Charles University

Faculty of Science

Study programme: Genetics, molecular biology and virology

Branch of study: Molecular biology and genetics of prokaryota and eukaryota

Bc. Aleksandr Melikov

Investigation of HSP70 oligomerization by structural mass spectrometry

Sledování oligomerizace proteinu HSP70 strukturní hmotnostní spektrometrií

Diploma thesis

Supervisor: RNDR. Petr Novák, Ph.D.

Prague, 2024

Declaration

I declare that I wrote and processed the thesis independently and that I have cited all the information sources and literature used. I also declare that this work has not been submitted elsewhere in order to obtain another or the same academic title.

In Prague, 2024 Bc. Aleksandr Melikov

Acknowledgements

I would like to thank my supervisor Dr. Petr Novák for the extensive help and consultations during the work on this diploma thesis. In addition, I would like to thank all the members of Petr Novák's research team and Dr. Zdeněk Kukačka in particular, for their support, teaching and advice regarding experimental work and data analysis. I would also like to express my gratitude to my external supervisor, Prof. Matthias Mayer, as well as Yang Zheng and Heidelberg University, for their collaboration on this project with our institution and providing us with the experimental material.

Further, I would also like to acknowledge our funding sources such as the Czech Science Foundation (22-27695S) and NPO-NEURO-EXCELLES (LX22NPO5107).

Finally, the greatest thanks to my family and my girlfriend, for the constant support during the whole Master studies.

Abstract

Heat shock cognate protein 70 (HSC70) is a 71 kDa chaperone protein belonging to the ubiquitous family of heat shock proteins 70 (Hsp70). The representatives of this protein family are considered as molecular machines with ATP-hydrolase activity facilitating correct folding of spatial protein structure, both in normal and stressful conditions (hypoxia, heat shock, pH fluctuations etc.) In addition, HSC70 was identified as an uncoating enzyme for triskelion meshwork on the surface of clathrin-coated vesicles. Among other roles, HSC70 prevents protein aggregation and assists the polypeptide maturation, it facilitates the protein transport into organelles, such as endoplasmic reticulum and mitochondria. It is involved in targeting proteins for lysosomal degradation and in many other dramatically important cellular processes related to protein homeostasis. Therefore, the regulation of HSC70 and other HSP70 proteins is believed to be dramatically important, especially in a context of cellular stress. Based on the experimental observation, the mechanism of inactivation through oligomerization was hypothesized. The dimer and trimer species of Hsp70 proteins were identified both in case of prokaryotic and eukaryotic homologs. It was also speculated that Hsp40 cofactors promote oligomerization to even higher-order oligomers. This and other possible oligomerization models of wild type HSC70 and the certain subset of HSC70 mutants were investigated by cross-linking mass spectrometry. The distance constraints between certain amino acid residues imposed by the cross-linker length allowed us to build structural models of Hsc70 oligomeric species. To decipher inter/intra molecular restraints and allow the precise mapping of identified cross-links, the studied proteins were produced and analyzed in a mixture containing ^{14}N- and ^{15}N-labeled form.

Keywords: Structural biology, Mass spectrometry, Protein structure, Protein-protein interaction, Allostery, Chaperones, Heat shock proteins, HSC70

Abstrakt

Příbuzný protein teplotního šoku (HSC70) je 71 kDa chaperon patřící ke všudypřítomné rodině proteinů teplotního šoku 70 (Hsp70). Zástupci této proteinové rodiny jsou považováni za molekulární stroje s ATPázovou aktivitou, které usnadňují správné sbalování prostorové struktury proteinů jak za normálních, tak i za stresových podmínek (hypoxie, tepelný šok, či kolísání pH). HSC70 navíc funguje jako enzym rozbalující vrstvu triskelionu na povrchu klatrinových váčků. Další rolí HSC70 je např. zamezení agregace proteinů, asistence při maturaci polypeptidového řetězce a usnadnění transportu proteinů do organel, jako jsou endoplazmatické retikulum a mitochondrie. HSC70 se účastní směrování proteinů určených k degradaci do lysosomů a v mnoha dalších kriticky důležitých buněčných procesech spojených s homeostází proteinů. Tudíž, regulace HSC70 a jiných proteinů HSP70 je považována za velmi důležitou, obzvlášť v kontextu buněčného stresu. Na základě experimentálních pozorování byl navržen mechanismus inaktivace za pomoci oligomerizace. Dimery a trimery proteinů Hsp70 byly identifikovány jak u prokaryotických, tak u eukaryotických homologů. Byla také diskutována role proteinového kofaktoru Hsp40 ve stimulaci oligomerizaci Hsp70 do vyšších oligomerů. Tento a jiné možné modely oligomerizace „divoké“ varianty HSC70 a určité sady mutantů HSC70 byly studovány pomocí hmotnostní spektrometrie s použitím síťovacích činidel. Díky omezování vzdáleností mezi určitými aminokyselinovými zbytky, které je určeno délkou zvoleného síťovadla, bylo možné vytvořit strukturní modely oligomerů HSC70. Abychom zlepšili prostorové rozlišení síťovacího experimentu a dokázali přesně namapovat identifikované zesítění, byly studované proteiny připraveny ve formách ^{14}N a ^{15}N.

Klíčová slova: Strukturní biologie, hmotnostní spektrometrie, Struktura proteinů, Interakce proteinů, Alosterie, Chaperony, Proteiny teplotního šoku, HSC70

Content

List of abbreviations

aa — amino acid
ATP — adenosine triphosphate
ADP — adenosine diphosphate
CBD — C-terminal binding domain
CID — collision induced dissociation
cryoEM — cryo electron microscopy
DSS — disuccinimidyl suberate
DSPU — disuccinimidyl dipropionic urea
ECD — electron capture dissociation
ER — endoplasmic reticulum
ESI — electrospray ionization
FT-ICR — Fourier-transform ion cyclotron resonance
HDX — hydrogen-deuterium exchange
HSC70 — heat shock cognate protein 70
Hsp70/90/60/100 — heat shock protein 70/90/60/100
HPLC —high-performance liquid chromatography
IMAC — immobilized metal affinity chromatography
IM — ion mobility
IMMS — ion mobility mass spectrometry
JDP — J-domain protein
LC — liquid chromatography
MALDI — matrix-assisted laser desorption ionization
MS — mass spectrometry
NBD — nucleotide-binding domain
NEF — nucleotide-exchange factor
NMR — nuclear magnetic resonance
PDB — protein databank
PTM — post-translational modification
RMSD — root mean square deviation
SBD — substrate-binding domain
SDS-PAGE — sodium dodecyl sulfate polyacrylamide gel electrophoresis
sHSPs — small heat shock proteins
TF — trigger factor
tims — trapped ion mobility spectrometry
TOF — time-of-flight mass analyzer
WT — wild type
XL-MS — cross-linking mass spectrometry

Introduction

The heat shock protein 70 (Hsp70) family is found in both prokaryotic and eukaryotic organisms, including *Drosophila* and *Homo sapiens* (*Dwornczak and Mirault 1987*). It consists of many homologous proteins, molecular chaperones, which share a similar set of functions - acting as a protection and support of the correct protein folding during the cell cycle and helping to prevent or alleviate the consequences of various cellular stresses, such as hypoxia, heat shock or pathological states (*Balchin et al. 2016; Mogk et al. 2018*). The Hsp70 family isn't the only one belonging to chaperones – there are many other families different in mechanism of action, but fulfilling the same general function: assist protein folding and prevent misfolding events *(Ellis and van der Vies, 1991).*

During various cellular stresses, the expression of certain chaperone genes (mostly Hsp70 family) is highly upregulated to secure a cell from the protein unfolding danger (*Lindquist 1986*). This natural reaction of cell physiology named heat shock response leads to high concentrations of translated chaperones in cytosol. After the stress is no longer acting, the abundant pool of Hsp70 must be eliminated or at least inactivated due to the substantial negative effect of Hsps on the cell growth, which can be explained by the formation of Hsp70-granules (*Feder et al. 1992*). Since the proteolytic machinery can't degrade the inactive pool of Hsp70-granules in a sufficiently short period of time, the rapid and reversible mechanism of inactivation has to be engaged. The first step of reversible inactivation is believed to be Hsp70 oligomerization, which might be then followed by the certain PTMs of Hsp70 molecules (*Preissler et al. 2015*). The exploration of this first step and elucidation of the formed oligomeric structures of Hsp70 is critical to understanding of the mechanism of reversible Hsp70 inactivation.

Mass spectrometry belongs among the modern structural biology methods. In the past decades, this technique became critically important in studying both the whole proteomes and individual proteins, e.g., conformational changes of a protein, protein-protein or protein-ligand interaction in noncovalent complexes (*Sharon 2010; Dyachenko et al. 2013*). One of such techniques implements the covalent fixation of protein-protein complexes using cross-linkers – organic compounds specifically reacting with certain amino acid residues in proteins and linking them together (*Sinz 2006*). In the case of studying homo oligomers by MS, it is necessary to exploit a labeling technique to distinguish intra-subunit and inter-subunit cross-linked sites, because the amino acid sequence of subunits in a complex is identical. One of the most reliable techniques is the protein isotope labeling, which takes an advantage of prototrophic bacterial

strains for heterologous protein expression, e.g. *E.coli*; the production of a protein containing heavy isotopes of widespread biogenic elements, such as ^{15}N or ^{13}C, is based on the bacterial cultivation on a minimal medium containing inorganic substrates as a source of these isotopes (*Marley et al. 2001; Morgner et al. 2015*).

We have successfully investigated and described the oligomerization of human HSC70, the target Hsp70 family member, and subsequently created the different HSC70 dimer models *in situ,* which was the aim of the work. In the following chapters I would like to give a theoretical insight into chaperones and the Hsp70 network, introduce various MS techniques (and XL-MS in particular) as a very promising tool of structural biology and discuss the experimental part.

Chaperone network: the concept and representatives

Proteins represent highly complex biological polymers made up of amino acids connected via amide, or peptide, bonds, which are formed by virtue of ribosomal catalysis during the process of translation. The biological function of a protein can be fulfilled only in the case if its primary structure, i.e., polypeptide chain adopts the native spatial configuration by reaching one of the thermodynamic energy minimums; nevertheless, for a given protein chain there are always multiple ways to fold into native state (*Dobson et al. 1998*). Several scenarios of this process have been devised in recent decades to attempt to explain the protein folding process, such as the framework model and the hydrophobic collapse model. In addition, a number of energy landscape models have been proposed to depict the process of folding based on the biophysical observations and calculations (*Dill et al. 1995*). The most widely accepted landscape model, the "rugged" landscape, takes into account the existence of multiple energy minima, as well as intermediate folded states adopting those minima. This type of spatial energy diagram is more plausible due to the existence of a huge number of possible short- and long-distance interactions in a protein, which cannot be all satisfied at once, therefore creating a palette of folding intermediates and partially folded protein species (*Bryngelson et al. 1995; Ferreiro et al. 2007*) . While folding intermediates are localized on declines of the energy funnel eventually leading to the natively folded states, many of the partially folded protein entities may lead to misfolded states. The "competing" weak interactions described above can create these local energy minima with high energy barriers; consequently, non-native protein configurations can be trapped in the local energy minima without reaching a biologically correct spatial structure. These partially folded protein species are often very prone to aggregation forming amorphous

aggregates, inactive oligomers or, in the extreme cases, amyloid fibrils (*Jahn and Radford 2005*). All of these protein aggregates are extremely toxic and dangerous to cells and are involved in their aging and death. In this context, the critical role in preventing these aggregation events and facilitating the correct folding of proteins belongs to molecular chaperones.

The term "molecular chaperones" can be defined as "a protein that binds to and stabilizes an otherwise unstable conformer of another protein and by controlled binding and release of the substrate protein, facilitates its correct fate *in vivo*: be it folding, oligomeric assembly, transport to a particular subcellular compartment, or controlled switching between active/inactive conformations" (*Hendrick and Hartl 1993*). According to this definition, a chaperone must bind to structural motifs of an unfolded protein or a monomer of oligomeric assembly, which are not present in their final native structure. These motifs are exposed to the environment and therefore can be bound by chaperones.. The increased proportion of partially folded or completely unfolded proteins is often observed in stressful conditions, such as heat shock, presence of certain compounds etc. As a consequence, an enhanced expression of certain chaperones in hyperthermal conditions was detected; therefore, those heat-induced chaperons acquired the name "heat shock proteins", or Hsps (*Nover and Scharf 1984*). Eventually, this name has stuck to all molecular chaperones present in a cell genome. To classify various types of chaperones, the values of their molecular weight in kilodaltons (kDa) were taken into account, so classes like Hsp70, Hsp60, Hsp90 emerged. It was then discovered that some of these proteins are present in other cell compartments as well, like endoplasmic reticulum (ER) (*Munro and Pelham 1986*). In the mentioned study, the authors initially proposed the possible mechanism of Hsps function: they may act via binding hydrophobic regions exposed to solvent, consequently releasing them and allowing them to refold in a native manner. In **Fig.1,** this process of "refolding assistance" is depicted in the context of the protein folding energy landscape discussed earlier. Before a partially folded protein falls into an energy trap and participates in an amorphous aggregate formation, molecular chaperones bind it and provide enough time for refolding; in the same time, while bound to a chaperone, a protein can't aggregate.

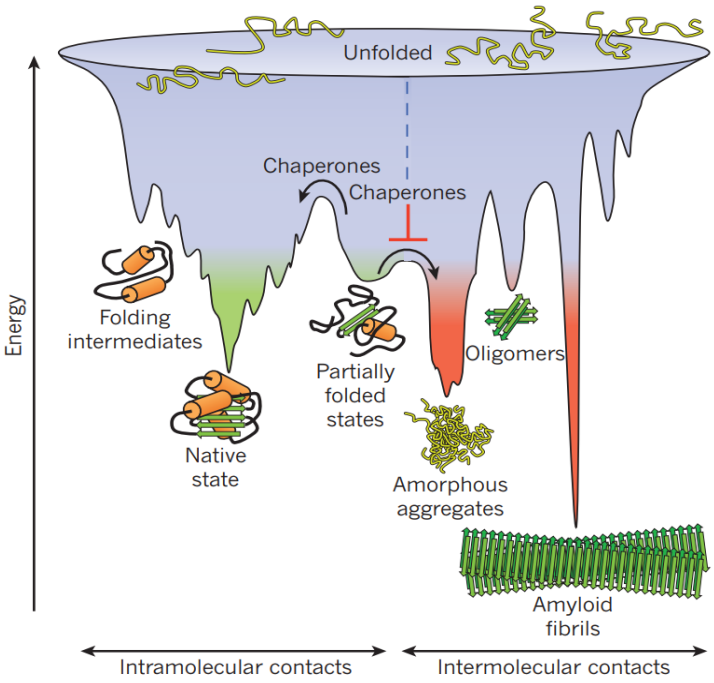


Figure 1. The protein folding energy landscape. The high-energy unfolded states can follow several scenarios: either follow the correct folding pathway through 1 or several folding intermediates, or become partially non-natively folded and eventually form aggregates or amyloid fibrils. Molecular chaperones act via binding of partially folded states and aborting aggregation events. In addition, they allow the bound protein to be refolded after a chaperone releases it. Courtesy of *Hartl, Bracher, and Hayer-Hartl 2011*.

Nonetheless, Hsps aren't expressed solely under stressful conditions, but also constitutively: e.g., the member of Hsp70 family, BiP, was identified in ER, where he facilitates folding of newly translated secretory proteins, such as immunoglobulins (*Haas et al. 1983; Gething and Sambrook 1992*). It indicates the generally important role of chaperones in *de novo* protein folding, when polypeptide chains are produced on ribosomes.

The crucial factor for Hsps correct function is the regulation of their working cycle: as a chaperone molecule binds to a target substrate, a mechanism for the substrate release must exist. This "switch" between the different states of certain Hsps during their working cycle is provided by adenosine triphosphate (ATP) - these chaperones belong to the ATP-dependent group (i.e., Hsp70, Hsp60, Hsp90 families), which are able to bind and hydrolyze ATP to bind and release unfolded substrates (*Flynn et al. 1991; Todd et al. 1994; Panaretou 1998*). On the contrary, there is the ATP-independent group of molecular chaperones acting via various mechanisms not involving any macroergic bond hydrolysis; among the well-known members of this group are small Hsps (sHsps), trigger factor, Spy-protein and SecB (*Mitra et al. 2022*). In addition to these, Hsp40s function with no need in ATP as well. However, their biochemical activity is tightly connected to the aforementioned Hsp70, when Hsp40 act as a substrate-presenting cofactors ("*co-chaperones*") for Hsp70 facilitating the substrate binding by Hsp70 and stimulating its ATPase activity; without these co-chaperone function the Hsp70 working cycle couldn't proceed (*Jiang et al. 2019*). The Hsp70-Hsp40 chaperone network, together with other participating proteins (e.g., nucleotide exchange factor, NEF) will be discussed later. Among other examples of "chaperone + co-chaperone" interactions is the bacterial GroEL-

GroES system. GroEL belongs to the Hsp60 proteins, which are generally present in symmetric oligomeric assemblies (14 subunits in the case of GroEL) forming 2 inner cavities for the client binding. Each subunit possesses an ATP-binding and ATPase activities, which are synchronized in the GroEL working cycle. After one half of the GroEL oligomer (a heptameric ring) is bound to 7 ATP molecules and to a client protein, the working cycle can't proceed without the interaction between GroES and the heptameric ring; the GroES docking onto the GroEL then induces ATP hydrolysis and the client refolding in the GroEL cavity (*Mayhew et al. 1996; Xu et al. 1997*). According to all the above, molecular chaperones do not function alone; instead, they very often require cofactors, either co-chaperones or small macroergic molecules, as ATP.

