal structure in class assignment. Biochem. Mol. Biol. Educ. *42*, 398–404.

Horowitz, S., Salmon, L., Koldewey, P., Ahlstrom, L.S., Martin, R., Quan, S., Afonine, P. V, van den Bedem, H., Wang, L., Xu, Q., et al. (2016a). Visualizing chaperone-assisted protein folding. Nat. Struct. Mol. Biol. *23*, 691–697.

Horowitz, S., Koepnick, B., Martin, R., Tymieniecki, A., Winburn, A.A., Cooper, S., Flatten, J., Rogawski, D.S., Koropatkin, N.M., Hailu, T.T., et al. (2016b). Determining crystal structures through crowdsourcing and coursework. Nat. Commun. *7*, 12549.

Horst, R., Fenton, W.A., Englander, S.W., Wüthrich, K., and Horwich, A.L. (2007). Folding Trajectories of Human Dihydrofolate Reductase Inside the GroEL–GroES Chaperonin Cavity and Free in Solution. Proc. Natl. Acad. Sci. *104*, 20788–20792.

Horwich, A.L., and Fenton, W. a (2009). Chaperonin-mediated protein folding: using a central cavity to kinetically assist polypeptide chain folding.

Horwich, A.L., Apetri, A.C., and Fenton, W. a (2009). The GroEL/GroES cis cavity as a passive anti-aggregation device. FEBS Lett. *583*, 2654–2662.

Hu, W., Walters, B.T., Kan, Z., Mayne, L., Rosen, L.E., Marqusee, S., and Englander, S.W. (2013). Stepwise protein folding at near amino acid resolution by hydrogen exchange and mass spectrometry. Proc. Natl. Acad. Sci. U. S. A. *110*, 7684–7689.

Huang, C., Rossi, P., Saio, T., and Kalodimos, C.G. (2016). Structural basis for the antifolding activity of a molecular chaperone. Nature *537*, 202–206.

Ignatova, Z., and Gierasch, L.M. (2006). Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant. Proc. Natl. Acad. Sci. *103*, 13357–13361.

Ignatova, Z., Krishnan, B., Bombardier, J.P., Marcelino, A.M.C., Hong, J., and Gierasch, L.M. (2007). Folding and Aggregation of a b-Clam Protein in the Cell. Pept. Sci. *88*, 157–163.

Ilardo, M., Meringer, M., Freeland, S., Rasulev, B., and Cleaves, H.J. (2015). Extraordinarily adaptive properties of the genetically encoded amino acids. Sci. Rep. *5*, 9414.

Jackson, S.E., ElMasry, N., and Fersht, A.R. (1993). Structure of the hydrophobic core in the transition state for folding of chymotrypsin inhibitor 2: A critical test of the protein engineering method of analysis. Biochemistry *32*, 11270–11278.

Jäger, M., Zhang, Y., Bieschke, J., Nguyen, H., Dendle, M., Bowman, M.E., Noel, J.P., Gruebele, M., and Kelly, J.W. (2006). Structure-function-folding relationship in a WW domain. Proc. Natl. Acad. Sci. U. S. A. *103*, 10648–10653.

James, R., Kleanthous, C., and Moore, G.R. (1996). The biology of E colicins: Paradigms and paradoxes. Microbiology *142*, 1569–1580.

Jelesarov, I., and Bosshard, H.R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. J. Mol. Recognit. *12*, 3–18.

Jewett, A.I., and Shea, J.E. (2006). Folding on the Chaperone: Yield Enhancement Through Loose Binding. J. Mol. Biol. *363*, 945–957.

Jiang, Y., Douglas, N.R., Conley, N.R., Miller, E.J., Frydman, J., and Moerner, W.E. (2011). Sensing cooperativity in ATP hydrolysis for single multisubunit enzymes in solution. Proc. Natl. Acad. Sci. U. S. A. *108*, 16962–16967.

Joachimiak, L.A., Walzthoeni, T., Liu, C.W., Aebersold, R., and Frydman, J. (2014). The Structural Basis of Substrate Recognition by the Eukaryotic Chaperonin TRiC/CCT. Cell *159*, 1042–1055.

Johnson, K.A., Simpson, Z.B., and Blom, T. (2009). Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data. Anal. Biochem. *387*, 20–29.

Kampinga, H.H., and Craig, E.A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat. Rev. Mol. Cell Biol. *11*, 579–592.

Karagöz, G.E., and Rüdiger, S.G.D. (2015). Hsp90 interaction with clients. Trends Biochem. Sci. *40*, 117–125.

Karagöz, G.E., Duarte, A.M.S., Akoury, E., Ippel, H., Biernat, J., Morán Luengo, T., Radli, M., Didenko, T., Nordhues, B.A., Veprintsev, D.B., et al. (2014). Hsp90-tau complex reveals molecular basis for specificity in chaperone action. Cell *156*, 963–974.

Karanicolas, J., and Brooks III, C.L. (2002). The origins of asymmetry in the folding transition states of protein L and protein G. Protein Sci. *11*, 2351–2361.

Karplus, M. (1997). The Levinthal paradox: yesterday and today. Fold. Des. *2*, S69–S75.

Karplus, M. (2011). Behind the folding funnel diagram. Nat. Chem. Biol. *7*, 401–404.

Kazmirski, S.L., Wong, K.B., Freund, S.M., Tan, Y.J., Fersht, a R., and Daggett, V. (2001). Protein folding from a highly disordered denatured state: the folding pathway of chymotrypsin inhibitor 2 at atomic resolution. Proc. Natl. Acad. Sci. U. S. A. *98*, 4349–4354.

Kellis, J.T., Nyberg, K., Sali, D., and Fersht, a R. (1988). Contribution of hydrophobic interactions to protein stability. Nature *333*, 784–786.

Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., et al. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell *122*, 209–220.

Khorasanizadeh, S., Peters, I.D., and Roder, H. (1996). Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. Nat. Struct. Biol. *3*, 193–205.

Kim, S., Seo, J.H., and Suh, Y.H. (2004). a-Synuclein, Parkinson's disease, and Alzheimer's disease. Park. Relat. Disord. *10*, 9–13.

Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Ulrich Hartl, F. (2013). Molecular Chaperone Functions in Protein Folding and Proteostasis.

Klein-Seetharaman, J., Oikawa, M., Grimshaw, S.B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L.J., Dobson, C.M., and Schwalbe, H. (2002). Long-Range Interactions Within a Nonnative Protein. Science (80-.). *295*, 1719–1722.

Knowling, S.E., Figueiredo, A.M., Whittaker, S.B.M., Moore, G.R., and Radford, S.E. (2009). Amino Acid Insertion Reveals a Necessary Three-Helical Intermediate in the Folding Pathway of the Colicin E7 Immunity Protein Im7. J. Mol. Biol. *392*, 1074–1086.

Ko, T.P., Liao, C.C., Ku, W.Y., Chak, K.F., and Yuan, H.S. (1999). The crystal structure of the DNase domain of colicin E7 in complex with its inhibitor Im7 protein. Structure *7*, 91–102.

Koczan, D., Hobom, G., and Seyfert, H.M. (1991). Genomic organization of the bovine alpha-S1 casein gene. Nucleic Acids Res. *19*, 5591–5596.

Koldewey, P., Stull, F., Horowitz, S., Martin, R., and Bardwell, J.C.A. (2016). Forces Driving Chaperone Action. Cell *166*, 369–379.

Kozlov, A.G., and Lohman, T.M. (2002). Stopped-flow studies of the kinetics of single-stranded DNA binding and wrapping around the Escherichia coli SSB tetramer. Biochemistry *41*, 6032–6044.

Krishna, M.M.G., Englander, S.W., and Englander, S.W. (2007). A unified mechanism for protein folding : Predetermined pathways with optional errors A unified mechanism for protein folding : Predetermined pathways with optional errors. 449–464.

Kubelka, J., Hofrichter, J., and Eaton, W.A. (2004). The protein folding "speed limit." Curr. Opin. Struct. Biol. *14*, 76–88.

Kwon, E., Kim, D.Y., Gross, C. a, Gross, J.D., and Kim, K.K. (2010). The crystal structure Escherichia coli Spy. Protein Sci. *19*, 2252–2259.

Labbadia, J., and Morimoto, R.I. (2015). The Biology of Proteostasis in Aging and Disease. Annu. Rev. Biochem. 1–30.

Lakowicz, J.R. (2006). Principles of fluorescence spectroscopy.

Lebowitz, J., Lewis, M.S., and Schuck, P. (2002). Modern analytical ultracentrifugation in protein science: a tutorial review. Protein Sci. *11*, 2067–2079.

Leidig, C., Bange, G., Kopp, J., Amlacher, S., Aravind, A., Wickles, S., Witte, G., Hurt, E., Beckmann, R., and Sinning, I. (2013). Structural characterization of a eukaryotic chaperone--the ribosome-associated complex. Nat. Struct. Mol. Biol. *20*, 23–28.

Leitner, A., Joachimiak, L.A., Bracher, A., Mo, L., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., Ludtke, S., et al. (2012). The Molecular Architecture of the Eukaryotic Chaperonin TRiC / CCT. 814–825.

Levinthal, C. (1968). Are there pathways for protein folding? J. Chim. Phys. Physico-Chimie Biol. *65*, 44–45.

Levinthal, C. (1969). How to fold graciously. Mossbauer Spectrosc. Biol. Syst. Proc. Univ. Illinois Bull. (University Illinois, Urbana, IL) 22–24.

Levy, E.D., De, S., and Teichmann, S. a (2012). Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. Proc. Natl. Acad. Sci. U. S. A. *109*, 20461–20466.

Li, W., Keeble, A.H., Giffard, C., James, R., Moore, G.R., and Kleanthous, C. (2004). Highly Discriminating Protein-Protein Interaction Specificities in the Context of a Conserved Binding Energy Hotspot. J. Mol. Biol. *337*, 743–759.

Li, Z., Hartl, F.U., and Bracher, A. (2013). Structure and function of Hip, an attenuator of the Hsp70 chaperone cycle. Nat. Struct. Mol. Biol. *20*, 929–935.

Libich, D.S., Tugarinov, V., and Clore, G.M. (2015). Intrinsic unfoldase/foldase activity of the chaperonin GroEL directly demonstrated using multinuclear relaxation-based NMR. Proc. Natl. Acad. Sci. U. S. A. *112*, 8817–8823.

Lin, Z., Schwarz, F.P., and Eisenstein, E. (1995). The hydrophobic nature of GroEL-substrate binding. J. Biol. Chem. *270*, 1011–1014.

Lin, Z., Madan, D., and Rye, H.S. (2008). GroEL stimulates protein folding through forced unfolding. Nat. Struct. {&} Mol. Biol. *15*, 303–311.

Lindorff-Larsen, K., Piana, S., Dror, R.O., and Shaw, D.E. (2011). How Fast-Folding Proteins Fold. Science (80-.). *334*, 517–520.

Lopez, T., Dalton, K., and Frydman, J. (2015). The Mechanism and Function of Group II Chaperonins. J. Mol. Biol. *427*, 2919–2930.

Luheshi, L.M., Crowther, D.C., and Dobson, C.M. (2008). Protein misfolding and disease: from the test tube to the organism. Curr. Opin. Chem. Biol. *12*, 25–31.

Maghawry, H. a, Mostafa, M.G.M., and Gharib, T.F. (2014). A New Protein Structure Representation for Efficient Protein Function Prediction. J. Comput. Biol. *21*, 1–11.

Maier, R., Scholz, C., and Schmid, F.X. (2001). Dynamic association of trigger factor with protein substrates. J. Mol. Biol. *314*, 1181–1190.

Makhatadze, G.I. (2005). Thermal Unfolding of Proteins Studied by Calorimetry. In Protein Folding Handbook, Five-Volume Set, pp. 70–98.

Marcinowski, M., Rosam, M., Seitz, C., Elferich, J., Behnke, J., Bello, C., Feige, M.J., Becker, C.F.W., Antes, I., and Buchner, J. (2013). Conformational selection in substrate recognition by Hsp70 chaperones. J. Mol. Biol. *425*, 466–474.