Major chaperone classes

In this paragraph I would like to briefly cover the most important chaperone families, including their structural and functional aspects, as well as their role in cell physiology. Beforehand, it's important to distinguish between so-called chaperonins - large oligomeric complexes with chaperoning activity - and chaperones.

Hsp60. Chaperonins

As was mentioned in the previous chapter, the Hsp60 family is mostly represented by the large oligomeric complexes, which are basically their functional form. Each monomer is approximately 60 kDa (therefore Hsp60). These ring complexes are symmetric and consist of 2 rings, each of them can be made up of 7 or 8 subunits and form a cavity capable of binding exposed hydrophobic regions of the non-native protein structures and facilitating the folding process (**Fig.2**). Chaperonins are ubiquitous and were identified in all three domains of life, as well as in semi-autonomous organelles (i.e., mitochondria), which highlights their importance in protein homeostasis in general (*Horwich et al. 2007*).

All Hsp60 complexes can be divided into 2 groups, according to their structural arrangement.

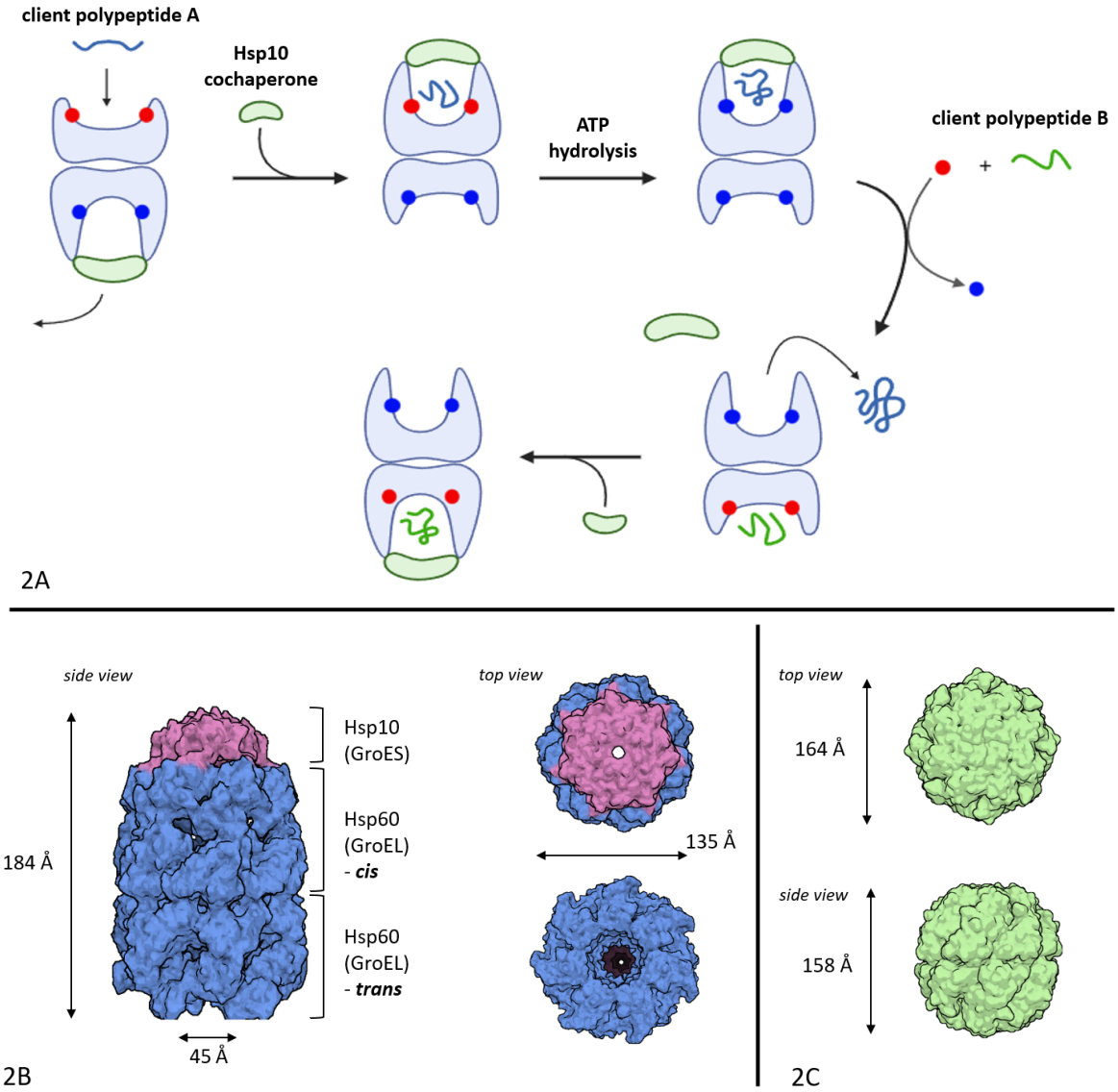


Fig.2. The Hsp60. **2A**. The working cycle of chaperonin (as example, type 1 Hsp60, e.g, GroEL, is depicted). After the *cis*-ring is bound to 7 ATP molecules, the cavity is ready to bind an unfolded client protein A and change the conformation . Thereafter, the Hsp10 molecule (GroES) docks onto the *cis*-ring and closes the aperture, allowing the folding catalysis and ATP hydrolysis. Then, simultaneous exchange of ADP for ATP in the *trans*-ring stimulates the dissociation of the Hsp10 and the folded client protein A. The *trans*-ring now becomes the new *cis* and can bind the client protein B. **2B**. The structure of the bacterial GroEL-GroES complex (PDB code 1AON). **2B**. The structure of the archaeal thermosome, type 2 Hsp60 (PDB code 1A6D). Created in Biorender® application. Adapted from *Horwich et al. 2007.*

Type 1 chaperonins are present in bacterial cytosol, mitochondria and plastids; each of the ring consist of 7 Hsp60 monomers, therefore the whole complex is 14-mer; in addition, type 1 chaperonins require cofactor, Hsp10, as was mentioned in the previous chapter. The most studied representative is the GroEL-GroES complex (**Fig.2B**), which functions via ATP-dependent cycle switching between its heptameric rings (*Xu et al. 1997*). Type 2 chaperonins were identified in eukaryotic and archaeal cytosol; their architecture slightly differs from the

type 1, namely their complexes are 16-mer (2 octameric rings) and they don't require Hsp10. In the **Fig.2C,** the structural model of the archaeal thermosome is depicted: notably, this is the ADP-state and the lids of both rings are closed. However, when the working cycle occurs and 8 ATP molecules are bound, each monomer on the one side "protrudes" its apical domain, so the overall conformational change results in the lid opening (mechanism similar to a camera aperture opening/closing) (*Ditzel et al. 1998, Pappenberger 2002*). Type 1 Hsp60s don't possess this apical domain - it's shortened instead; therefore they require an external "lid", which is Hsp10.

The functional cycle of both chaperonin groups is quite similar, despite evident structural differences. Multimeric chaperonin complexes are characterized by positive allosteric coupling of monomers within a ring, but between 2 rings there is a negative allostery - that is why it's possible to distinguish *cis*- (extended, with a client polypeptide inside) and *trans*-ring (no client bound) (**Fig.2A**). The switch between *cis*- and *trans*-state of a ring occurs in an ATP-dependent manner: the binding of a client protein and ATP molecules to a *cis*-ring triggers the conformational change and simultaneous ATP hydrolysis. This highly coordinated allostery allows chaperonins to rapidly proceed in their working cycle (*Rye et al. 1997*).

The physiological importance of chaperonins was proved by the experiment with deletion mutants: for instance, the deletion of GroEL gene in *E. coli* turned out to be lethal at all tested temperatures, i.e., from 17 to 30°C and at 42°C, which makes this gene essential (*Fayet et al. 1989*). Dramatically important proteins were identified among Hsp60 clients, such as actin and tubulin (for eukaryotic chaperonin TRiC/CCT) or mitochondrial enzymes, i.e., isocitrate dehydrogenase (for mitochondrial Hsp60) (*Dubaquie et al. 1998; Llorca et al. 2000*). It's noteworthy that chaperonins extensively interact with Hsp70 and work in conjunction with them. The transfer of newly translated polypeptide clients from ribosomes to DnaK (bacterial Hsp70) and consequently to GroEL-GroES was observed in bacteria (*Teter et al. 1999*). In similar fashion, the interaction between eukaryotic Hsp70 and TRiC/CCT (eukaryotic chaperonin) and the sequential substrate transfer was detected, with actin and firefly luciferase as substrates (*Lewis et al. 1992.; Frydman et al. 1994; Frydman and Hartl 1996*). It indicates the communication between various chaperone classes within the cell chaperone network.

sHsps

Small heat shock proteins (sHsps) are the ATP-independent, diverse group of chaperones, which molecular weight varies in the range from 12 to 43 kDa. They were identified in all

domains of life; a number of identified sHsps homologs increases from prokaryotes to eukaryotes with the highest number found in plants (around 20 in *A. thaliana*) (*Haslbeck et al. 2005*). The major mechanism underlying the physiological function of sHsps is preventing the aggregation process, especially under heat stress conditions (*Jakob et al. 1993*). Noteworthy, these chaperones tend to form large oligomeric assemblies, which represent their functional form: for instance, the Hsp16.5 oligomer from *Methanocaldococcus jannaschii* contains 24 subunits, whereas other representative complexes contain variable number of subunits (**Fig. 3A**) (*Kim et al. 1998*). sHsps assemblies are capable of binding several unfolded substrates at once through the exposed hydrophobic regions; the formed "substrate-oligomer" complexes are then very stable at normal physiological conditions. For the release of bound client proteins, interactions with other ATP-dependent chaperone systems are necessary, such as Hsp70-Hsp40 or Hsp100; without those interactions, a "chaperone - client" complex can't dissociate (*Lee et al. 1997; Mogk et al. 2003*).

Relatively low substrate specificity of sHsps makes them potentially very important chaperones for a substantial fraction of the cell proteome. For instance, in *Saccharomyces cerevisiae,* approximately 33% of cytosolic proteins are kept soluble under the heat stress conditions by 2 main sHsps, Hsp42 and Hsp26 (*Haslbeck et al. 2004*). It was also shown in humans that sHsps play an important role in various pathophysiological states: protection of neuronal and cardiomyocyte proteomes by Hsp27 from ischemia consequences; an increase in neurotoxicity of amyloid-β in neurons of Alzheimer's disease patients by αB-crystallin; the stabilization of abnormal proteins in case of autoimmune and misfolding diseases, and so on (*Latchman 2002; Stege et al. 1999; Clark and Muchowski 2000*).

Hsp90

Another dramatically important chaperone, Hsp90, was identified in eukaryotic and bacterial cytoplasm, as well as in ER and mitochondria (*Spence et al. 1990; Chen et al. 2005*). It is approximately a 90 kDa protein, which functions as a homodimer. It possesses the general chaperoning activity, i.e., preventing unfolded protein aggregation via binding exposed hydrophobic regions (*Wiech 1992*). Interestingly, it was shown that Hsp90 is less promiscuous in substrate choice than other ubiquitous chaperones and prefers a certain subset of cellular proteins. Namely, Hsp90 is crucial for the correct folding and maturation of enzymes and

receptors involved in signal transduction, e.g., glucocorticoid receptors, p53, cyclin-dependent kinases and so on. In addition, Hsp90 can either maintain a signal protein in its inactive form, or assist the signaling process by stabilizing complex formation (*Picard 2002*). The studies also indicate the role of Hsp90 in cancer development due to its ability to stabilize incorrect conformations of proteins involved in proliferation signal pathways. Therefore, Hsp90 can biochemically "buffer" mutated signaling proteins of cancer cells and help them to persist, despite their genetic instability (*Whitesell and Lindquist 2005*).

Hsp90 belongs to ATP-dependent chaperones, its working cycle and structure is depicted in **Fig.3D.** It's noteworthy that Hsp90 has numerous co-chaperones with various functions: Aha1, the ATPase activator; p23/Sba1, ATPase inhibitor (role in steroid receptor maturation); Hop1/Sti1, the adaptor linking Hsp70 and Hsp90 system; and many other (*Wandinger et al. 2008*). The latter, Hop1/Sti1, is responsible for physical transfer of a substrate protein from Hsp70 to Hsp90, thus Hsp90 machinery works downstream from Hsp70 (similarly to Hsp60-Hsp70 interaction). In this cooperation, Hsp90 plays a rather protective role for a client and induces the correct folding nucleation after release from Hsp70 (*Wegele et al. 2006; Morán Luengo et al. 2018*).

Ribosome-associated chaperones

According to the name of this chaperone group, they are situated on ribosomes or in their vicinity and somehow facilitate the folding of a newly translated polypeptide as it emerges from the ribosomal tunnel. In all life domains, there is a specific class of chaperones, which are involved in interaction with ribosomes and cotranslational folding: bacterial trigger factor (TF), eukaryotic ribosome-associated complex (RAC), eukaryotic and archaeal nascent polypeptide-associated complex (NAC) with numerous cofactors. The downstream chaperone machinery includes not only the direct ribosomal interactor, but also Hsp70-Hsp40 and chaperonin systems, as well as some other, like prefoldin *(Preissler and Deuerling 2012*). I would like to briefly describe solely TF, since "the ribosome-associated chaperones" is a very broad definition and comprises many different non-homologous proteins of different properties and structures. TF is the most studied representative found only in bacteria and chloroplasts. The 97.5 kDa moiety, it contains 3 domains, one of which possesses a peptidyl-prolyl *cis-trans* isomerase (PPIase) activity, and the rest 2 domains (N- and C-) are required for interaction with the large ribosomal unit (particularly the tunnel exit) and an emerging polypeptide (**Fig.3B**).

The N- and C-domains together form a cavity, which is lined with hydrophobic residues and can accommodate a polypeptide chain via hydrophobic interactions; through this mechanism TF prevents the risk of the initial protein aggregation or targeting for degradation, as well as delays the folding process (*Ferbitz et al. 2004; Hoffmann et al. 2006*). The studies have shown the substantial overlap between TF clients and the substrate pool of cytosolic Hsp70 and chaperonins – if the deletion of TF alone didn't cause much of protein aggregation, the additional suppression of Hsp70 activity was followed by the dramatic increase of misfolded and aggregated proteins. For the bacterial chaperonin GroEL, the simultaneous knock-out of DnaK (Hsp70) and TF led to a high substrate overload of the chaperonin (*Deuerling et al. 1999; Kerner et al. 2005*).

Hsp100

One of the largest chaperones (in terms of monomer MW), Hsp100 belongs to the ATP-dependent disaggregases, i.e., proteins capable of disassembling large protein complexes or aggregates potentially toxic to cells. It's another representative of ring-shaped oligomeric chaperones (among chaperonins and sHsps); each subunit has 1 or more ATP-hydrolase domains, the whole assembly typically hexameric (**Fig.3C**) (*Schirmer et al. 1996*). The Hsp100 complexes can have a central pore (not necessarily present), which is, in case of yeast Hsp104, only 25 Å wide and can't allow even a partially folded substrate to pass through (versus 45 Å in the GroEL-GroES complex). Instead, the whole hexameric assembly dissolves protein aggregates and apparently pulls an unfolded polypeptide chain through itself, which is strictly controlled by the ATP-driven working cycle (*Parsell et al. 1994; Deville et al. 2017*). Hsp100 were found in all life domains and belong to the wide class of AAA+ ATPase superfamily and their activities are often coupled to proteolysis, for instance ClpA and ClpX chaperones, which associate with the ClpT protease in bacteria (*Schirmer et al. 1996*). It was discovered that Hsp100 oligomers interact with Hsp70, which mostly involves the substrate transfer during the aggregate dissolution process; the substantial number of possible mechanisms of Hsp100-Hsp70 interplay was discovered in recent years (*Doyle and Wickner 2009*).