Mashaghi, A., Bezrukavnikov, S., Minde, D.P., Wentink, A.S., Kityk, R., Zachmann-Brand, B., Mayer, M.P., Kramer, G., Bukau, B., and Tans, S.J. (2016). Alternative modes of client binding enable functional plasticity of Hsp70. Nature.

Mattoo, R.U.H., and Goloubinoff, P. (2014). Molecular chaperones are nanomachines that catalytically unfold misfolded and alternatively folded proteins. Cell. Mol. Life Sci. *71*, 3311–3325.

Mayer, M.P. (2013). Hsp70 chaperone dynamics and molecular mechanism. Trends Biochem. Sci. *38*, 507–514.

Mayer, M.P., and Breton, L. Le (2015). Review Hsp90 : Breaking the Symmetry. Mol. Cell *58*, 8–20.

McConkey, E.H. (1982). Molecular evolution, intracellular organization, and the quinary structure of proteins. Proc. Natl. Acad. Sci. U. S. A. *79*, 3236–3240.

Micevski, D., and Dougan, D.A. (2013). Regulated Proteolysis in Microorganisms. In Regulated Proteolysis in Microorganisms, pp. 105–128.

Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., and Kardos, J. (2015). Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc. Natl. Acad. Sci. U. S. A. *112*, E3095-103.

Miller, S.B.M., Mogk, A., and Bukau, B. (2015). Spatially organized aggregation of misfolded proteins as cellular stress defense strategy. J. Mol. Biol. *427*, 1564–1574.

Morawe, T., Hiebel, C., Kern, A., and Behl, C. (2012). Protein homeostasis, aging and Alzheimer's disease. Mol. Neurobiol. *46*, 41–54.

Motojima, F. (2015). How do chaperonins fold protein? Biophysics (Oxf). *11*, 93–102.

Nam, S.-H., and Walsh, M.K. (2003). Characterization of interactions between Escherichia coli molecular chaperones and immobilized caseins. Prep. Biochem. Biotechnol. *33*, 321–339.

Narayanan, A., Pullepu, D., and Kabir, M.A. (2016). The interactome of CCT complex – A computational analysis. Comput. Biol. Chem. *64*, 396–402.

Noeske, J., and Cate, J.H.D. (2012). Structural basis for protein synthesis: Snapshots of the ribosome in motion. Curr. Opin. Struct. Biol. *22*, 743–749.

Onuchic, J.N., Wolynes, P.G., Luthey-Schulten, Z., and Socci, N.D. (1995). Toward an outline of the topography of a realistic protein-folding funnel. Proc. Natl. Acad. Sci. *92*, 3626–3630.

Pace, C.N., Fu, H., Fryar, K.L., Landua, J., Trevino, S.R., Shirley, B.A., Hendricks, M.M., Iimura, S., Gajiwala, K., Scholtz, J.M., et al. (2011). Contribution of hydrophobic interactions to protein stability. J. Mol. Biol. *408*, 514–528.

Park, E.S., Fenton, W.A., and Horwich, A.L. (2005). No evidence for a forced-unfolding mechanism during ATP/GroES binding to substrate-bound GroEL: No observable protection of metastable Rubisco intermediate or GroEL-bound Rubisco from tritium exchange. FEBS Lett. *579*, 1183–1186.

Park, S.J., Borin, B.N., Martinez-Yamout, M. a, and Dyson, H.J. (2011a). The client protein p53 adopts a molten globule-like state in the presence of Hsp90. Nat. Struct. Mol. Biol. *18*, 537–541.

Park, S.J., Kostic, M., and Dyson, H.J. (2011b). Dynamic interaction of Hsp90 with its client protein p53. J. Mol. Biol. *411*, 158–173.

Pashley, C.L., Morgan, G.J., Kalverda, A.P., Thompson, G.S., Kleanthous, C., and Radford, S.E. (2012). Conformational properties of the unfolded state of Im7 in nondenaturing conditions. J. Mol. Biol. *416*, 300–318.

Peisker, K., Chiabudini, M., and Rospert, S. (2010). The ribosome-bound Hsp70 homolog Ssb of Saccharomyces cerevisiae. Biochim. Biophys. Acta - Mol. Cell Res. *1803*, 662–672.

Perrett, S., Zahn, R., Stenberg, G., and Fersht, A.R. (1997). Importance of electrostatic interactions in the rapid binding of polypeptides to GroEL. J. Mol. Biol. *269*, 892–901.

Pincus, D. (2016). Size doesn't matter in the heat shock response. Curr. Genet. 1–4.

Pirchi, M., Ziv, G., Riven, I., Cohen, S.S., Zohar, N., Barak, Y., and Haran, G. (2011). Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein. Nat. Commun. *2*, 493.

Powers, E.T., and Balch, W.E. (2013). Diversity in the origins of proteostasis networks--a driver for protein function in evolution. Nat. Rev. Mol. Cell Biol. *14*, 237–248.

Powers, E.T., Morimoto, R.I., Dillin, A., Kelly, J.W., and Balch, W.E. (2009). Biological and chemical approaches to diseases of proteostasis deficiency. Annu. Rev. Biochem. *78*, 959–991.

Privalov, P.L., and Makhatadze, G.I. (1993). Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. J. Mol. Biol. *232*, 660–679.

Priya, S., Sharma, S.K., Sood, V., Mattoo, R.U.H., Finka, A., Azem, A., De Los Rios, P., and Goloubinoff, P. (2013a). GroEL and CCT are catalytic unfoldases mediating out-of-cage polypeptide refolding without ATP. Proc. Natl. Acad. Sci. U. S. A. *110*, 7199–7204.

Priya, S., Sharma, S.K., and Goloubinoff, P. (2013b). Molecular chaperones as enzymes that catalytically unfold misfolded polypeptides. FEBS Lett. *587*, 1981–1987.

Ptitsyn, O.B. (1995). Molten Globule and Protein Folding. Adv. Protein Chem. *47*, 83–229.

Pugh, S.D., Gell, C., Smith, D.A., Radford, S.E., and Brockwell, D.J. (2010). Single-molecule studies of the Im7 folding landscape. J. Mol. Biol. *398*, 132–145.

Quan, S., and Bardwell, J.C.A. (2012). Chaperone discovery. BioEssays *34*, 973–981.

Quan, S., Koldewey, P., Tapley, T., Kirsch, N., Ruane, K.M., Pfizenmaier, J., Shi, R., Hofmann, S.,

Foit, L., Ren, G., et al. (2011). Genetic selection designed to stabilize proteins uncovers a chaperone called Spy. Nat. Struct. Mol. Biol. *18*, 262–269.

Quan, S., Wang, L., Petrotchenko, E. V, Makepeace, K.A., Horowitz, S., Yang, J., Zhang, Y., Borchers, C.H., and Bardwell, J.C. (2014). Super Spy variants implicate flexibility in chaperone action. Elife *3*, e01584.

Raffa, R.G., and Raivio, T.L. (2002). A third envelope stress signal transduction pathway in Escherichia coli. Mol. Microbiol. *45*, 1599–1611.

Raivio, T.L., Laird, M.W., Joly, J.C., and Silhavy, T.J. (2000). Tethering of CpxP to the inner membrane prevents spheroplast induction of the Cpx envelope stress response. Mol. Microbiol. *37*, 1186–1197.

Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. J. Mol. Biol. *7*, 95–99.

Reissmann, S., Joachimiak, L.A., Chen, B., Meyer, A.S., Nguyen, A., and Frydman, J. (2012). A Gradient of ATP Affinities Generates an Asymmetric Power Stroke Driving the Chaperonin TRIC/CCT Folding Cycle. Cell Rep. *2*, 866–877.

Rodriguez-Mendieta, I.R., Spence, G.R., Gell, C., Radford, S.E., and Smith, D.A. (2005). Ultraviolet resonance Raman studies reveal the environment of tryptophan and tyrosine residues in the native and partially folded states of the E colicin-binding immunity protein Im7. Biochemistry *44*, 3306–3315.

Roh, S.H., Kasembeli, M., Galaz-Montoya, J.G., Trnka, M., Lau, W.C.Y., Burlingame, A., Chiu, W., and Tweardy, D.J. (2016a). Chaperonin TRiC/CCT modulates the folding and activity of leukemogenic fusion oncoprotein AML1-ETO. J. Biol. Chem. *291*, 4732–4741.

Roh, S.H., Kasembeli, M.M., Galaz-Montoya, J.G., Chiu, W., and Tweardy, D.J. (2016b). Chaperonin TRiC/CCT Recognizes Fusion Oncoprotein AML1-ETO through Subunit-Specific Interactions. Biophys. J. *110*, 2377–2385.

Röhl, A., Rohrberg, J., and Buchner, J. (2013). The chaperone Hsp90: Changing partners for demanding clients. Trends Biochem. Sci. *38*, 253–262.

Rose, G.D., Fleming, P.J., Banavar, J.R., and Maritan, A. (2006). A backbone-based theory of protein folding. Proc. Natl. Acad. Sci. U. S. A. *103*, 16623–16633.

Roth, D.M., and Balch, W.E. (2011). Modeling general proteostasis: Proteome balance in health and disease. Curr. Opin. Cell Biol. *23*, 126–134.

Rüdiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. EMBO J. *16*, 1501–1507.

Rüßmann, F., Stemp, M.J., Mönkemeyer, L., Etchells, S.A., Bracher, A., and Hartl, F.U. (2012). Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC / CCT. *109*.

Rutherford, B.J., Dahl, R.H., Price, R.E., Szmidt, H.L., Benke, P.I., Mukhopadhyay, A., and Keasling, J.D. (2010). Functional genomic study of exogenous n-butanol stress in Escherichia coli. Appl. Environ. Microbiol. *76*, 1935–1945.

Saibil, H.R., Fenton, W. a, Clare, D.K., and Horwich, A.L. (2013). Structure and allostery of the chaperonin GroEL. J. Mol. Biol. *425*, 1476–1487.

Saio, T., Guan, X., Rossi, P., Economou, A., and Kalodimos, C.G. (2014). Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. Science *344*, 597–610.

Salmon, L., and Blackledge, M. (2015). Investigating protein conformational energy landscapes and

atomic resolution dynamics from NMR dipolar couplings: a review. Rep. Prog. Phys. *78*, 126601.

Salmon, L., Ahlstrom, L.S., Horowitz, S., Dickson, A., Brooks, C.L., and Bardwell, J.C.A. (2016). Capturing a Dynamic Chaperone-Substrate Interaction Using NMR-Informed Molecular Modeling. J. Am. Chem. Soc. *138*, 9826–9839.

Scalbert, A. (1991). Antimicrobial properties of tannins. Phytochemistry *30*, 3875–3883.

Schechter, Y., Patchornik, a, and Burstein, Y. (1973). Selective reduction of cystine 1-8 in alpha-lactalbumin. Biochemistry *12*, 3407–3413.

Schmidpeter, P.A.M., and Schmid, F.X. (2015). Prolyl isomerization and its catalysis in protein folding and protein function. J. Mol. Biol. *427*, 1609–1631.

Scholl, Z.N., Yang, W., and Marszalek, P.E. (2014). Chaperones rescue luciferase folding by separating its domains. J. Biol. Chem. *289*, 28607–28618.

Schreiber, G., Haran, G., and Zhou, H.-X. (2009). Fundamental aspects of protein-protein association kinetics. Chem. Rev. *109*, 839–860.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys. J. *78*, 1606–1619.

Schumann, W. (2016). Regulation of bacterial heat shock stimulons. Cell Stress Chaperones 1–10.

Schwalm, J., Mahoney, T.F., Soltes, G.R., and Silhavy, T.J. (2013). Role for Skp in LptD assembly in Escherichia coli. J. Bacteriol. *195*, 3734–3742.

Schweizer, R.S., Aponte, R.A., Zimmermann, S., Weber, A., and Reinstein, J. (2011). Fine Tuning of a Biological Machine: DnaK Gains Improved Chaperone Activity by Altered Allosteric Communication and Substrate Binding. ChemBioChem *12*, 1559–1573.

Sciore, A., Su, M., Koldewey, P., Eschweiler, J.D., Diffley, K.A., Linhares, B.M., Ruotolo, B.T., Bardwell, J.C.A., Skiniotis, G., and Marsh, E.N.G. (2016). Flexible, symmetry-directed approach to assembling protein cages. Proc. Natl. Acad. Sci. *113*, 8681–8686.