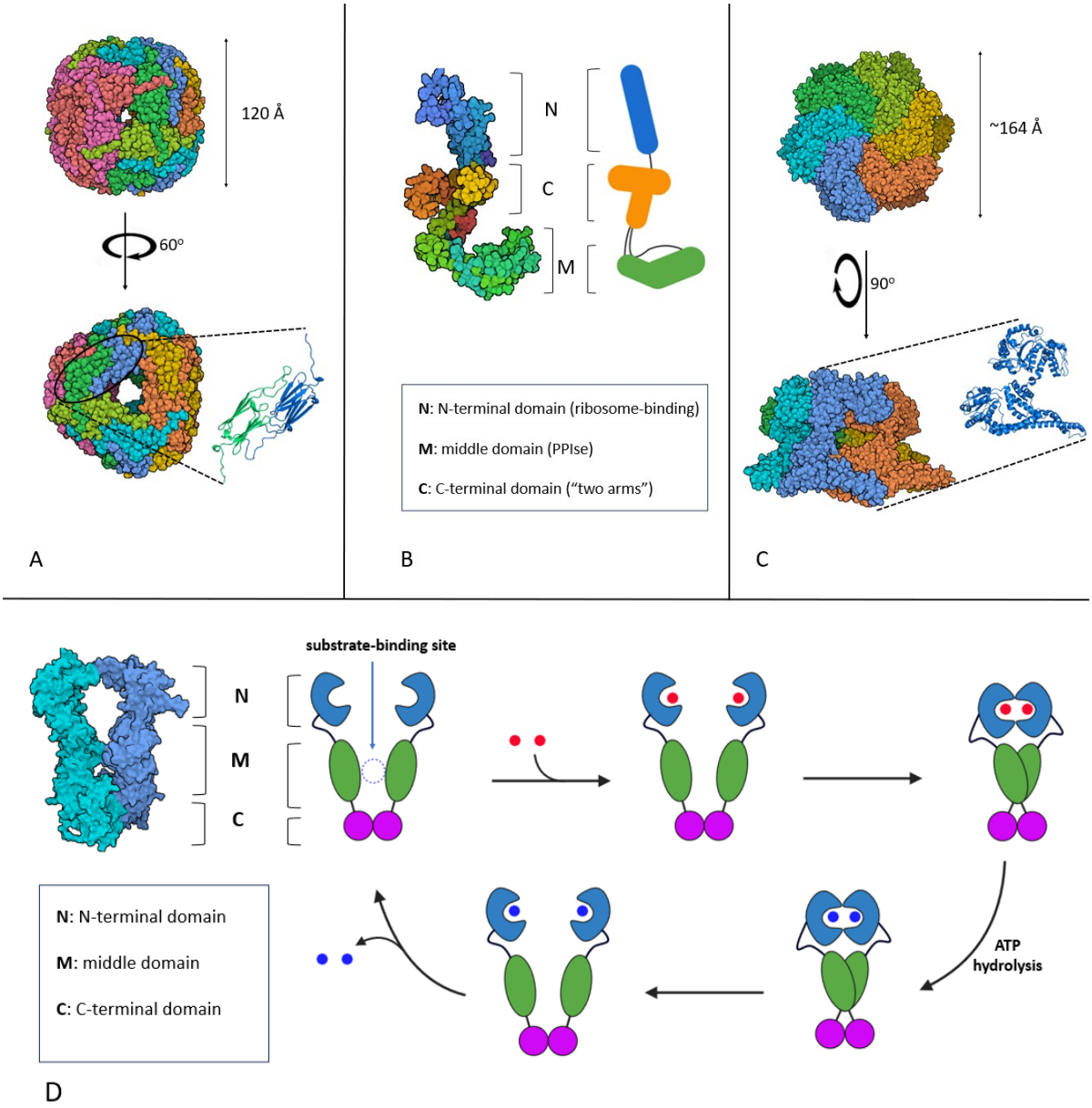


Fig.3. The representatives of the various chaperone classes.**A.** The oligomeric assembly formed by Hsp16.5 from *M. jannaschii,* PDB code 1SHS. The complex consists of 24 identical subunits; the assembly occurs via dimer intermediate (below), which is normally present in cell cytosol. **B.** The bacterial TF (*E.coli*), PDB code 1W26. **C.** The ClpB complex (Hsp100, *E.coli*), hexameric, in ATP-bound state. **D.** The structure and ATP-driven cycle of Hsp90. Its functional form is the almost parallel homodimer (PDB code 2IOP), where each subunit is capable of ATP binding and hydrolysis (N-domain). A client polypeptide is bound between 2 adjacent M-domains, while the C-domains are responsible for constitutive dimerization. After 2 ATP molecules (*red circles*) binding, Hsp90 undergoes substantial conformational changes, which result in N-domains docking onto M-domains and the twist of the whole assembly. When ATP is hydrolyzed, the conformation returns into its initial state and a client protein and ADP (*blue circles*) molecules are released. Created in `Biorender® application and PyMOL®`. *Ferbitz et al. 2004; Kim et al. 1998; Deville et al. 2017*; *Wandinger et al. 2008*.

The Hsp70 chaperone family wasn’t described in this chapter, since it’s the main target of this study and it will be described in detail in the following sections.

The Hsp70 protein family

While describing all major chaperone classes in the previous section, each time there was evidence of Hsp70 involvement in the interactions with all of Hsps above. One of the most crucial and well-studied chaperone systems, Hsp70, was identified in cytosol of bacteria, certain archaea (*T. maritima, A. pyrophilus*) and eukaryotes, as well as in the various eukaryotic compartments (i.e., mitochondria, plastids, ER, nucleus) and even plasma membrane (*Lindquist and Craig 1988; Michels et al. 1997; Gribaldo et al. 1999; Gehrmann et al. 2005*). Members of the Hsp70 family were reported to participate in a wide range of intracellular processes related to protein folding and homeostasis; these Hsp70 functions can be divided into "housekeeping" (i.e., under normal physiological conditions) and "stress-related". The first group includes *de novo* folding immediately after translation (in cooperation with other ribosome-associated chaperons); assembly and disassembly of oligomeric complexes; protection from proteolytic degradation; membrane translocation into organelles like mitochondria or plastids etc.. The stress-related functions comprise refolding of misfolded proteins; protein targeting to degradation; impeding protein aggregation or disaggregation and so on (*Rosenzweig et al. 2019*). All of the major Hsp70 cellular functions are depicted in **Fig.4,** not mentioning some specific ones. Due to its widespread occurrence and essential functions, it's not unexpected that all Hsp70 shares a very high homology level and sequence identity (not lower than 40-50%) (*Lindquist and Craig 1988*). Both in the bacterial and eukaryotic genomes there are usually several Hsp70 encoded, for instance 3 genes in *E. coli* K-12 strain and at least 13 in *H. sapiens* (*Genevaux et al. 2007; Radons 2016*). Such redundancy in the Hsp70 gene number per genome emphasizes specific functions fulfilled by certain homologs and provides an additional guarantee for the presence of functional Hsp70 molecules in cells.

In this chapter I would like to thoroughly discuss the structural and sequence aspects of Hsp70 representatives, their interactome (with the focus on the immediate Hsp70 cofactors) with relation to their working cycle, the latest finding regarding their oligomerization properties and the specific features and role of human HSC70, which is the object of this study.

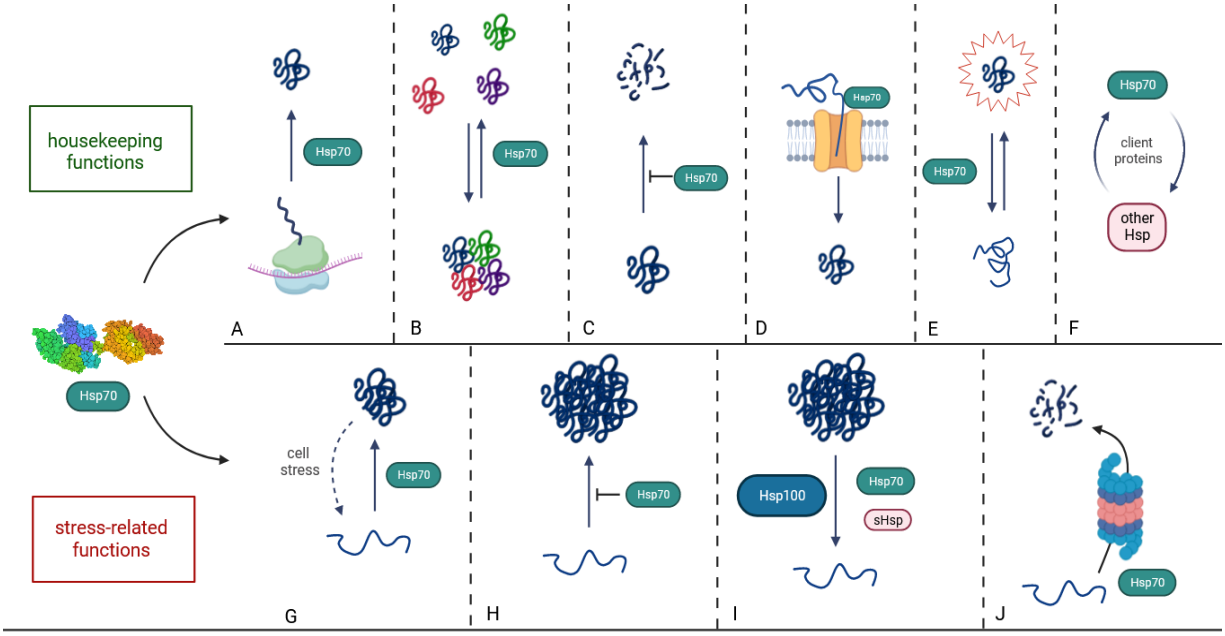


Fig.4. The diverse Hsp70 functions in cells, divided into 2 main categories - housekeeping (**A-F**) and stress-related (**G-J**). **A.** *de novo* protein folding (associated with ribosomes); **B.** assembly/disassembly of multimeric protein complexes; **C**. protection from proteolysis; **D.** assisting membrane transfer of proteins; **E.** regulation of protein activity via conformational changes; **F.** cooperation with other Hsp systems; **G.** refolding of denatured proteins; **H.** preventing protein aggregation; **I.** disaggregation (in collaboration with sHsps and Hsp100); **J.** protein targeting to proteolytic degradation (e.g., to proteasomes). Adapted from *Rosenzweig et al. 2019*.

Hsp70 structure and domains

As was mentioned above. all Hsp70 share high sequence identity and their structural organization is even more conserved in evolution (**Fig.5A**) As a model example, the bacterial chaperone DnaK (*E.coli*) will be exploited to show the general domain organization of Hsp70.

The 66.6 kDa protein, DnaK, consists of 2 distinctive domains. The larger 45 kDa N-terminal domain is named NBD (nucleotide-binding domain) for its ability to bind and hydrolyze ATP; the smaller 15 kDa C-terminal domain is SBD (substrate-binding domain) responsible for a client polypeptide binding. 2 domains are connected via short disordered linker primarily consisting of hydrophobic amino acids; in addition, a disordered C-terminal tail of variable length protrudes from SBD (*Rosenzweig et al. 2019*). However, the binding of a single ATP molecule to the NBD entails the drastic conformational change in SBD: the α-helical lid of SBD gets opened, whereby a client protein can fit into the SBD cleft (**Fig. 5B-D**).

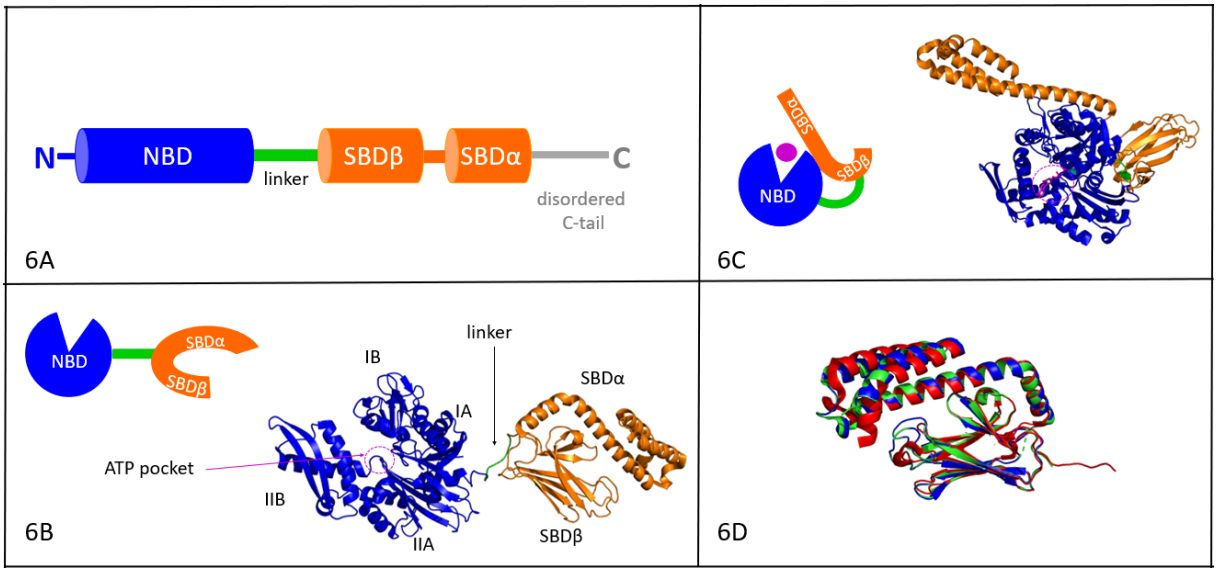


Fig.5. The Hsp70 structure. **A.** The Hsp70 general domain organization. **B.** *E.coli* DnaK structure and cartoon model in ADP/apo state **(**PDB code 2KHO). **C.** DnaK in ATP state (PDB code 4B9Q); magenta - ATP molecule. **D.** Structural alignment of Hsp70 SBD coming from *H. sapiens* (HSP72, green, PDB code 4PO2), *E.coli* (Hsp70, red, PDB code 4JNF) and *S. cerevisiae* (Ssb1, blue, homology model created on SWISS-MODEL server); it illustrates the very high level of homology within the Hsp70 family. Created in PyMOL®.

The thorough structural analysis of both domains revealed their mechanism of action. The X-ray diffraction model of NBD showed it consists of 2 lobes divided into 4 subdomains - IA, IB, IIA and IIB. An ATP molecule is accommodated in the deep cleft between the lobes, which both allosterically cooperate for ATP binding and hydrolysis. It was additionally supported by the comparison of ATP and ADP states of NBD, where in case of bound ATP the lobes were packed slightly tighter. The overall NBD fold is nearly identical to the actin-fold and its nucleotide-binding core is similar to that of hexokinase; the same amino-acid positions in the core were identified also in hexokinase, which might suggest a similar mechanism of action (even though NBD possesses no kinase activity). Therefore, the authors proposed the same protein ancestor for Hsp70 NBD, actin and hexokinase (*Flaherty et al. 1990; Flaherty et al. 1991*).

The ATPase cycle of NBD is highly connected to the SBD, the second major domain capable of binding exposed hydrophobic regions of unfolded polypeptides. SBD consists of 2 subdomains: SBDβ, which is basically a twisted 2-layer β-sheet of 8 antiparallel β-strands, and SBDα, the purely α-helical structure, which serves as the overly flexible “lid” (*Morshauser et al. 1995*). The peptide with high affinity to DnaK identified with phage display (sequence NRLLLTG) was used to analyze “SBD + model peptide” complex. It was discovered that mainly Van der Waals hydrophobic interactions are responsible for SBD affinity to the peptide, with first two Leu residues of the peptide contributing the most; in addition, several hydrogen bonds were detected with the SBD backbone groups, however, the importance of hydrophobic residues like Leu, Ile and Met in a client sequence was highlighted (*Zhu et al. 1996*).

These 2 functionally unrelated domains, NBD and SBD, are joined by the short linker, which proceeds from the IIB NBD subdomain to SBDβ. This interdomain linker is highly conserved among all Hsp70 homologs and is absolutely essential for the correct Hsp70 functional cycle - it allosterically couples ATP binding and hydrolysis, which is carried out by NBD, to substrate binding and release by SBD. The linker itself is composed of 6-8 amino acids (in DnaK and in most other homologs residues 386-394), with 4 almost exclusively hydrophobic residues in the middle part (Leu, Val, Ile, Met) flanked by 2 hydrophilic residues, namely by conserved Asp or Glu (**Fig.6**). It was shown that mutation of D393 leads to the failure of transitioning the allosteric signal from NBD to SBD; similar results were obtained in case of the preceding hydrophobic amino acids in the linker (*Vogel et al. 2006*). In the ADP state of Hsp70, the linker remains disordered, so 2 domains are very mobile relative to each other, resembling "2 beads on the string" (**Fig. 5B**). However, upon the ATP binding event, the linker docks onto NBD forming a β-sheet with IIA subdomain of NBD, facilitating the docking of SBDα onto NBD (**Fig. 5C**). The overall conformational changes of Hsp70 during the allosteric cycle are discussed in detail in the following chapter.

The disordered C-terminal tail is the least conserved region of Hsp70; its length slightly varies among homologs, in DnaK being 35 residues and human HSP72 around 27-28 residues (Uniprot entries P0A6Y8 and P0DMV8 respectively). In eukaryotic Hsp70s, the C-terminus contains a conserved motif (sequence EEVD) interacting with tetratricopeptide repeat (TPR), which is present in the sequences of many Hsp70 cofactors, e.g., Hop, Hip, class III JDP and CHIP (*M. P. Mayer and Bukau 2005*). It was discovered that mutations in DnaK C-tail lead to the significant decrease in DnaK refolding activity *in vitro* and thermotolerance of cells. Moreover, according to the same study, the DnaK C-tail isn't responsible neither for the interaction with co-chaperones or client proteins, nor for maintaining the interdomain allostery (*Smock et al. 2011*). This is also in agreement with the fact that bacteria don't have analogs of TPR-containing Hsp70 cofactors and there is no EEVD motif in bacterial Hsp70 C-terminus. In addition, in yeasts the C-tail of Ssa1 (*S. cerevisiae* cytosolic Hsp70) was shown to enhance substrate recognition; mutations in this region led to impaired thermotolerance, similarly to the DnaK case mentioned above (*Gong et al. 2018*).

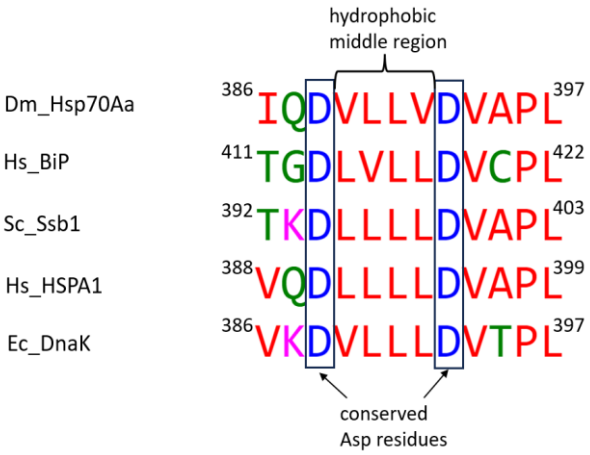


Fig.6. Multiple sequence alignment (MSA) of the Hsp70 linker region from the various Hsp70. The middle hydrophobic regions mainly contain Leu and Val; it is flanked by highly conserved Asp residues. Dm - *Drosophila melanogaster,* Hs - *Homo sapiens,* Sc - *Saccharomyces cerevisiae,* Ec - *Escherichia coli.* MSA was performed on the Clustal Omega (EMBL) server, the linker sequences were extracted from the Uniprot database.

To get a better understanding of the domain interplay of Hsp70, it's necessary to discuss the Hsp70 allosteric cycle and how the main co-chaperones are involved in it. According to all the above, every structural moiety of Hsp70 is necessary for the complete chaperone activity, and it will be supported even more by the following chapters.

Hsp70 working cycle. Role of co-chaperones.

Despite minor dissimilarities between various Hsp70 homologs, especially between prokaryotic and eukaryotic variants, the general features are conserved in all of them. The substrate binding and release by SBD is driven by ATP turnover in NBD (i.e., binding, hydrolysis and exchange). As was discussed previously, Hsp70 molecule can exist in 2 main conformational states: 1) the ADP/apo form, also named *the high-affinity state*, when the SBDα-lid is docked onto SBDβ (closed), trapping a client protein in SBD and not allowing it to refold; 2) the ATP form, or *the low-affinity state*, when the lid is opened and docked onto NBD, leaving the SBD-pocket exposed and allowing the previously bound protein substrate dissociate and proceeding with the binding of a next one (**Fig.7**) (*M. P. Mayer and Bukau 2005*). The mentioned affinity of Hsp70 to a client protein in ATP state is approximately 10-fold lower than in ADP form, since the binding of ATP to NBD increases the rate of dissociation of "SBD-client" complex 1000-fold and the rate of association around 100-fold (*Rosenzweig et al. 2019*). In the ATP-state, the linker and SBD are bound to NBD: it's noteworthy that while ATP binding entails the relative rotation of all 4 NBD-lobes, so the ATP-pocket gets smaller. The SBDβ then stabilizes ATP-NBD in such a way that ATP hydrolysis isn't energetically favored - the corresponding residues in the ATP pocket aren't suitably oriented for that. This explains why the intrinsic ATPase

activity of Hsp70 is quite low (for DnaK *in vitro* - approximately 1 ATP molecule per 1 DnaK molecule per minute) in the absence of necessary co-chaperones or protein substrates (*Liberek et al. 1991; Kityk et al. 2012*). When a client protein binds to SBD in this state, the occurring hydrophobic interactions stimulate the conformational shift in SBDβ, which then propagates into detachment of SBDβ and the linker from NBD. Consequently, NBD isn't clamped by SBDβ anymore, so it can proceed with the NBD lobes relative rotation into the favorable for ATP hydrolysis conformation. However, the increase of DnaK ATPase activity upon substrate binding was only 1.5- to 3.0-fold and *in vivo* is significantly enhanced by the co-chaperone, DnaJ (*Kityk et al. 2015*). Immediately after the ATP hydrolysis, SBDα dissociates from NBD and closes the SBD-pocket with the bound substrate, physically trapping it in an unfolded state (*Zhu et al. 1996*). The release of a substrate occurs much more probably when the lid is opened: the studies conducted on the lidless form of DnaK demonstrate that the absence of SBDα increases the dissociation and association rate of a client peptide more than 100-fold (*Buczynski et al. 2001*). The lid opening is triggered by the exchange of ADP for ATP, which entails the binding of SBDβ to NBD and the SBDα docking onto NBD. It allows a substrate to dissociate from SBD and the cycle can repeat again (**Fig.7**).

As was mentioned above, the allostery cycle is highly dependent on several cofactors, which mediate substrate binding and stimulate ADP/ATP exchange or ATP hydrolysis. One of the most important and extensively studied classes are J-domain proteins (JDPs) and nucleotide-exchange factors. I would like to shortly review their structural organization and their role in Hsp70 functionality.

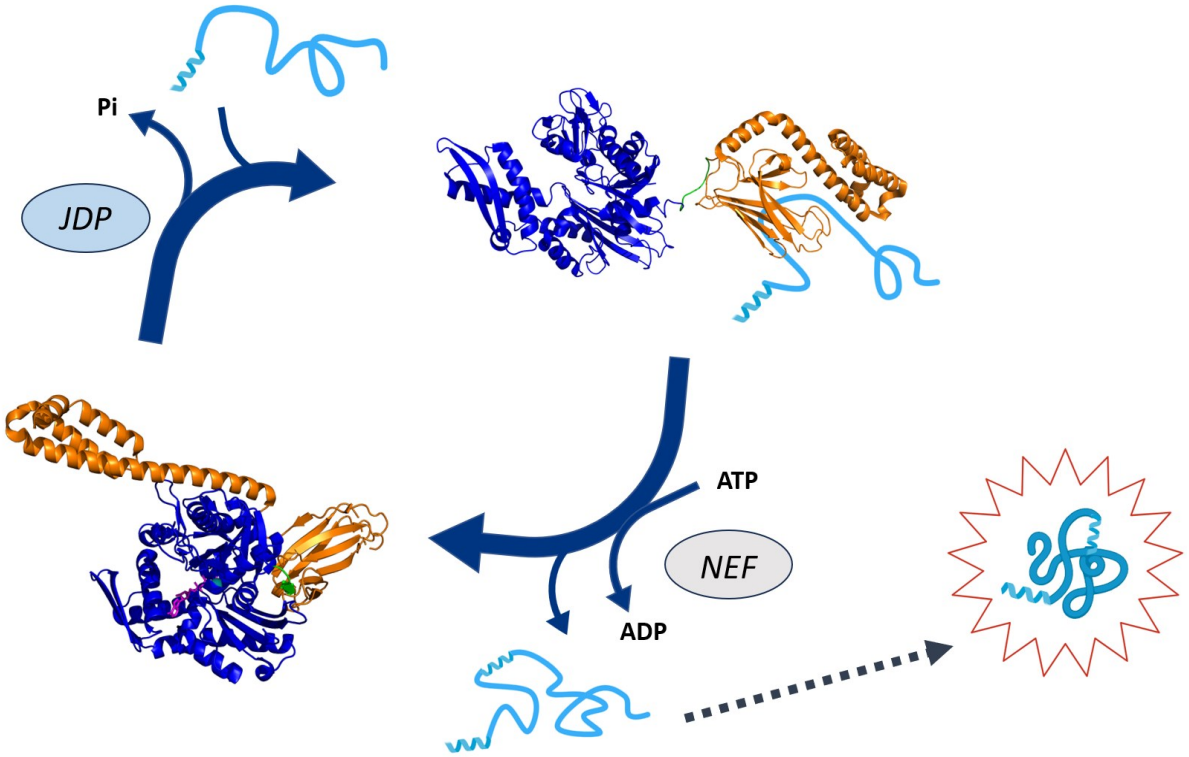


Fig.7. Allosteric cycle of Hsp70. In the low-affinity, ATP-bound state, Hsp70 is ready to bind a misfolded/unfolded protein substrate, which is coupled to the ATP-hydrolysis event; the whole step is facilitated by a JDP. Once the lid is closed and the client is bound to SBD (high-affinity, ADP-bound state), the ADP/ATP exchange enhanced by NEFs is required for the substrate release and the Hsp70 conformational switch. The released partially-folded client is therefore able to proceed in the native folding pathway and acquire its functional form. Created in PyMOL ®. Adapted from *Rosenzweig et al. 2019*.

JDPs

These Hsp70 protein cofactors, often referred to as Hsp40, are ubiquitous at the similar extent as Hsp70 itself, because their activities are closely intertwined. The key structural feature of all JDPs is the presence of 70-amino acid J-domain, which is basically an α-helical hairpin loop and is usually located at the N-terminus of JDP polypeptide. This domain is the most important, since it is able to stimulate ATPase activity of Hsp70 (*Kelley 1998*). JDPs are divided into 3 main families: A, B and C, based on their similarity to the best-studied member of JDPs, bacterial DnaJ. The DnaJ itself belongs to the class A and consists of: N-terminal 70-aa J-domain, followed by the Glycine/Phenylalanine-rich region of 30-40 aa (G/F-rich); the β-sandwich C-terminal domain 1 and 2 (CTD1 and CTD2), whereas the Zinc-finger-like domain (ZFLD) is included into CTD1; the C-terminal dimerization domain (DD) with the disordered C-tail (**Fig.8A**). While the N-terminal J-domain is responsible for interaction with Hsp70 and

stimulation of its ATPase activity, C-terminal domains are necessary for client polypeptide binding and dimerization, since JDPs function in the dimer form (*Goffin and Georgopoulos 1998; Kelley 1998; Y. Wu et al. 2005*). Class B JDPs has a very similar domain organization as class A, except usually longer G/F region and absence of ZFLD domain; class C is the most diverse and least conserved, with only J-domain homologically shared with DnaJ (*Cheetham and Caplan 1998*).

JDPs generally interact with both Hsp70 domains, NBD and SBD: the Histidine-Proline-Aspartate (HPD) motif present in J-domain forms multiple hydrophobic and polar contact with 2 interfaces of Hsp70, the NBD-SBDβ and the linker-SBDβ interfaces. As a result, the HPD motif is in close contact with the ATPase catalytic center of Hsp70 NBD, promoting its activity significantly upon the substrate binding to Hsp70 SBD (**Fig.8B**) (*Kityk et al. 2018*). As was mentioned above, the other, equally important function of JDPs is the substrate presenting to Hsp70. The promotion of ATP hydrolysis is tightly coupled to this role of JDP, since together they represent an integrated stimulus leading to the conformational switch of Hsp70 described earlier, from low-affinity to high-affinity state. The affinity of JDPs for unfolded substrates is located within C-terminal domains; the peptide affinity screening for DnaJ revealed a core motif consisting of 8 mainly hydrophobic and aromatic residues, very similar binding sites for DnaK as well (*Rudiger 2001*). JDPs therefore serve as a sort of "substrate scanners", preselecting client polypeptides for Hsp70. The binding of a substrate occurs in JDP dimer via 4 sites (4 CTD domains), each one of them has low-affinity binding properties and the sufficiently stable complex is formed only if all 4 sites are engaged in this process (*Jiang et al. 2019*). The other significant aspect of such mode of binding is its transient nature: modular low-affinity binding enables more rapid substrate transfer from Hsp40 to Hsp70. The interactions between Hsp40 and Hsp70 remain dramatically important for the whole refolding machinery; it was experimentally shown that two ZFLD in class A JDPs are also quite important for that. In fact, ZFLDs provide additional interaction sites for DnaK, which ensures the client transfer from Hsp40 machinery to Hsp70; it seems ZFLD fulfills the similar function within the whole class A JDPs (*Linke et al. 2003*). It's noteworthy that biochemical comparison of class A and B JDPs from yeast cytosol revealed the 4-times more efficient refolding of a model substrate, luciferase, in case of class A JDP. It might be due to more extensive interaction between Hsp70 and Hsp40 via the aforementioned ZFLDs, which are absent in class B JDPs (*Lu and Cyr 1998*). Interestingly, the disordered C-tail in class A JDPs also seems to be involved in this substrate transfer: it exhibits a competitional

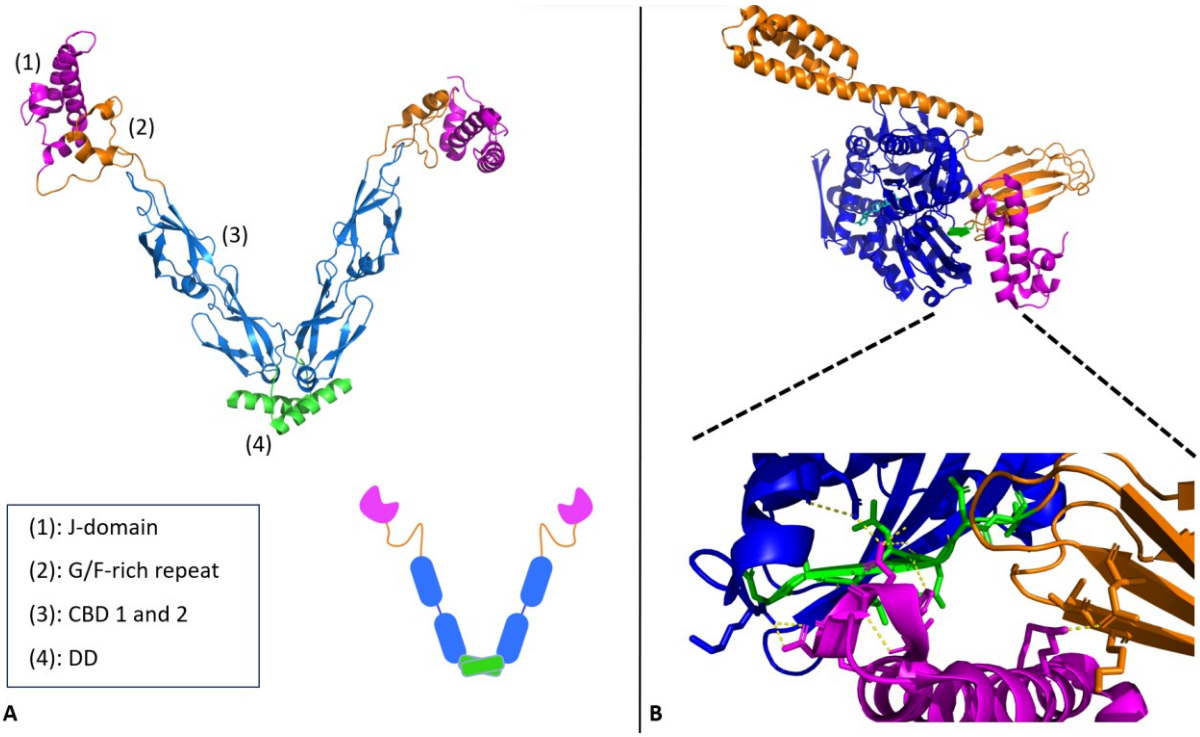


Fig.8. The structure of JDP dimer and interactions with Hsp70. **A.** The model and the cartoon of DnaJ dimer (bacteria Thermus thermophilus, PDB code 4J80). The individual colors and numbers indicate the domain organization. The disordered C-tail isn't depicted. **B.** The complex of Hsp70 in ATP-state and J-domain (E.coli, PDB code 5NRO). J-domain interacts with the NBD-SBDβ and the linker-SBDβ interfaces, engaging H-bonding and polar contacts between amino acid backbone atoms and sidechains; magenta - J-domain, blue - NBD, orange - SBD, green - Hsp70 linker, cyan - ATP molecule. Created in PyMOL® and Biorender® . Based on Mayer 2021.

binding to the one of CBD substrate-binding domains, therefore displacing the bound protein substrate and working as an autoinhibitory mechanism for JDPs (*Jiang et al. 2019*).

Hsp70 NEFs

Nucleotide-exchange factors, or NEFs, are necessary for Hsp70 cycle continuation, since they accelerate the exchange of ADP for ATP and therefore SBDα lid opening and substrate release. For GrpE, NEF homolog, the observed increase in ADP/ATP exchange was 5000-fold; however, the Hsp70 cycle proceeding is strictly dependent on simultaneous presence of both JDP and NEF, since they are responsible for different phases of "Hsp70-nucleotide" interactions (*Liberek et al. 1991; Packschies et al. 1997*). One of the most studied NEF type, GrpE-like, is represented by a parallel homodimer assembly (each subunit is around 21.8 kDa) with a long N-terminal α-helical domain serving for dimerization and C-terminal small β-sheet domain used for the interaction with Hsp70-NBD (*Harrison et al. 1997*) (**Fig.9).** GrpE-like NEFs are known as prokaryotic, mitochondrial and plastid Hsp70 NEFs; the eukaryotic

counterparts belong to 3 other known classes - Hsp110 type, Bag type and Armadillo type, which barely share homology, but nevertheless similarly acting via stabilization of the open low-affinity Hsp70 conformation (reviewed in *Rosenzweig et al. 2019*). I would like to briefly describe the mode of action only for GrpE-like NEFs, since the Hsp70 NEFs aren't the object of this study.

As was discussed earlier, a NEF functions upon the ATP hydrolysis performed by NBD of Hsp70, which is stimulated by the substrate binding and interaction with JDP. In the ADP-state, there are 6 sites of interaction between GrpE and DnaK: the main 2-3 are located between the GrpE β-sheet domain and IB and IIB subdomains of DnaK NBD, from the both sides of the nucleotide-binding cleft (**Fig.9B**). The rest is situated along the long GrpE α-helical domain, which contacts the IIA subdomain of NBD. The contacts between GrpE and NBD include salt bridges, hydrophobic interactions and H-bonds. The length of α-helical domain is sufficient to interact with DnaK SBD, facilitating the substrate release; nonetheless, even N-terminally truncated GrpE was able to assist the model substrate refolding, luciferase (*Szabo and Hartl 1994; Harrison et al. 1997*). For the efficient interaction, the GrpE-signature motif, a loop in the IIB subdomain of NBD, is particularly essential and specific for GrpE-like-NEF-interacting Hsp70 according to the data (*Brehmer et al. 2001*). The mechanism itself is underpinned by simple induction of NBD cleft opening and lowering the activation energy for the nucleotide exchange. This occurs via insertion of one of the GrpE β-sheet domain into the nucleotide-binding cleft of NBD and widening it (*Packschies et al. 1997; Harrison et al. 1997*).