Sekhar, A., and Kay, L.E. (2013). NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. Proc. Natl. Acad. Sci. U. S. A. *110*, 12867–12874.

Sharma, U.K., and Chatterji, D. (2010). Transcriptional switching in Escherichia coli during stress and starvation by modulation of σ70 activity. FEMS Microbiol. Rev. *34*, 646–657.

Sharma, S., Chakraborty, K., Müller, B.K., Astola, N., Tang, Y.C., Lamb, D.C., Hayer-Hartl, M., and Hartl, F.U. (2008). Monitoring Protein Conformation along the Pathway of Chaperonin-Assisted Folding. Cell *133*, 142–153.

Sharma, S.K., De los Rios, P., Christen, P., Lustig, A., and Goloubinoff, P. (2010). The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. Nat. Chem. Biol. *6*, 914–920.

Shoemaker, B.A., Portman, J.J., and Wolynes, P.G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc. Natl. Acad. Sci. U. S. A. *97*, 8868–8873.

Shtilerman, M., Lorimer, G.H., and Englander, S.W. (1999). Chaperonin Function: Folding by Forced Unfolding. Science (80-.). *284*, 822–825.

Sirur, A., and Best, R.B. (2013). Effects of interactions with the groel cavity on protein folding rates. Biophys. J. *104*, 1098–1106.

Song, M.C., and Scheraga, H.A. (2000). Formation of native structure by intermolecular thiol-disulfide exchange reactions without oxidants in the folding of bovine pancreatic ribonuclease A. FEBS Lett. *471*, 177–181.

Southall, N.T., Dill, K.A., and Haymet, A.D.J. (2002). A view of the hydrophobic effect. J. Phys. Chem. B *106*, 521–533.

Spitzer, J., and Poolman, B. (2009). The role of biomacromolecular crowding, ionic strength, and physicochemical gradients in the complexities of life's emergence. Microbiol. Mol. Biol. Rev. *73*, 371–388.

Srivastava, S.K., Lambadi, P.R., Ghosh, T., Pathania, R., and Navani, N.K. (2014). Genetic regulation of spy gene expression in Escherichia coli in the presence of protein unfolding agent ethanol. Gene *548*, 142–148.

Strop, P., and Mayo, S.L. (2000). Contribution of surface salt bridges to protein stability. Biochemistry *39*, 1251–1255.

Stull, F., Koldewey, P., Humes, J.R., Radford, S.E., and Bardwell, J.C.A. (2016). Substrate protein folds while it is bound to the ATP-independent chaperone Spy. Nat. Struct. Mol. Biol. *23*, 53–58.

Sun, P., Tropea, J.E., and Waugh, D.S. (2011). Heterologous Gene Expression.

Sutto, L., Lätzer, J., Hegler, J. a, Ferreiro, D.U., and Wolynes, P.G. (2007). Consequences of localized frustration for the folding mechanism of the IM7 protein. Proc. Natl. Acad. Sci. U. S. A. *104*, 19825–19830.

Svanström, A., and Grantham, J. (2016). The molecular chaperone CCT modulates the activity of the actin filament severing and capping protein gelsolin in vitro. Cell Stress Chaperones *21*, 55–62.

Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K.D., Karras, G.I., and Lindquist, S. (2012). Quantitative analysis of Hsp90-client interactions reveals principles of substrate recognition. Cell *150*, 987–1001.

Takei, Y. (2015). From aquatic to terrestrial life: evolution of the mechanisms for water acquisition. Zoolog. Sci. *32*, 1–7.

Taldone, T., Patel, H.J., Bolaender, A., Patel, M.R., and Chiosis, G. (2014). Protein chaperones: a composition of matter review (2008 – 2013). Expert Opin Ther Pat. *24*, 501–518.

Tam, S., Spiess, C., Auyeung, W., Joachimiak, L., Chen, B., Poirier, M. a, and Frydman, J. (2009). The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation. Nat. Struct. Mol. Biol. *16*, 1279–1285.

Tandford, C. (1979). The hydrophobic effect and the organization of living matter. Science (80-.). *200*, 1012–1018.

Tapley, T.L., Franzmann, T.M., Chakraborty, S., Jakob, U., and Bardwell, J.C. a (2010). Protein refolding by pH-triggered chaperone binding and release. Proc. Natl. Acad. Sci. U. S. A. *107*, 1071–1076.

Teixeira, F., Castro, H., Cruz, T., Tse, E., Koldewey, P., Southworth, D.R., Tomás, A.M., and Jakob, U. (2015). Mitochondrial peroxiredoxin functions as crucial chaperone reservoir in Leishmania infantum. Proc. Natl. Acad. Sci. U. S. A. 1419682112-.

Terwilliger, T.C., Bunkóczi, G., Hung, L., Zwart, P.H., Smith, J.L., Akey, D.L., and Adams, P.D. (2016). Can I solve my structure by SAD phasing? Planning an experiment, scaling data and evaluating the useful anomalous correlation and anomalous signal. Acta Crystallogr. Sect. D Struct. Biol. *72*, 359–374.

Tian, G., Vainberg, I.E., Tap, W.D., Lewis, S.A., and Cowan, N.J. (1995). Specificity in chaperonin-mediated protein folding. Nature *375*, 250–253.

Tompa, P., and Tompa, P. (2002). Intrinsically unstructured proteins. Trends Biochem. Sci. *27*, 527–533.

Treweek, T.M., Meehan, S., Ecroyd, H., and Carver, J.A. (2015). Small heat-shock proteins: Important players in regulating cellular proteostasis. Cell. Mol. Life Sci. *72*, 429–451.

Truscott, K.N., Bezawork-Geleta, A., and Dougan, D.A. (2011). Unfolded protein responses in bacteria and mitochondria: A central role for the ClpXP machine. IUBMB Life *63*, 955–963.

Tyagi, N.K., Fenton, W.A., Deniz, A.A., and Horwich, A.L. (2011). Double mutant MBP refolds at same rate in free solution as inside the GroEL/GroES chaperonin chamber when aggregation in free solution is prevented. FEBS Lett. *585*, 1969–1972.