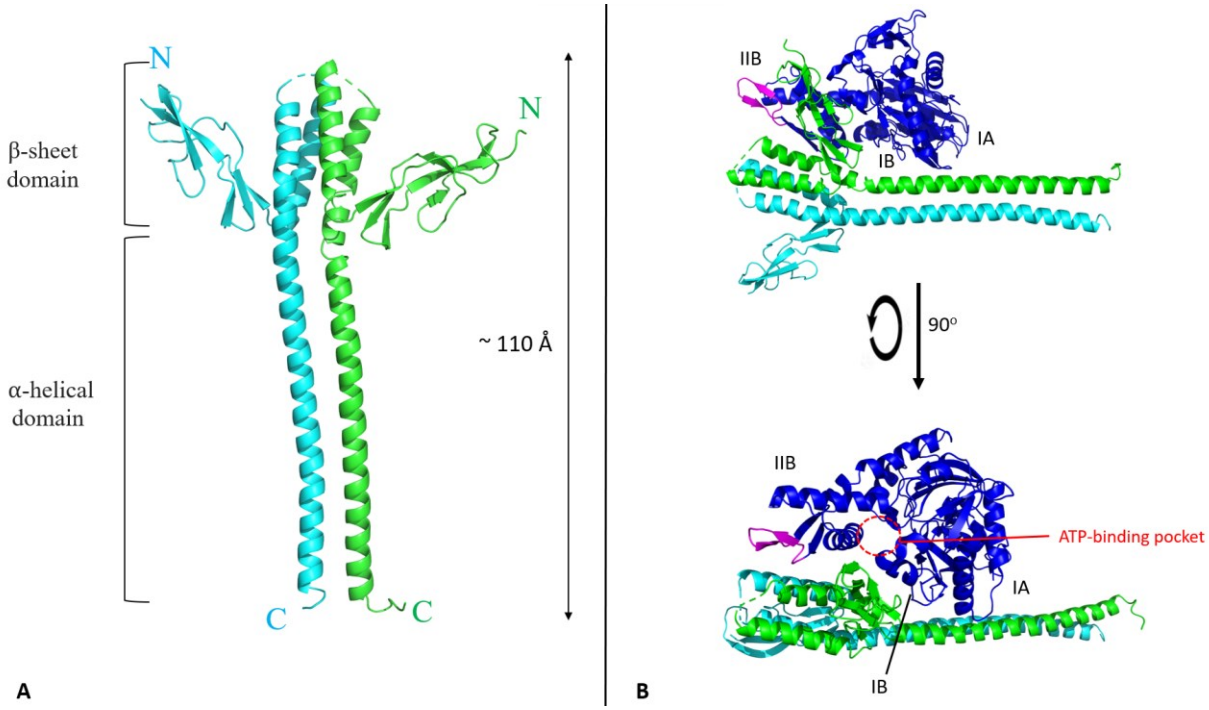


Fig.9. The structure of GrpE, the NEF for DnaK. **A.** The functional homodimer model of *E.coli* GrpE. Due to the long α-helical domain, the overall length of the structure is about 110 Å. **B.** The complex of GrpE and DnaK NBD. N-terminal β-sheet domain of GrpE and subdomains IIB and IB of NBD create the most important contact interface between the molecule; β-sheet domain gets inserted between the NBD lobes, therefore widening the nucleotide pocket and supporting the nucleotide exchange. *Blue* - DnaK NBD, *cyan and green* - GrpE subunits, *magenta* - GrpE-signature motif of DnaK. PDB code - 1DKG, created in PyMOL®. *Harrison et al. 1997.*

The Hsp70 dimerization

Decades ago it was discovered that Hsp70 molecules are able to dimerize in ADP-bound state, meanwhile the addition of ATP significantly decreases the dimer fraction *(Schmid et al. 1985*). These first evidences were then taken into account and investigated further with the ER-specific chaperone BiP: upon the substrate release, BiP is post-translationally modified, namely phosphorylated and ADP-ribosylated, which leads to its inactivation and formation of dimeric and higher-oligomeric BiP species. In addition, this mechanism is reversible - the peptide substrate binding and subsequent ATP hydrolysis by BiP stimulates the conversion into an active monomeric form capable of refolding activity (*Freiden et al. 1992; Blond-Elguindi et al. 1993*). The additional factor affecting the oligomerization process is JDP: bovine brain and yeast hsp70 were able to oligomerize into larger species in the presence of YDJ1, yeast homolog of DnaJ, and ATP, even though it was previously reported that ATP breaks down Hsp70 oligomers (*King et al. 1995*). From all the above it becomes clear that the members of Hsp70 family can undergo a reversible formation of dimer or even higher oligomers, which is regulated in a complex manner by multiple factors - PTMs, ATP or ADP, protein substrates and JDPs.

These data raised several questions regarding the exact mechanism of oligomerization, its precise regulation and the structure of Hsp70 oligomers.

Thompson et al. applied chemical cross-linking by glutaraldehyde to fixate the DnaK oligomers and analyze them biochemically *in vitro* and natively *in vivo.* It was discovered that: 1) heat stress elevated the oligomeric DnaK fraction in *E. coli*; 2) the cross-linked DnaK oligomers retain ATPase activity and prevents denatured substrates from aggregation, without refolding them (holdase activity); 3) the oligomers are still responsive to GrpE stimulation, but not to DnaJ ATPase stimulation (*Thompson et al. 2012*). The other interesting finding of the described study is that the ATPase-deficient mutant of DnaK (T199A) is primarily in a monomeric form. The discovery of retention of certain DnaK activities in oligomeric form *in vitro* confronts the earlier conclusion that Hsp70 oligomers represent an inactive chaperone pool *in vivo,* even though it's partially true - the primary activity of Hsp70, as a refoldase, in oligomeric state is abrogated and substituted to holdase activity.

The first attempts to define the structure of simplest oligomeric Hsp70 assemblies, dimers, were undertaken using the crystallographic symmetry of a Hsp70 homolog crystal. Upon the obtaining of ADP-bound GkDnaK (DnaK from *Geobacillus kaustophilus*) crystal, it was discovered that the region of the extended hydrophobic linker in one GkDnaK molecule perfectly fits into the SBD substrate pocket of another chaperone molecule; based on that, the initial model was built *in silico* (**Fig.10A**) (*Chang et al. 2008*). It is consistent with the fact that the linker sequence vastly matches the consensus SBD-binding peptide sequence of Hsp70-clients (see above). This *"substrate-like" dimer* model is further supported by the aforementioned finding about Hsp70 dimerization: a) ATP induces the dimer dissociation, which is similar to the native mechanism of Hsp70 substrate release; b) JDPs promote Hsp70 oligomerization, which putatively occurs via presenting of one Hsp70 molecule to another as a substrate; c) in ADP-bound state, SBD lid is closed, trapping the bound substrate, which in this case is a Hsp70 linker. Substrate-like dimer model containing DnaK from *G. kaustophilus* was identified by X-ray crystallography in later study - the assembly was stabilized by interaction with GrpE dimer, where both DnaK molecules are interacting by their NBDs with 2 N-terminal GrpE domains and simultaneously retaining substrate-like binding between DnaK proteins (*C.-C. Wu et al. 2012*). Other study indicating this particular type of dimer was conducted with BiP and its two mutants: 1) ADDA-BiP, with substituted 4 hydrophobic linker residues to Ala-Asp-Asp-Ala and therefore uncoupled interdomain allostery; 2) V461F-BiP, with mutated SBD and therefore unable to bind substrates. The linker mutant was also resistant to SubA digestion, which cleaves at the linker site of BiP, and can't be bound as a substrate by other Hsp70

molecules. Vice versa, SBD-mutant is susceptible to SubA digestion and can't bind an Hsp70 molecule as a substrate. According to analytical SEC data, the mixture of the linker and the SBD mutants form a dimer structure, which is resistant to SubA digestion; the monomeric fraction, corresponding to free SBD-mutant, which was additionally labeled by luminescent dye, has been completely digested, leaving the high-intensity luminescent peak corresponding to free SBD. Similarly, the linker mutant alone is primarily monomeric, because it's unable to dimerize with itself, similarly to SBD-mutant alone. To sum up, all the aforementioned data from the independent studies support the "substrate-like" dimer model for Hsp70 (*Preissler et al. 2015*).

Nevertheless, other Hsp70 oligomer models were recently discovered. With the help of native mass spectrometry and chemical cross-linking, antiparallel arrangement of human Hsp70 dimer was determined: with both Hsp70 molecules in ATP-bound state with SBD/SBD interface (**Fig.10B)** and with both molecules in ADP-bound state with the interface along the whole molecule (contacts between NBD and SBD, **Fig.10C**). Authors also considered the substrate-like Hsp70 dimer model; nevertheless, according to their data, the double Hsp70 mutant, which is unable to bind substrates and has mutated key phosphorylation site, was still able to dimerize (*Morgner et al. 2015*). In this study, the role of JDP in promoting Hsp70 dimerization was confirmed; in addition, a putative model of "Hsp70 dimer-JDP" was built, with each Hsp70 molecule in ADP-state interacting with the opposite JDP subunits via their C-terminal disordered tail. The importance of newly identified phosphorylation site in Hsp70 SBDβ subdomain, T501, was established: the MS signal intensity corresponding to Hsp70 dimer significantly decreased in case of dephosphorylated Hsp70. Another dimer model was created on the basis of DnaK-ATP X-ray crystallography data (**Fig.10D).** This assembly primarily involves the interface between NBD of 2 DnaK molecules, with minority of contacts occurring between one DnaK SBDα and second DnaK NBD; however, this model was considered biologically irrelevant and isn't apparently present at some detectable level in solution (*Kityk et al. 2012*).

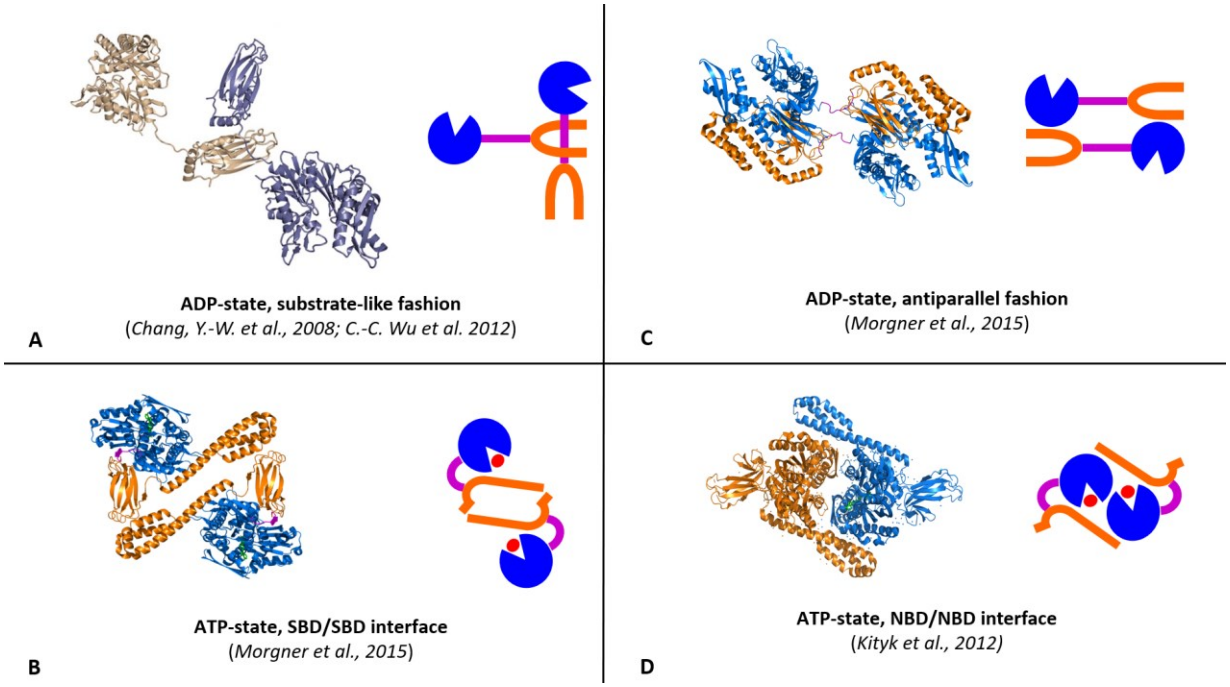


Fig.10. The putative Hsp70 dimer models. **A.** Substrate-like model, based on crystallography data (PDB code 4ANI); **B.** ATP-bound dimer model with SBD/SBD interface, based on cross-linking mass spectrometry (XL-MS) data (PDB code 4B9Q); **C.** ADP-bound antiparallel dimer model from XL-MS data (PDB code 2KHO); **D.** ATP-bound dimer model with NBD/NBD interface from crystallography data (PDB code 4B9Q).

The target Hsp70 homolog: human HSC70

The heat shock cognate protein 70, or simply HSC70, encoded by the gene *HSPA8*, represents one of the 13 human HSP70 subfamilies and shares the overall similarity in the domain organization and functional cycle described above (*Radons 2016*). However, there are some noteworthy differences and specific features of HSC70, which I would like to discuss in this chapter.

HSC70 is the constitutively expressed cytosolic chaperone; the sequence identity between HSC70 and "classic" stress-induced HSP70 (HSPA1A) is 85.6%, where the least similarity is observed for the variable C-terminal disordered region (pairwise alignment, EMBOSS Needle). The C-terminus of HSC70 was found to be a key interactor with Heat shock factor 1 (HSF1): under the heat stress, HSC70 is able to enter nucleus and modulate the activation of the transcription factor HSF1, which is responsible for the induction of heat shock related genes (e.g., stress-inducible chaperones) (*Ahn et al. 2005*). Among other differences, HSC70 has the

pattern of interaction with lipid bilayer different from HSP70 - it induces the aggregation of lipid vesicles in presence of calcium ions at much higher extent than HSP70 does (*Arispe et al. 2002*). Moreover, HSC70 promotes tumor cell growth more extensively, probably because of its constant presence in cell cytosol, oppositely to HSP70: in the knock-out experiments it was discovered that the deletion of HSC70 immediately kills both cancer and non-cancer cell lines, whereas deletion of stress-induced HSP70 only affected the viability of cancer cells (*Rohde et al. 2005*).

The variety of HSC70 functions is tightly connected to its localization in the cell. As was mentioned, it is mainly situated in the cell cytosol, but is also able to translocate to the nucleus for HSF1 activation. It can be located at the cell membrane periphery as well, participating in clathrin-mediated endocytosis. HSC70 was originally described as an ATPase, which promotes the dissociation of clathrin triskelions from clathrin-coated vesicles (*Chappell et al. 1986*). It was later established that disruption of HSC70 ATPase activity leads to the accumulation of clathrin-coated vesicles and also empty triskelion cages *in vivo*, which suggests the prominent role of HSC70 in the overall clathrin-mediated vesicle cycle (*Newmyer et al. 2001*). HSC70 is an important mediator of polypeptide transfer into various organelles (*Deshaies et al. 1988*). For instance, the cooperation of HSC70 and HSP90 allows them to interact with the import mitochondrial channel TOM40 via the TPR motif discussed earlier in this work. The binding of transferred mitochondrial preprotein by HSC70 and HSP90 allows to keep it in an unfolded state for the effective import and prevents its aggregation as first; at the same time, the chaperones stimulate the assembly of TOM40 complex on the outer mitochondrial membrane and mediate the preprotein recognition by TOM40 (*Sheffield et al. 1990; Young et al. 2003*). In the case of ER, the primary function of the HSC70/HSP40 complex is to post-translationally target the proteins with the C-terminal signal sequence (tail-anchor) into the ER membrane (*Rabu et al. 2008*). Much evidence of vast interaction between HSC70 and nucleus was found: 1) both HSC70 and HSP70 play an important role in maintaining the protein transport into nucleus; 2) the HSC70 and its ATPase activity is necessary for the efficient β-importin turnover, namely for its transfer back to the cytosol; 3) in plant cells, HSC70 mediates the targeting of nuclear pore associated WIT proteins, which are important for the proper RanGAP localization in the vicinity of nuclear pores (*Shi and Thomas 1992; Kose et al. 2005; Meier et al. 2010*). To conclude, the chaperone of our interest is profoundly important in membrane transport and associated processes.

HSC70 fulfills a standard set of Hsp70-associated activities in protein homeostasis, i.e., assisting the correct protein folding via binding exposed polypeptide hydrophobic regions in

cooperation with HSP40 co-chaperones (*Frydman and Hartl 1996*). Non-surprisingly, some fundamental cell proteins were detected among its clients. For instance, HSC70 is indispensable for the proper folding of cyclin D1 and the D1/CDK4 complex assembly, ensuring its stability and activation during the cell cycle (*Diehl et al. 2003*). The other example is actin monomers (G-actin), which acquires its native structure through the sequential interactions with HSC70 and TRiC/CCT, an eukaryotic chaperonin; the HSC70/HSP40 system is the initial binder, which then transfers actin to the chaperonin complex (*Frydman and Hartl 1996*). In vertebrates, HSC70, together with HSP90, is apparently involved in some intermediate steps of myosin folding during the striated myocytes maturation (*Srikakulam and Winkelmann 2004*). The substrate repertoire of HSC70 is quite broad - every protein above 20 kDa serves as a potential client for this cytosolic chaperones; it is additionally associated with *de novo* folding of polypeptide chains newly emerging from cytosolic ribosomes, among other chaperones involved in cotranslational folding (*Thulasiraman et al. 1999*).

The opposite side of protein homeostasis, the proteolytic degradation, also involves HSC70 activity. The ubiquitin-proteasome degradation pathway engages HSC70 as a factor for the correct E3-enzyme reaction (ubiquitin-ligase): the chaperone may act via binding a client and exposing the site for E3. Many important substrates for this reaction were identified, e.g., actin, histone H2A, p53, glyceraldehyde-3-phosphate dehydrogenase, α-lactalbumin, E2A transcription factor etc. (*Bercovich et al. 1997; Huang et al. 2004; Esser et al. 2005*). In addition, HSC70 is involved in recruitment of substrates for lysosomal degradation. Misfolded proteins are bound by HSC70 and then delivered to the lysosomal membrane receptor, LAMP2A, which is responsible for the uptake into the lysosomal compartment (*Cuervo and Dice 1996*). The later studies indicate the presence of HSP90/HSC70/HIP/HOP complex at the lysosomal membrane responsible for protein targeting to lysosomes (*Agarraberes and Dice 2001*).

According to the above, the constitutively expressed cytosolic chaperone HSC70 is involved in numerous processes in the cell, beyond just supporting protein folding. It is an essential component of cytosol and its activity is absolutely vital for the cell. It has been an object of investigation for many decades, as well as the whole Hsp70 family. Among described HSC70 features is its self-association, or oligomerization abilities, which were observed in case of other Hsp70 homologs (described in the previous section). Benaroudj et al. performed glutaraldehyde cross-linking, size exclusion chromatography and analytical ultracentrifugation experiments on HSC70 species: the presence of the dimer, trimer and higher oligomer fractions were confirmed by all three techniques (*Benaroudj et al. 1995*). In their paper, authors speculate of functional

implications of HSC70 oligomeric complexes: since the oligomers are stabilized in ADP-state, such complexes might have a higher affinity to substrates. This is also in agreement with the data that HSC70 is able to bind large oligomeric complexes as substrates. Nonetheless, structural models of HSC70 oligomers were never described.

Structural mass spectrometry of proteins

In recent decades, mass spectrometry (MS) has become a powerful tool not only for identification of biomolecules in complex samples, but also for determining their structure. The basic principle behind MS analysis is transferring biomolecules into gas phase with their concurrent ionization and separating them in the mass analyzer chamber according to their mass-to-charge ratio (m/z), which is eventually registered by a detector. The universal output is the mass spectra - individual m/z values and their corresponding intensities. In the modern tandem-MS setups, several mass analyzers are coupled together for individual ion selection, their subsequent fragmentation in gas phase by numerous techniques (via collision with inert gas - CID, by ultraviolet irradiation - UVPD etc.) and analyzing the emerged fragments; it allows to improve identification and perform selective MS experiments on biomolecules of your interest. The obtained mass spectra of such fragments are called MS/MS spectra or even higher (MS^3, MS^4 etc.), depending on the number of fragmentation steps performed. The typical mass spectrometer arrangement is depicted in **Fig.11;** nonetheless, the variety of possible setups is innumerable and their discussion is beyond the scope of this chapter (*Hoffmann and Stroobant 2007*).

The analysis of proteins by MS represents one of the major subfields within MS itself. There are several ways to study proteins by MS: 1) in an intact form using *native MS*, 2) ionizing an intact protein and fragmenting it in gas phase by a chosen technique (CID, ECD etc.) - *top-down approach*; 3) performing a preliminary enzymatic/chemical cleavage of a protein and analyzing the produced peptides by MS; 4) the peptides from enzymatic/ chemical digestion are additionally fragmented in the mass spectrometer - *bottom-up approach.*

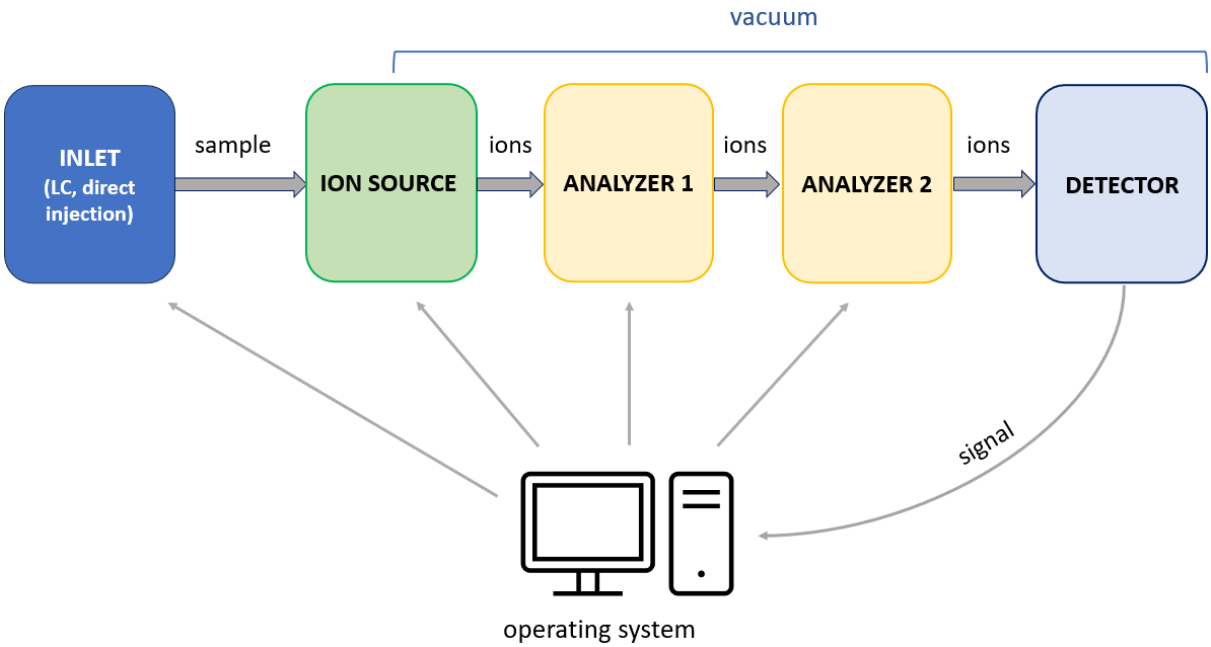


Fig.11. Schematic arrangement of a mass spectrometer with 2 mass analyzers. A sample enters the system through a sample inlet, which can be a direct probe or LC system (liquid chromatography); the sample is further ionized in an ion source and enters the high vacuum area; the produced ions are subsequently injected into series of mass analyzers, which can function as a selective filter for a certain subset of ions; the analyzed ions are registered at the detector, the acquired signal is then transformed into mass spectra by the operating system and proper data analysis software. The operating computer is able to send feedback to the individual mass spectrometer modules, modifying and adjusting the measurement in real time. Adapted from *Hoffmann and Stroobant 2007*.

The real MS data can be compared to theoretical fragmentation/digestion spectra of a protein, which ensures a proper identification and quantification in proteomics experiments (*reviewed by Wysocki et al. 2005*). Rather than discussing all possible techniques and setups in protein MS, I would like to focus on how the information about native protein structure can be obtained from mass spectra and which advantages the structural MS approaches possess (**Fig.12)**.

Native mass spectrometry and ion mobility

In this case, the whole protein, either monomeric or in a quaternary complex, retains its biologically related conformation while entering the gas-phase. The main advantages of this approach are very high sensitivity (analysis of picomole amounts, which is suitable for

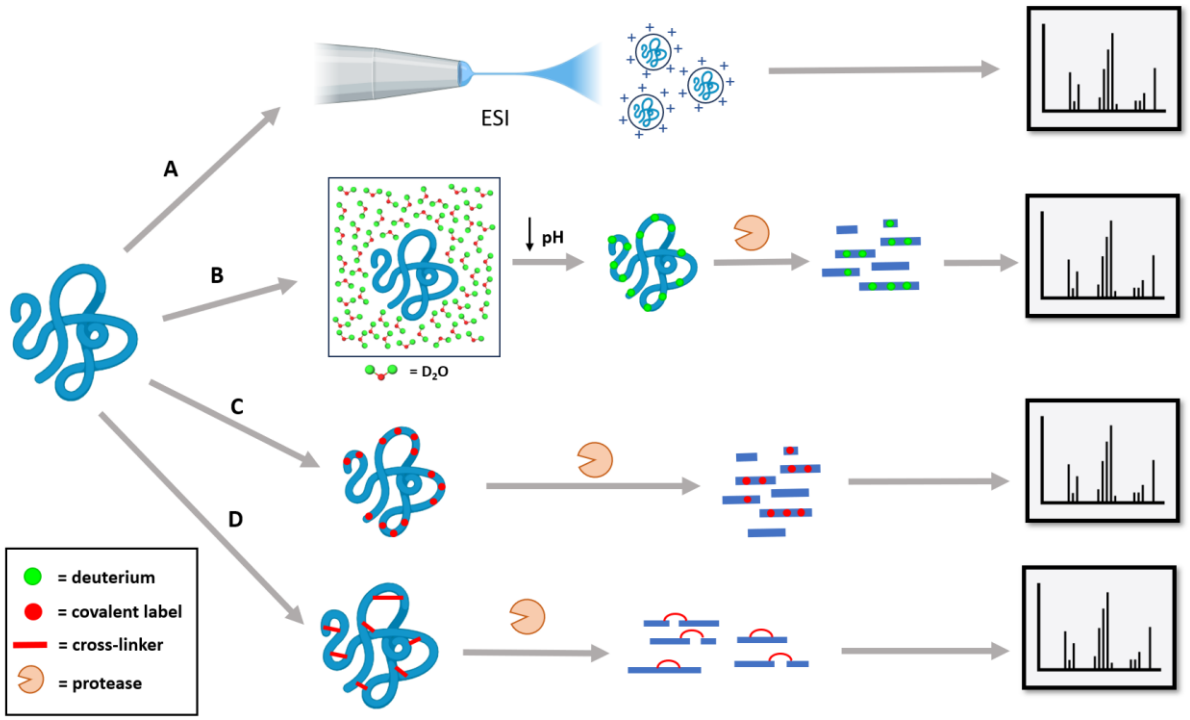


Fig.12. Structural mass spectrometry approaches. **A.** Native MS. **B.** HDX. **C.** Covalent labeling. **D.** XL-MS. Created in Biorender®. Based on *Vandermarliere et al. 2016.*

endogenously expressed proteins), robustness, high speed and selectivity. Despite these advantages, it provides less detailed information compared to standard structural methods, i.e., X-ray crystallography or NMR. Another drawback is the necessity of the purification of a protein/complex of interest, as well as using the specific buffer conditions for effective desolvation and ionization (*reviewed by Heck 2008*). Since masses of intact proteins are way bigger than peptide masses, the mass spectrometer suitable for native MS has to provide a wide m/z range and be capable of soft ionization of protein complexes while retaining the non-covalent interactions. It became possible with the introduction of electrospray ionization (ESI) and time-of-flight (TOF) or hybrid quadrupole-time-of-flight (Q-TOF) mass analyzers (*Heuvel and Heck 2004*). The upper limit of a protein complex size is theoretically unlimited - for instance, bacteriophage HK97 virion of 18 MDa was studied with native MS (*Snijder et al. 2013*). Mass spectrometers are nowadays often coupled with *ion mobility* (IM) - a technique that allows to investigate compactness of a desired protein in the gas phase. In the simplest setup, drift-tube ion mobility (DT-IM), an ionized protein traverses a linear cell filled with an inert gas (e.g., nitrogen) serving as an "obstacle" for the moving protein ions; the ion movement itself is ensured by a linear electric field. The various proteins and their complexes in different conformational states are progressing in such ion mobility cells with different rates - it depends on their compactness: the more the studied protein packed, the more easily it moves through the inert gas cloud. These IM-measurements allow to define a particle's collision cross-section (CCS) and separate different conformational species, which can further be analyzed by MS.

This and other possible setups of native MS coupled to ion mobility were described by *Konijnenberg et al. 2013.*

Hydrogen-deuterium exchange (HDX)

This approach takes advantage of the fact that hydrogen atoms present on proteins in H-N, H-O and H-S bonds can exchange with H-atoms of solvent, which is typically water. Placing a protein into heavy water, i.e., D_2O or deuterium oxide, leads to the exchange of the aforementioned H-atoms on a protein for D-atoms, which ultimately increases the mass of protein (+1 Da per such exchange event). The less protein H-atoms are exposed to a solvent and more buried into protein spatial structure, e.g., hidden in a protein hydrophobic core, the slower is the process of exchange; therefore, a protein conformation or its changes can be probed in a very detailed way by HDX-MS (*Hvidt and Linderstrøm-Lang 1955; Konermann et al. 2011*).

Among the bonds of the biggest interest for HDX are H-N on the amide backbone of proteins. In secondary structure, backbone amide H-atoms are often involved in hydrogen-bonding and thus become less accessible for the exchange. Another factor affecting the exchange rate is solvent accessibility, which was mentioned earlier. The choice of H-N as a target for HDX is its even distribution across a polypeptide chain (each peptide bond) and the fact that sidechain H-atoms capable of HDX have too high rate of exchange when exposed to a solvent, which makes it extremely difficult to measure using a standard HDX pipeline (*Englander et al. 1985*). This pipeline may vary substantially, depending on the aim of an experiment, e.g., studying one or more conformations present in a solution. The general protocol includes: 1) incubation of a target protein in D_2O-solution, which leads to HDX process; 2) quenching of HDX reaction by acidifying the solution and decreasing the temperature; this step can be performed at the different time points, if the conformational changes in time are investigated; 3) proteolytic digestion with low-pH tolerant protease (optional); 4) LC separation and ESI-MS measurement; 5) data analysis. For time-dependent measurements, the MS data are then used to create a plot depicting the percentage of deuteration against the time, which is interpreted as a rate of conformational changes of a studied protein, which makes HDX a perfect complementary technique for studying protein structural dynamics (*Masson et al. 2019*).

Covalent labeling

Even though the term "covalent labeling" refers to any covalent modification implemented on certain amino acid residues in a protein, which also includes the aforementioned HDX, I would like to use this definition to indicate *irreversible* covalent modifications. The irreversibility of amino acid chemical modifications ensures their stability and robust detection in mass spectra; moreover, such labeling is usually performed on amino acid side chains and therefore may provide a different type of structural information unavailable from HDX data. On the other hand, these modifications alter the biophysical and biochemical properties of an amino acid and thus affect the native protein structure (*Mendoza and Vachet 2009*).

The untargeted labeling approach utilizes the similar principle as HDX - the modification emerges on the solvent-accessible residues, whereas the residues involved in contact with a binding partner or buried in protein structure are typically less modified. Therefore, coupled to MS, it can be used as a powerful tool to investigate the protein interactions and conformational changes (*Guan and Chance 2006*). The untargeted labeling is mostly implemented as oxidative radical footprinting, the technique of labeling using the oxygen radicals. For example, hydroxyl radicals can be derived by the X-rays or γ-rays irradiation of water or by using the reaction of Fe(II)-EDTA complex with hydrogen peroxide; the emerged ·OH particles react with protein side chains, followed by an enzymatic digestion and LC-MS or LC-MS/MS (*Guan and Chance 2005*). Similar method, fast photochemical oxidation of proteins (FPOP), attempts to avoid the drawbacks of hydroxyl radical labeling, i.e., the longer reaction time and exposure of a protein to radicals leading to the false positive results and modification-induced unfolding. In FPOP setup, the protein solution with H_2O_2 is passed through a capillary, where in the certain region it is irradiated by a short laser pulse (nanoseconds). The pulsed laser beam induces the hydrogen peroxide homolytic cleavage and thus generation of radicals, which react with the accessible residues within a sub-millisecond interval. FPOP therefore allows to capture fast conformational changes and binding events (*Zhang et al. 2018*).

The targeted covalent labeling typically carried out on the specific amino acid side chain groups or free terminal carboxyl and amino groups, e.g., -SH (Cys), -OH (Ser, Thr, Tyr), -COOH (Asp, Glu, C-terminus), $-NH_2$ (Lys, N-terminus), $-CONH_2$ (Gln, Asn) and aromatic ring (Phe, Tyr, Trp, His). The description of all existing labeling reagents is out of scope of this work; they are usually classified according to their specificity to a certain side chain. For instance, cysteine thiol group is a strong nucleophile and is targeted by the nucleophilic agents such as iodoacetamide and iodoacetate or by maleimide derivatives (e.g., N-ethylmaleimide) (*Smythe*

1936; Gregory 1955). Other example, the aromatic ring, is susceptible to fluoroalkyl radical reaction with the Togni reagents, which is induced by ascorbate, ot by light in case of tryptophan (*Rahimidashaghoul et al. 2021; Fojtík et al. 2021*). Finally, the lysine ε-amino group or N-terminal amino group undergoes the reactions with N-hydroxysuccinimide esters (NHS esters) (*Blumberg and Vallee 1975*).