Udgaonkar, J.B. (2008). Multiple routes and structural heterogeneity in protein folding. Annu. Rev. Biophys. *37*, 489–510.

Uversky, V.N. (2016). Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. J. Biol. Chem. *291*, 6681–6688.

Uversky, V.N., Oldfield, C.J., and Dunker, a K. (2008). Intrinsically Disordered Proteins in Human Diseases: Introducing the D 2 Concept. Annu. Rev. Biophys. *37*, 215–246.

Verba, K.A., Wang, R.Y.-R., Arakawa, A., Liu, Y., Shirouzu, M., Yokoyama, S., and Agard, D.A. (2016). Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. Science (80-.). *352*, 1542–1547.

Villebeck, L., Moparthi, S.B., Lindgren, M., Hammarström, P., and Jonsson, B.H. (2007). Domain-specific chaperone-induced expansion is required for β-actin folding: A comparison of β-actin conformations upon interactions with GroEL and tail-less complex polypeptide 1 ring complex (TRiC). Biochemistry *46*, 12639–12647.

Voelz, V.A., Singh, V.R., Wedemeyer, W.J., Lapidus, L.J., and Pande, V.S. (2010). Unfolded-State Dynamics and Structure of Protein L Characterized by Simulation and Experiment. J Am Chem Soc *132*, 4702–4709.

Vogt, S.L., and Raivio, T.L. (2012). Just scratching the surface: An expanding view of the Cpx envelope stress response. FEMS Microbiol. Lett. *326*, 2–11.

Voth, W., Schick, M., Gates, S., Li, S., Vilardi, F., Gostimskaya, I., Southworth, D.R., Schwappach, B., and Jakob, U. (2014). The protein targeting factor Get3 functions as ATP-Independent chaperone under oxidative stress conditions. Mol. Cell *56*, 116–127.

Walter, P., and Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. Science (80-.). *334*, 1081–1086.

Walters, B.T., Mayne, L., Hinshaw, J.R., Sosnick, T.R., and Englander, S.W. (2013). Folding of a large protein at high structural resolution. Proc. Natl. Acad. Sci. U. S. A. *110*, 18898–18903.

Walton, T. a, Sandoval, C.M., Fowler, C.A., Pardi, A., and Sousa, M.C. (2009). The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. Proc. Natl. Acad. Sci. U. S. A. *106*, 1772–1777.

Wang, D., and Fierke, C. a (2013). The BaeSR regulon is involved in defense against zinc toxicity in E. coli. Metallomics *5*, 372–383.

Wang, F., Cazzolli, G., Wintrode, P., and Faccioli, P. (2016). Folding Mechanism of Proteins Im7 and Im9: Insight from All-Atom Simulations in Implicit and Explicit Solvent. J. Phys. Chem. B *7*,

acs.jpcb.6b05819.

Wang, J.D., Herman, C., Tipton, K.A., Gross, C.A., and Weissman, J.S. (2002). Directed evolution of substrate-optimized GroEL/S chaperonins. Cell *111*, 1027–1039.

Wang, X., Song, X., Zhuo, W., Fu, Y., Shi, H., Liang, Y., Tong, M., Chang, G., and Luo, Y. (2009). The regulatory mechanism of Hsp90alpha secretion and its function in tumor malignancy. Proc. Natl. Acad. Sci. U. S. A. *106*, 21288–21293.

Weaver, J., and Rye, H.S. (2014). The C-terminal tails of the bacterial chaperonin GroEL stimulate protein folding by directly altering the conformation of a substrate protein. J. Biol. Chem. *289*, 23219–23232.

Wedemeyer, W.J., Welker, E., and Scheraga, H.A. (2002). Proline cis-trans isomerization and protein folding. Biochemistry *41*, 14637–14644.

Whitesell, L., and Lindquist, S.L. (2005). HSP90 and the chaperoning of cancer. Nat. Rev. Cancer *5*, 761–772.

Whitmore, L., and Wallace, B.A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. Biopolymers *89*, 392–400.

Whittaker, S.B.M., Clayden, N.J., and Moore, G.R. (2011). NMR characterisation of the relationship between frustration and the excited state of Im7. J. Mol. Biol. *414*, 511–529.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. *67*, 235–242.

Winter, J., and Jakob, U. (2004a). Beyond transcription--new mechanisms for the regulation of molecular chaperones. Crit. Rev. Biochem. Mol. Biol. *39*, 297–317.

Winter, J., and Jakob, U. (2004b). Beyond transcription--new mechanisms for the regulation of molecular chaperones. Crit. Rev. Biochem. Mol. Biol. *39*, 297–317.

Wooh, J.W., Kidd, R.D., Martin, J.L., and Kobe, B. (2003). Comparison of three commercial sparse-matrix crystallization screens. Acta Crystallogr. - Sect. D Biol. Crystallogr. *59*, 769–772.

Xia, K., Zhang, S., Bathrick, B., Liu, S., Garcia, Y., and Col??n, W. (2012). Quantifying the kinetic stability of hyperstable proteins via time-dependent SDS trapping. Biochemistry *51*, 100–107.

Xu, W., and Neckers, L. (2007). Targeting the molecular chaperone heat shock protein 90 provides a multifaceted effect on diverse cell signaling pathways of cancer cells. Clin. Cancer Res. *13*, 1625–1629.

Xue, B., Dunker, A.K., and Uversky, V.N. (2012). Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. J. Biomol. Struct. Dyn. *30*, 137–149.

Yam, A.Y., Xia, Y., Lin, H.-T.J., Burlingame, A., Gerstein, M., and Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. Nat. Struct. Mol. Biol. *15*, 1255–1262.

Yamamoto, K., Ogasawara, H., and Ishihama, A. (2008). Involvement of multiple transcription factors for metal-induced spy gene expression in Escherichia coli. J. Biotechnol. *133*, 196–200.

Yesselman, J.D., Horowitz, S., Brooks, C.L., and Trievel, R.C. (2015). Frequent side chain methyl carbon-oxygen hydrogen bonding in proteins revealed by computational and stereochemical analysis of neutron structures. Proteins Struct. Funct. Bioinforma. *83*, 403–410.

Zarrine-Afsar, A., and Davidson, A.R. (2004). The analysis of protein folding kinetic data produced in protein engineering experiments. Methods *34*, 41–50.

Zhang, X., and Kelly, J.W. (2014). Chaperonins resculpt folding free energy landscapes to avoid kinetic traps and accelerate protein folding. J. Mol. Biol. *426*, 2736–2738.

Zhang, H., Davies, K.J.A., and Forman, H.J. (2015). Oxidative stress response and Nrf2 signaling in aging. Free Radic. Biol. Med. *88*, 1–23.

Zheng, X., Bi, C., Li, Z., Podariu, M., and Hage, D.S. (2015). Analytical methods for kinetic studies of biological interactions: A review. J. Pharm. Biomed. Anal. *113*, 163–180.

Zhuravleva, A., and Gierasch, L.M. (2015). Substrate-binding domain conformational dynamics mediate Hsp70 allostery. Proc. Natl. Acad. Sci. U. S. A. *112*, E2865-73.

Zimmerman, S.B., and Trach, S.O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. J. Mol. Biol. *222*, 599–620.

Ziv, G., and Haran, G. (2009). Protein folding, protein collapse, and Tanford's transfer model: Lessons from single-molecule FRET. J. Am. Chem. Soc. *131*, 2942–2947.

Zoetendal, E.G., Smith, A.H., Sundset, M. a, and Mackie, R.I. (2008). The BaeSR two-component regulatory system mediates resistance to condensed tannins in Escherichia coli. Appl. Environ. Microbiol. *74*, 535–539.

6 Abbreviations

ATP	Adenosine tri-phosphate
ADP	Adenosine di-phosphate
Bae	Envelope stress response pathway of *E. coli*
BLI	Bio-layer interferometry
CCT/TRiC	Chaperonin Containing TCP-1 complex/ TCP-1 Ring Complex; chaperonin system of eukaryotes and some archaea
CD	Circular dichroism spectroscopy
Cdc37	Hsp90 co-chaperone
CDK4	Cyclin-dependent kinase 4; Hsp90 client protein
DegP	Periplasmic protease of *E. coli*
Cpx	Envelope stress response pathway of *E. coli*
CpxP	Periplasmic regulator of the Cpx envelope stress response pathway of *E. coli* and a paralog of the protein Spy
CsgA	Major curlin subunit; structural subunit of the curli fimbriae of *E. coli*
CTD	C-terminal domain of Hsp90
DapA	Dihydrodipicolinate synthase
ΔC_p	Thermodynamic constant depicting the change in heat capacity due to a reaction
Get3	Golgi to ER traffic protein 3 of yeast; redox-regulated chaperone
ΔG	Gibb's free energy
ΔH	Thermodynamic constant depicting the change in enthalpy due to a reaction
DNase	Desoxyribonuclease
ΔS	Thermodynamic constant depicting the change in entropy due to a reaction
E. coli	Escherichia coli; gram negative bacterium
GroEL/ES	Chaperonin system of *E. coli*
HdeA	Acid induced periplasmic molecular chaperone of *E. coli*
Hsp	Heat shock protein
IDP	Intrinsically disordered protein
Im7	Colicin E7 immunity protein
ITC	Isothermal Titration Calorimetry
K	Equilibrium folding constant

Abbreviations

K_d	Dissociation constant
kDa	kilo Dalton
k_{IN}	Folding rate constant for the conformational transition from the intermediate to the native folding state
K_{IN}	Folding equilibrium constant for the conformational transition from the intermediate to the native folding state
k_{obs}	Observed rate constant
k_{off}	Release rate constant
k_{on}	Binding rate constant
k_{UI}	Folding rate constant for the conformational transition from the unfolded to the intermediate folding state
K_{UI}	Folding equilibrium constant for the conformational transition from the unfolded to the intermediate folding state
MBP	Maltose binding protein; periplasmic protein of *E. coli*
MD	Middle domain of Hsp90
mRNA	Messenger ribonucleic acid
NEF	Nucleotide exchange factor of Hsp70
NMR	Nuclear magnetic resonance spectroscopy
NTD	N-terminal domain of Hsp90
p53	Eukaryotic transcriptional regulator; client protein of Hsp90
PAGE	Polyacrylamide gel electrophoresis
pI-Phe	4-Iodo-phenylalanin
PN	Proteostasis network
RNase A/H	Ribonuclease A; endonuclease that catalyzes the cleavage of RNA
Rcs	Envelope stress response pathway of *E. coli*
r.m.s.d	Root-mean-square deviation
rpm	Rounds per minute
RpoH	RNA polymerase sigma factor; activates heat shock gene expression
SBD	Substrate binding domain of Hsp70
SDS	Sodium dodecyl sulfate
SecA/B	Components of the *E. coli* secretion pathway SEC; A: Tanslocase subunit; B: chaperone that transfers unfolded proteins to SecA for trans-membrane transport

SeMet	Seleno-L-methionine
Se-SAD	Single-wavelength anomalous diffraction; X-ray crystallography technique used to obtain phases via anomalous signal of selenium (Se)
SH3	SRC Homology 3 domain
Skp	Periplasmic molecular chaperone of *E. coli*
Spy	Spheroblast protein Y; molecular chaperone
Ssz1	Hsp70 homolog; subunit of the ribosome-association complex of *saccharomyces cerevisiae*
SUMO	Small ubiquitin-like modifier
SurA	Periplasmic proline-isomerase and molecular chaperone of *E. coli*
SV	Sedimentation velocity analytical ultracentrifugation
TS	Transition state of a reaction
ULP1	SUMO-specific protease
VDW	van der Waals interactions, induced by transient dipoles due to fluctuations in the electron density interacting groups
Zrap	Molecular chaperone and paralog of the *E. coli* protein Spy

7 Acknowledgements

First, I would like to thank Prof. Dr. James C.A. Bardwell, for the opportunity to join his lab. Without his steady support, encouragement, and guidance me throughout my PhD, I wouldn't have come that far. It was a very good experience for me to work in his lab. Thank you Jim!

I also want to thank Prof. Dr. Johannes Buchner, who kindly accepted to be my supervisor at the TU Munich. Without him this thesis wouldn't have been possible.