Cross-linking mass spectrometry (XL-MS)

Among targeted covalent labeling methods is the chemical cross-linking (XL) — the technique allowing covalently link the residues situated in the vicinity in a native protein structure. Typically, a cross-linker is an organic compound bearing 2 functional groups joined by a spacer and is capable of specifical reaction with chosen residue side chains. The cross-linkers of various specificity were developed in recent decades and are being developed nowadays: e.g., amino-specific for lysine and N-terminus and thiol-specific for cysteine. Both functional cross-linker groups can be the same in case of *homobifunctional cross-linkers,* which therefore react with 2 similar side chains (e.g., Lys-Lys, Cys-Cys); on the contrary, *heterobifunctional cross-linkers* are designed to covalently attach different side chains, e.g. Cys-Lys etc.. The spacer length (varying for different compounds) restricts the maximal distance of 2 neighboring residues in a protein, "fixing" their mutual position; the information obtained from the cross-linked residues is of a high value for building a structural protein model and analyzing protein-protein interactions in a complex. In addition, spacer can be uncleavable or uncleavable by MS/MS fragmentation (*Sinz 2003*). Some of the widespread cross-linkers of variable length and the corresponding reactivity with protein side chains are depicted in **Fig.13.**

As well as for HDX and covalent labeling described earlier, the same principle is valid: the cross-linker would react with the solvent-accessible side chains not completely buried within the hydrophobic protein core: it is perfectly demonstrated in the XL-MS study of αB-crystallin, where a fraction of lysins within the protein weren't cross-linked by DTSSP reagent (*Peterson et al. 2004*). In addition, this study exemplifies a standard XL-MS experiment pipeline: 1) optimizing of cross-linker:protein molar ratio, buffer composition and

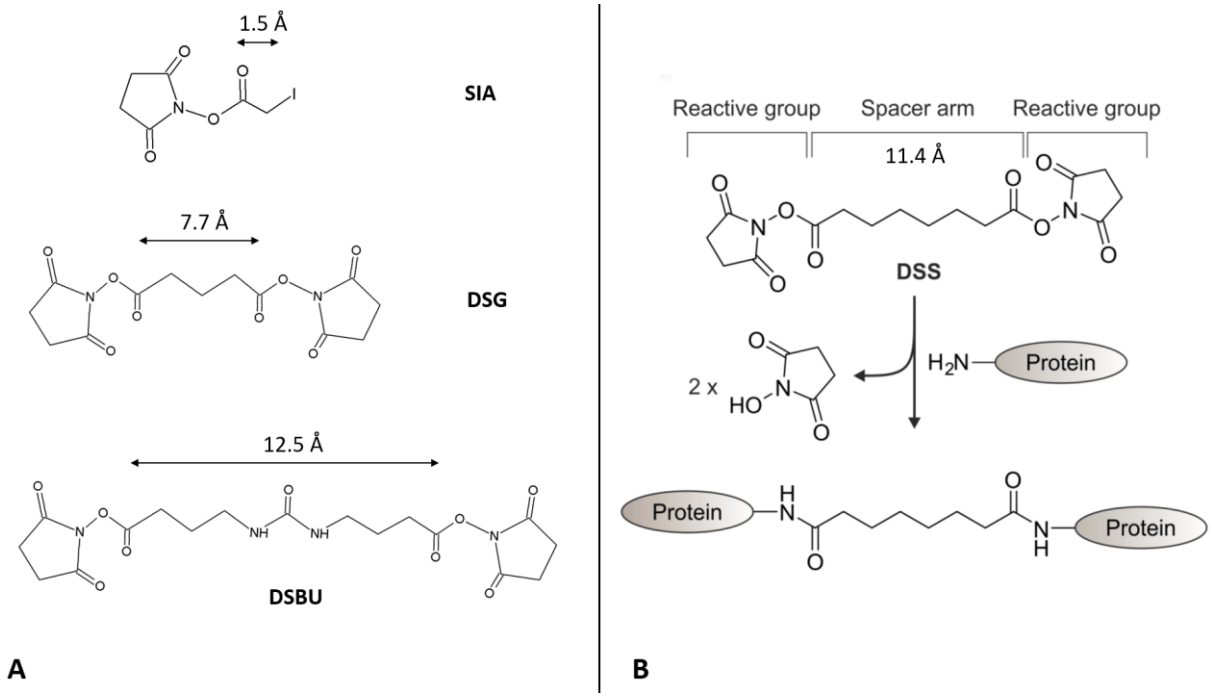
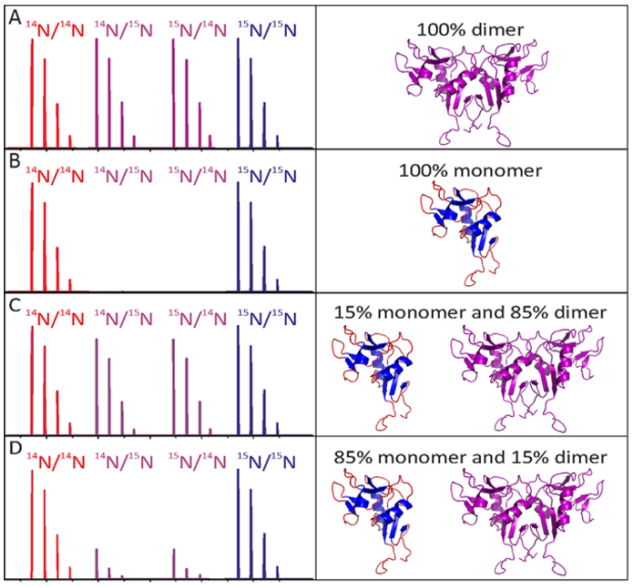


Fig.13. A. Examples of chemical cross-linkers. The spacer length represents the distance between side-chain functional groups of the covalently joined amino acids. **SIA -** succinimidyl iodoacetate, a bifunctional cross-linker (Lys-Cys); **DSG -** disuccinimidyl glutarate (Lys-Lys); **DSBU** - disuccinimidyl dibutyric urea, a cleavable cross-linker (Lys-Lys);. **B.** The NHS (Lys-Lys) cross-linkers reaction with amino groups on the example of **DSS (**disuccinimidyl suberate). The nucleophilic attack of amino groups leads to formation of amide bonds between the lysine side chains and the cross-linker (on both sides) and dissociation of N-hydroxysuccinimide. Based on *Zybailov 2013 and Liu et al. 2020.*

other XL-reaction conditions; 2) performing the XL-reaction; 3) an optional fractionation step to separate the cross-linked species, e.g., with the help of HPLC or SDS-PAGE; 4) proteolytic digestion; 5) analysis of the obtained XL-peptide mixture by MS (LC-MS, LC-MS/MS, MALDI-TOF etc.) The individual experiment design is dependent on the research goals and which type of structural information is sought: to probe the conformational state, mainly intra-cross-links (occurring within 1 polypeptide chain) are analyzed, therefore monomeric species are selected during the fractionation step; on the other hand, if the goal is to determine the interaction interface between several proteins, inter-cross-links (between 2 polypeptide chains) are favored and analyzed (*Sinz 2003*). In the second case, the data analysis of emerging XL-species is complicated: it is impossible to distinguish between intra- and inter-cross-links. To solve this problem, the heavy isotope labeling can be utilized (^{15}N, ^{13}C etc.): one of the subunits from the studied oligomer is produced recombinantly on a the heavy isotope containing medium (e.g., $^{15}NH_4Cl$ or ^{13}C-glucose for prototrophic producing strain, ^{15}N-amino acids for auxotrophic strains). The heavy atoms are therefore incorporated *in vivo* into the produced protein; the following incubation procedure prior to XL includes mixing the studied proteins in "light" (e.g., ^{14}N) and "heavy" (e.g., ^{15}N) forms in the defined molar ratio, usually 1:1. The

acquired mass spectra, after the general XL-protocol described above, contain triplets or quadruplets of MS peaks, which correspond to the possible inter-XL-peptide forms - $^{14}N/^{14}N$ (light-light, LL), $^{14}N/^{15}N$ (light-heavy, LH), $^{15}N/^{14}N$ (heavy-light, HL) and $^{15}N/^{15}N$ (heavy-heavy, HH). Commonly, in this type of study, LH and HL-peptides are of the biggest interest, because they originate from a dimer. The signal intensities of the individual XL-peptide forms can be used for the calculation of inter/intra ratio according to the peak intensities (**Fig.14**), which was already implemented in some XL-MS data analysis softwares, such as LinX (*Taverner et al. 2002; Kukačka et al. 2021*).

The XL-MS approach proved to be a powerful tool of structural biology among other major techniques like NMR, cryoEM or X-ray diffraction. Very frequently, it's difficult to obtain a protein crystal, especially for huge protein complexes; sometimes it's possible to crystallize and analyze individual domains and subunits, however the structural information of the whole protein/protein complex is absent. Therefore, XL-MS is particularly useful in solving structures of oligomeric assemblies, filling these informational "gaps": it allowed to elucidate structures of 14-meric RNA polymerase I and 12-meric RNA polymerase II with and without bound TFIIF transcription initiation factor, 26S and 19S yeast proteasome complex, 40S-eIF1-eIF3 translation initiation complex, even the whole prokaryotic ribosome (*Jennebach et al. 2012; Chen et al. 2010; Kao et al. 2012; Lasker et al. 2012; Erzberger et al. 2014; Lauber and Reilly 2011*). In many of the aforementioned studies, the previously obtained structural models of the individual subunits were used to complete the model of the whole assembly: the known distance restrictions imposed by a cross-linker spacer length to guide *in silico* docking. Moreover, in case of single protein conformational studies, XL-MS data facilitate structure prediction (*Orbán-Németh et al. 2018*). Another possible way to implement XL-MS data is to validate the docking outcome by the known distance restraints (*Van Dijk et al. 2005*). These and many other applications of XL-MS in structural biology and proteomics are reviewed in *Leitner et al. 2020.*



$$\text{inter/intra ratio} = \frac{I_{1415}+I_{1514}}{I_{1414}+I_{1515}} = \frac{I_{\mathrm{LH}}+I_{\mathrm{HL}}}{I_{\mathrm{LL}}+I_{\mathrm{HH}}}$$

Fig.14. The theoretical mass spectra from an XL-MS experiment with the equimolar mixture of ^{14}N- and ^{15}N- labeled subunits of a putative dimer. The ratio is calculated as a sum of LH and HL intensities divided by the sum of LL and HL intensities. Based on this ratio, the relative occurrence of an identified cross-link in the dimer-monomer equilibrium can be concluded. **A.** The ratio is approximately 1, the cross-link is 100% from a dimer; **B.** The LH and HL XL-peptides weren't identified at all, the cross-link is 100% from a monomer; **C.** The intensities of LH and HL are a bit lower than in case of LL and HH forms, the ratio is 0.85 (85% comes from a dimer); **D.** Similarly, the ratio is 0.15 (15% comes from a dimer). Based on *Kukačka et al. 2021*.

Aims of the work

1. Overexpression and purification of the recombinant HSC70 WT and the chosen mutants in ^{14}N and ^{15}N forms;
2. Study of the ATP-dependent HSC70 dimerization and optimizing the time point for the XL-reaction by mass photometry;
3. Collection of the dimer subunit distance restraints for HSC70 WT and its mutants by structural mass spectrometry;
4. Building the HSC70 WT dimer model by the protein-protein docking *in silico* on the basis of identified XL-sites.

Methods

Used chemicals and reagents

- acetic acid (Fluka, Switzerland)
- acetonitrile (ACN) (Sigma-Aldrich, USA)
- acrylamide (Sigma-Aldrich, USA)
- adenosine 5'- triphosphate disodium salt hydrate (ATP) (Sigma-Aldrich, USA)
- agar (Oxoid, USA)
- ammonium chloride (Lach-Ner, Czech Republic)
- ammonium-^{15}N chloride (Sigma-Aldrich, USA)
- ammonium persulfate (APS) (Sigma-Aldrich, USA)
- aprotinin (Roche, Switzerland)
- BME vitamins 100x solution (Sigma-Aldrich, USA)
- bromophenol blue (Serva, Germany)
- α-cyano-4-hydroxycinnamic acid (Bruker Daltonics, Germany)
- Coomassie Brilliant Blue R-250 (VWR International, USA)
- calcium chloride (Fluka, Switzerland)
- 2-chloroacetamide (Sigma-Aldrich, USA)
- ethanol (Lach-Ner, Czech Republic)
- 4-ethylmorpholine (Sigma-Aldrich, USA)
- deionized water, LC-MS grade (LC-MS dH_2O) (Merck, Germany)
- deionized water provided in house; sterilized by autoclaving, filtered with filter paper and degassed if stated
- dimethylsulfoxide (DMSO) (Fluka, Switzerland)
- disodium hydrogen phosphate heptahydrate (Thermo Fisher Scientific, USA)
- dithiothreitol (Thermo Fisher Scientific, USA)
- D-glucose monohydrate (Lach-Ner, Czech Republic)
- DNAse I (Thermo Fisher Scientific, USA)
- glycerol (VWR International, USA)
- hydrochloric acid 37% (Sigma-Aldrich, USA)
- 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (Carl Roth, Germany)

- immersion oil Immersol 518 F (Zeiss, Germany)
- imidazole (Fluka, Switzerland)
- iron(III) chloride (Sigma-Aldrich, USA)
- isopropanol (Lach-Ner, Czech Republic)
- isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich, USA)
- kanamycin monosulfate (Sigma-Aldrich, USA)
- leupeptin (Sigma-Aldrich, USA)
- magnesium chloride hexahydrate (Lach-Ner, Czech Republic)
- magnesium sulfate (Merck, Germany)
- methanol (Lach-Ner, Czech Republic)
- *N,N'*-methylenebisacrylamide (Sigma-Aldrich, USA)
- Peptide Calibration Standard II (Bruker Daltonics, USA)
- pepstatin A (Sigma-Aldrich, USA)
- phenylmethylsulfonyl fluoride (PMSF) (Serva, Germany)
- Pierce™ Unstained Protein MW Marker (Thermo Fisher Scientific, USA)
- potassium chloride (Lach-Ner, Czech Republic)
- potassium dihydrogen phosphate (Carl Roth, Germany)
- sodium chloride (Carl Roth, Germany)
- sodium dodecyl sulfate (SDS) (BioRad, USA)
- sodium hydroxide (Lach-NEr, Czech Republic)
- SUMO protease Ulp1 with His_6 tag from *S. cerevisiae,* recombinant (Sigma-Aldrich, USA)
- tetramethylethylenediamine (TEMED) (Sigma-Aldrich, USA)
- Trace Metal Mix A5 with Co 1000x (Merck, Germany)
- trifluoroacetic acid (TFA) (Sigma-Aldrich, USA)
- tris (VWR, USA)
- tris(2-carboxyethyl)phosphine (TCEP) (Sigma-Aldrich, USA)
- Trypsin/Lys-C Mix, Mass Spec Grade (Promega, USA)
- tryptone (Oxoid, USA)
- yeast extract (Oxoid, USA)

Used solutions, buffers and media

- acrylamide mix for SDS-PAGE (30%)
    - 1% (w/v) *N,N'*-methylenebisacrylamide, 29% (w/v) acrylamide, dH_2O
- 10 % APS (v/w) solution
- ATP washing buffer for IMAC
    - Tris lysis buffer + 5 mM ATP
- cleavage buffer
    - 90% (v/v) LC-MS dH_2O, 150 mM 4-ethylmorpholine, 10% (v/v) ACN, pH = 8.35 (calibrated with acetic acid)
- destaining solution for SDS-PAGE gels
    - 10% (v/v) acetic acid, 35% (v/v) ethanol, dH_2O
- elution buffer for IMAC
    - Tris lysis buffer + 300 mM imidazole
- 20% (v/v) ethanol solution for chromatography column storage
    - filtered with filter paper, degassed
- high salt washing buffer for IMAC
    - 50 mM Tris-HCl (pH = 7.9) 50 mM, 1 M NaCl, 5 mM $MgCl_2$, 10% (v/v) glycerol, sterile dH_2O
- 1M HEPES/NaOH stock solution, pH = 7.5 (calibrated with NaOH, filtered with filter paper)
- HKM150 buffer for XL reaction
    - 25 mM HEPES/NaOH (pH = 7.5), 150 mM KCl, 5 mM $MgCl_2$, LC-MS dH_2O
- HKMG150 buffer
    - 25 mM HEPES/NaOH (pH = 7.5), 150 mM KCl, 5 mM $MgCl_2$, 10% (v/v) glycerol, sterile dH_2O
- 1000x IPTG stock solution (1 M, filtered with 0.22 μm filter)
- 1000x Kanamycin stock solution (50 mg/ml, filtered with 0.22 μm filter)
- LB medium (sterilized by autoclaving)
    - 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, sterile dH_2O, pH = 7.4 (calibrated with NaOH)