I would also like to express my thanks to Prof. Dr. Ursula Jakob for her support. Without her very helpful advice, especially during lab meetings, my work wouldn't have developed to the level that it has.

I would like to thank Prof. Dr. Bruce Palfey and Prof. Dr. David Ballou for the critical discussions that helped me foster my knowledge in the field of transient kinetics and stopped-flow.

I also want to thank Prof. Dr. Sheena Radford and Julia Humes for the great collaboration and advice during my project.

I would like to thank Prof. Dr. Erik Zuiderweg for his very critical comments, which in the end helped improve my first manuscript.

I also very much want to thank Dr. Scott Horowitz and Dr. Frederick Stull. Without your support, the inspiring conversations and our great collaborative work, this thesis wouldn't have reached the level that it has. I would like to especially thank Dr. Scott Horowitz for introducing me into the world of X-ray crystallography and the many fun trips to the synchrotron. I will miss working with you guys!

I would also like to thank Dr. Titus Franzmann and Prof. Dr. Borries Demeler for the introduction and support in Analytical Ultracentrifugation. This technique really helped me to progress with my thesis and to establish many fruitful collaborations.

Many thanks also to Dr. Linda Foit, Dr. Shu Quan, and Dr. Christopher Lennon for inspiring conversations and for all the help with scientific issues. Thanks guys!

Special thanks also to Ke Wan for the numerous protein preps. You were a big time saver!

I would like to thank Dr. James Windack for helping me with the ESI-MS.

Thanks also goes to all my many collaborators for all the great work we did together.

Of course I also thank the whole Bardwell and Jakob lab for the great working atmosphere, the help with the daily scientific issues and the fun time in and outside the lab. Especially Tsinatkeab Tadesse,

Antje Müller-Schickert, and Wilhelm Voth I would like to thank for the many good conversations and fun moments that kept me going!

I have to thank our lab managers Frederic Amenta, Julia Hrycko, and Molly Wysong and the work study team for the steady support. Thank you guys! You made the daily work in the lab a whole lot easier.

I have to express my sincerest thanks to Boehringer Ingelheim Fonds for the financial support and invitations to very helpful meetings and workshops. I am very grateful for that!

I want to thank Filipa Teixeira for the love and support. Without you, I wouldn't have made it! Thank you!

Finally, I want to declare my deepest thanks to my family for the support and the belief in me through all the years of my academic education. Your steady love and trust encouraged me to go on with my studies. Thank you so much!

8 List of publications

* Authors equally contributed to the publication.

1. Dahl JU, **Koldewey P**, Bardwell JC, Jakob U. *Detection of the pH-dependent Activity of Escherichia coli Chaperone HdeB In Vitro and In Vivo*. **J Vis Exp**. 2016 Oct 23;(116). doi: 10.3791/54527

2. Horowitz S, Koepnick B, Martin R, Tymieniecki A, Winburn AA, Cooper S, Flatten J, Rogawski DS, Koropatkin NM, Hailu TT, Jain N, **Koldewey P**, Ahlstrom LS, Chapman MR, Sikkema AP, Skiba MA, Maloney FP, Beinlich FR; Foldit Players; University of Michigan students, Popović Z, Baker D, Khatib F, Bardwell JC. *Determining crystal structures through crowdsourcing and coursework.* **Nat Commun.** 2016 Sep 16;7:12549. doi: 10.1038/ncomms12549.

3. Aaron Sciore A, Min Su M, **Koldewey P**, Eschweiler JD, Diffley KA, Linhares BM, Ruotolo BT, Bardwell JCA, Skiniotis G, Marsh ENG. *A flexible, symmetry-directed approach to assembling protein cages.* **Proc Natl Acad Sci U S A.** 2016 Aug 2;113(31):8681-6. doi: 10.1073/pnas.1606013113.

4. Horowitz S*, Salmon L*, **Koldewey P***, Ahlstrom LS, Martin R, Quan S, Afonine PV, van den Bedem H, Wang L, Xu Q, Trievel RC, Brooks CL 3rd, Bardwell JC. *Visualizing chaperone-assisted protein folding.* **Nat Struct Mol Biol.** 2016 Jul;23(7):691-7. doi: 10.1038/nsmb.3237.

5. **Koldewey P**, Stull F, Horowitz S, Martin R, Bardwell JCA. *Forces Driving Chaperone Action.* **Cell.** 2016 Jul 14;166(2):369-79. doi: 10.1016/j.cell.2016.05.054.

6. Stull F*, **Koldewey P***, Humes JR, Radford SE, Bardwell JC. *Substrate protein folds while it is bound to the ATP-independent chaperone Spy.* **Nat Struct Mol Biol.** 2016 Jan;23(1):53-8. doi: 10.1038/nsmb.3133.

7. Teixeira F, Castro H, Cruz T, Tse E, **Koldewey P**, Southworth DR, Tomás AM, Jakob U. *Mitochondrial peroxiredoxin functions as crucial chaperone reservoir in Leishmania infantum.* **Proc Natl Acad Sci U S A.** 2015 Feb 17;112(7):E616-24. doi: 10.1073/pnas.1419682112.

8. Dahl JU, **Koldewey P**, Salmon L, Horowitz S, Bardwell JC, Jakob U. *HdeB functions as an acid-protective chaperone in bacteria.* **J Biol Chem.** 2015 Jan 2;290(1):65-75. doi: 10.1074/jbc.M114.612986.

9. Horowitz S, **Koldewey P**, Bardwell JC. *Undergraduates improve upon published crystal structure in class assignment.* **Biochem Mol Biol Educ.** 2014 Sep-Oct;42(5):398-404. doi: 10.1002/bmb.20811.

10. Quan S, **Koldewey P**, Tapley T, Kirsch N, Ruane KM, Pfizenmaier J, Shi R, Hofmann S, Foit L, Ren G, Jakob U, Xu Z, Cygler M, Bardwell JC. *Genetic selection designed to stabilize proteins uncovers a chaperone called Spy.* **Nat Struct Mol Biol.** 2011 Mar;18(3):262-9. doi: 10.1038/nsmb.2016.

9 Appendix

Note that all the papers listed above in section 8, including those relevant for this thesis, have been published and are available online.