- LB+agar medium (sterilized by autoclaving)
    - LB medium with 1.25% (w/v) agar
- M9-glucose minimal medium (either ^{14}N or ^{15}N)
    - 1x M9 salts (either ^{14}N or ^{15}N) , 0.4% (w/v) D-glucose, 2 mM $MgSO_4$, 0.1 m M $CaCl_2$, 0.1 mM $FeCl_3$, 1x BME vitamins, 1x Trace Metal Mix A5 with Co, 50 μg/ml kanamycin, bacterial overnight culture (1000x diluted), sterile dH_2O
    - the stock solutions of $MgSO_4$, $CaCl_2$, $FeCl_3$ and D-glucose were filtered with filter paper, without autoclaving
- 5x M9 salts (either ^{14}N or ^{15}N) (filtered with filter paper and sterilized by autoclaving)
    - 6.4 % (w/v) $Na_2HPO_4 \cdot 7H_2O$, 1.5 % (w/v) KH_2PO_4, 0.25 % (w/v) NaCl, 0.5 % (w/v) $^{14}NH_4Cl$ or $^{15}NH_4Cl$, sterile dH_2O
- 10% (w/v) SDS solution
- 5x SDS-PAGE sample buffer (loading dye)
    - 4% (w/v) SDS, 250 mM Tris-HCl (pH = 6.8), 0.06 % (w/v) bromophenol blue, 0.077 % (w/v) DTT, 30% (v/v) glycerol, dH_2O
- staining solution for SDS-PAGE gels
    - 0.25 % (w/v) Coomassie Brilliant Blue R-250, 45 % (v/v) methanol, 10 % (v/v) acetic acid, dH_2O
- 1 M Tris-HCl stock solution, pH = 7.9 (calibrated with HCl, filtered with filter paper)
- 1 M Tris-HCl stock solution, pH = 6.8 (calibrated with HCl, filtered with filter paper)
- 1.5 M Tris-HCl stock solution, pH = 8.8
  (calibrated with HCl, filtered with filter paper)
- Tris lysis buffer
    - 50 mM Tris-HCl (pH = 7.9), 300 mM NaCl, 5 mM $MgCl_2$, 10% (v/v) glycerol, sterile dH_2O
- 10x Tris-glycine SDS-PAGE running buffer, commercial (Bio-Rad, USA)

Used bacterial strains

- *E.coli* TOP10 (Thermo Fisher Scientific, USA)
    - genotype: F⁻*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K λ⁻*rps*L(Str^R) *end*A1 *nup*G

- *E.coli* BL21 Rosetta gami 2 (Novagen, Germany)
    - genotype: Δ(*ara-leu*)*7697* Δ*lacX74* Δ*phoA Pvu*II *phoR araD139 ahpC galE galK rpsL* (DE3) F′[*lac*$^+$ *lacI*^{q} *pro*] *gor522*::Tn*10 trxB* pRARE2 (Cam^R, Str^R, Tet^R)

Used equipment and instruments

- Analytical scale AL54-IC (Mettler Toledo, Switzerland)
- Automatic pipettes and corresponding tips (Gilson, USA; Thermo Scientific, USA)
- Centrifuge 5920R, exchangeable rotors (Eppendorf, Germany)
    - fixed angle rotor
    - fixed angle rotor
- Centrifuge 5415R (Eppendorf, Germany)
- Centrifuge Avanti J-26 XP, exchangeable rotors (Beckman Coulter, USA)
    - fixed-angle rotor JLA-9.1000
- Centrifugal concentrators (Millipore, USA)
    - Amicon® Ultra Centrifugal Filter (15 ml, 30000 MWCO)
    - Amicon® Ultra Centrifugal Filter (0.5 ml, 30000 MWCO)
- Centrifugal tubes (Beckman Coulter, USA)
    - J-lite 1 l, PP
- Centrifugal falcon tubes - 15 ml, 50 ml (Corning, USA)
- ChemiDoc MP Gel Documentation System (Bio-Rad, USA)
- Cuvettes for spectrophotometer, 10mm light path (BRAND, Germany)
- Dialysis tubing cellulose membrane, 14 kDa (Sigma-Aldrich, USA)
- Elmasonic S30H Ultrasonic Unit (Elma, Germany)
- EmulsiFlex-C3 High Pressure Homogenizer (Avestin, Canada)
- FPLC system NGC Quest 10 (Bio-Rad, USA)
- Hamilton KF needle (ga22/51mm/pst3) (Hamilton, USA)

- Magnetic stirrer Variomag Maxi Direct (Thermo Fisher Scientific, USA)
- MALDI-TOF Autoflex Speed mass spectrometer (Bruker Daltonics, Germany)
- Mass photometer Refeyn Two^{MP} and Accurion vibration-isolation bench (Refeyn, UK)
- Membrane filters 0.22 μm (Millipore, USA)
- Microscopic glass slides (VWR International, USA)
- MP-250V Electrophoresis Power Supply (Cleaver Scientific, UK)
- MTP 384 target plate ground steel BC for MALDI-TOF (Bruker Daltonics, Germany)
- Ice maker machine Powericer XL (Klarstein, Germany)
- IMAC columns EconoFit Nuvia, Ni-Charged, 5 ml and 1 ml (Bio-Rad, USA)
- Incubation shaker Multitron Pro (Infors, Switzerland)
- Incubator Incucell (BMT, Czech Republic)
- pH electrode InLab Expert (AND, USA)
- pH meter Orion Star A111 (Thermo Fisher Scientific, USA)
- Pipette controller Accu-Jet (BRAND, Germany)
- Pipettes — 2 ml, 5 ml, 10 ml (VWR, USA)
- SEC column ENrich™ SEC 650 10 x 300, 24 ml (Bio-Rad, USA)
- Silicon 6-well gaskets for mass photometry (Refeyn, UK**)**
- SDS-PAGE kit (Bio-Rad, USA)
- 15T solariX FT-ICR mass spectrometer (Bruker Daltonics, Germany)
- Spectrophotometer DeNovix DS-11 FX+ (DeNovix, USA)
- Syringe plastic 2 ml (B.Braun, Germany)
- Table minicentrifuge Minispin (Eppendorf, Germany)
- Thermostat with shaker Thermomixer comfort (Eppendorf, Germany)
- Vortex (Scientifica, Italy)
- UHPLC system Agilent 1290 Infinity II (Agilent, USA)
    - desalting precolumn Luna Omega Polar C18 (5 µm, 100 Å, 0.3 × 30 mm) (PhenomeneX, USA)
    - analytical column Luna Omega Polar C18 (3 µm, 100 Å, 0.3 × 150 mm) (PhenomeneX, USA)
- Zeba™ Spin Desalting Columns, 7 kDa MWCO, 0.5 ml (Thermo Fisher Scientific, USA)

Used softwares

- for mass photometry measurements and data analysis:
    - AcquireMP (Refeyn, UK)
    - DiscoverMP (Refeyn, UK)
- for SDS-PAGE gel imaging:
    - ImageLab 6.1 (Bio-Rad, USA)
- for mass spectrometry measurements (Bruker Daltonics, USA):
    - Compass HyStar 4.1
    - ftmsControl 2.1.0
    - flexControl 3.4
- for MS data analysis:
    - Compass Data Analysis 4.4 (Bruker Daltonics, USA)
    - LinX 2.0 (*Kukačka et al. 2021*)
- for protein sequence analysis
    - GPMAW 12.2
- for protein structure analysis, XL mapping and visualization:
    - PyMOL2 (Schrödinger, USA)
    - UCSF ChimeraX (RVBI, USA)
    - XMAS package for ChimeraX (*Lagerwaard et al. 2022*)
    - xiVIEW (*Graham et al. 2019*)
- for FPLC control and chromatogram export:
    - ChromLab (Bio-Rad, USA)
- for Venn diagram generation:
    - Venn diagram web tool (Van de Peer Lab)
- for DeNovix DS-11 FX+ spectrophotometer control:
    - EasyApps™ (DeNovix, USA)
- for pairwise sequence alignment:
    - EMBOSS Needle web server (EMBL-EBI, UK)
- for homology modeling:
    - Modeller 10.4 (*Šali and Blundell 1993*)
- for protein-protein docking:
    - HADDOCK 2.4 web server (*Van Zundert et al. 2016; Honorato et al. 2021*)
    -

- for protein molar extinction coefficient calculation:
    - ProtParam tool (Expasy, Swiss Institute of Bioinformatics)

Proteins and plasmids

For the overexpression of the proteins in the bacterial strain, *E. coli* plasmid vector pCA528 was used. This vector carries *KanR* gene for kanamycin resistance, *lacI* for a cloned gene expression regulation, *ori* and the MCS (multicloning site). The MCS contains *lac* operator (regulation by *lacI*), T7 promoter and RBS (ribosome binding site) upstream a cloned gene, and SUMO-His_6-tag downstream a cloned gene for the IMAC affinity purification and the higher yield of the overexpressed protein (*Peroutka III et al. 2011*). The sequences of the human HSC70 WT, HSC70 V438F mutant (SBD-mutant) and HSC70 L392E,L392E mutant (linker mutant, allostery-deficient) were cloned into MCS of pCA528 and provided to us by Matthias Mayer research team (ZMBH, Heidelberg, Germany). The protein sequences are enlisted in **Appendix 1.**

Spectrophotometric measurements of protein concentration

Concentrations of all proteins were measured on the DeNovix DS-11 FX+ spectrophotometer at the wavelength 280 nm according to the Lambert-Beer law (equation below). The molar extinction coefficients were calculated in Expasy ProtParam web tool by using the protein sequences as an input (**Appendix 1**). Prior to measurements, the spectrophotometer was blanked to the corresponding buffer, in which a target protein is dissolved (e.g., HKMG150 or HKM).

$$A = C \cdot \varepsilon \cdot l$$

Bacterial transformation

For all used *E. coli* strains and plasmids, the same transformation protocol was used and the whole procedure was carried out in sterile conditions (next to the Bunsen burner or in the

laminar flow hood) and with the sterile tools/chemicals. The competent cell aliquots were prepared in advance, stored at -80 ºC and thawed prior to transformation on ice; LB and LB-agar were prepared in advance. 1 μl of a plasmid (70-90 ng/μl) was added to a competent cell aliquot (40 μl), properly mixed and left to incubate for 30 min on ice. Then, the aliquot was placed into the incubator for 45 s at 42 ºC; immediately after 500 μl of LB medium was added and properly mixed in the aliquot by pipetting. The aliquot was subsequently incubated in the Thermomixer at 37 ºC and 300 RPM for 60 min. The aliquots were spinned down on the centrifuge (1 min, 3500g), the supernatant is mostly discarded with only 40-50 μl left in the tube. With this volume, the bacterial pellet was resuspended and plated onto LB-agar-kanamycin Petri dishes with a sterile hockey-stick cell spreader. The ready Petri dishes were stored at 37 ºC upside-down overnight; the Petri dish slits were covered with a parafilm layer. For further storage, the Petri dishes with bacterial colonies were stored in a fridge at 4 ºC.

Bacterial culture growing and protein induction

After the transformation procedure and growing the BL21 Rosetta gami 2 strain on LB-agar plates with kanamycin (see section "*Bacterial transformation*"), one colony was picked up by a pipette tip and dissolved in 15 ml of sterile LB medium supplemented with kanamycin (50 μg/ml). The primary culture was incubated in the Multitron Pro shaker at 37ºC and 200 RPM overnight. The next day, 100 μl of the primary culture was passaged to 15 ml of M9 medium with kanamycin (50 μg/ml) and grown overnight at the same conditions. This procedure was repeated once again, to ensure the $^{14}N/^{15}N$ turnover in cells.

After all the passages (15 ml LB → 15 ml M9 → 15 ml M9), 600 μl of bacterial culture was dissolved in 600 ml of M9 medium and grown at 37ºC and 180 RPM until reaching the O.D. = 0.8 - 0.9, which took approx. 3-4 h. The O.D. was measured spectrophotometrically at $\boldsymbol{\lambda}$ = 600 nm in cuvettes (against the corresponding medium as a blank). Then, the temperature in the shaker was decreased to 25ºC, the culture was left for an additional 20-30 min. After, 600 μl of 1 M IPTG was added to the culture (final concentration in culture 1mM), which was grown overnight at 25ºC and 180 RPM. The following day, the bacterial pellets were harvested by centrifugation (20 min, 4 ºC, 14000g). The supernatant was sterilized and discarded, the pellets were collected manually by a spatula into 50 ml falcon tubes and stored at -80 ºC. To verify the induction, SDS-PAGE of bacterial culture samples before and after the induction was carried

out. The SDS-PAGE samples were prepared in the following fashion: 1) dilute the culture to O.D. = 0.3, 100 μl ; 2) spin down by centrifugation (5000g, 1 min); 3) discard the supernatant and resuspend the pellet in 20 μl of water/buffer/medium; 4) add 5 μl of 5x SDS loading dye, mix and heat at at 95 ^{o}C for 5 min.

Protein purification

All the proteins were produced with SUMO-His_6-tag on their C-terminus, which enabled the IMAC purification with chromatographic columns with Ni-NTA stationary phase. The tag was cleaved off in the following steps of the purification procedure.

1) **Pellet resuspension.** The pellet from -80 ^{o}C was thawed on ice and resuspended in 50 ml of Tris lysis buffer supplemented with protease inhibitors (10 μg/ml aprotinin, 10 μg/ml leupeptin, 8 μg/ml pepstatin, 1mM PMSF) and DNase I (0.2 U/ml). The suspension was left at 4 ^{o}C on the magnetic stirrer for 30-40 min until the pellet was completely dissolved.
2) **Cell lysis.** For softer lysis, the EmulsiFlex-C3 High Pressure Homogenizer was used. The sample container was washed (run through the homogenizer) 3x with cold dH_2O and 2x with the lysis buffer. After, the air regulator pressure was set to 2 bar and regulated during the procedure, so the pressure in the homogenizer was kept at 1000-1100 bar. The bacterial suspension was run through the homogenizer 5-6x. After the procedure, the sample container was washed 1x with warm tap water, 2x with cold dH_2O and 1x with 40% isopropanol. Finally, 20-30 ml of 40% isopropanol was poured into the sample tank (for the storage purposes).
3) **1st IMAC.** The first IMAC run was carried out on 5 ml EconoFit Nuvia (Ni-NTA) on the FPLC system NGC Quest 10. Protein fractions were detected by the UV detector at 280 nm.
    a) The system pumps were initially washed with 50 ml of degassed and filtered dH_2O and 50 ml of degassed lysis buffer.
    b) After the column connection, it was washed with 25 ml of degassed and filtered dH_2O (flow rate 0.5 ml/min), then with 25 ml of degassed lysis buffer (1.5 ml/min, constant for all the following steps); maximum column pressure was set to 1 MPa.

c) The sample (lysed bacterial suspension, approx. 50 ml) was manually loaded through the pump A inlet, flow-through was collected.

d) The column was washed with 35 ml of the degassed high salt buffer, the flow-through was collected.

e) The column was washed with 20 ml of the ATP buffer, the flow-through was collected.

f) The column was washed with 40 ml of the degassed lysis buffer.

g) The gradient elution with the elution buffer (lysis buffer + 300 mM imidazole) was performed by the preset method created in ChromLab operating software (**Tab.2**). The pump B was inserted in the elution buffer, the pump A - lysis buffer. The fraction volume was set to 2 ml. The collected fractions were analyzed by SDS-PAGE.

h) The column was washed with 25 ml of dH_2O and 15 ml of 20% ethanol (degassed, filtered).

Tab.2. The elution method for the 1st IMAC procedure.

Step	Total volume, ml	Step Description	Fraction of the flow via pump B, %	Phase Name
1	5	Isocratic flow	0	Equilibration
2	5	Zero Baseline (of the UV detector)	0	Equilibration
3	105	Gradient flow (fraction size 2 ml, start on intensity 50 mAU, end on intensity 50 mAU)	0-100	Elution
4	130	Isocratic flow	100	Column wash

4) Tag cleavage and dialysis. The SUMO-His_6-tag was cleaved off by Ulp1 (recombinant, SUMO-specific, His-tagged protease). The chosen IMAC fractions from the previous step were joined together (25-30 ml), mixed with the Ulp1 (0.25 mg/ml), placed into the 14 kDa cellulose membrane for the overnight dialysis at 4 ^oC on the magnetic stirrer. The HKMG150 buffer was used for the dialysis (2 l, cold).

5) 2nd IMAC (reverse). Prior to the 2nd IMAC run, the small aliquot for SDS-PAGE was taken. 5 ml EconoFit Nuvia (Ni-NTA) was used for this procedure to get rid of the Ulp1

protease and the cleaved SUMO-His_6-tag. The column was washed and equilibrated similarly to the 1st IMAC step (10 ml of degassed and filtered dH_2O with the flow rate 0.5 ml/min; 10 ml of HKMG150 buffer with the flow rate 1 ml/min). The protein solution was manually loaded into the pump A; the column was additionally washed with 2-4 ml of the HKMG150 buffer until the absorbance at 280 nm hadn't dropped to the noise level. Consequently, the column was washed with 10 ml of the elution buffer, the collected eluate was analyzed by SDS-PAGE, as well as the initial flow-through (the purified protein).

6) **Protein concentrating and SEC.** Since chromatographic techniques lead to a severe dilution of the sample, the concentrating step prior to the SEC run was carried out. For this step, 30 kDa MWCO Amicon centrifugal filters (15 ml) were used (6000g, 15-30 min). The concentration of the protein solution collected in the previous was measured spectrophotometrically; the concentrating time was adjusted accordingly. The ENrich™ SEC 650 10 x 300 column was applied for this step. For the SEC run, the loop injection was used; the injection volume was set to 2 ml, the protein concentration for 1 injection was 2 mg/ml (maximal column load is 5 mg of protein). The maximal pressure was set to 4 MPa, the flow-rate was constant during the run (1 ml/min). Prior to the sample injection, the column and the injection loop were washed with 48 ml of degassed and filtered dH_2O (flow rate 0.5 ml/min) and after with 48 ml of degassed HKMG150 supplemented with 1 mM TCEP (1 ml/min). The injection into the loop was carried out manually by the syringe injection (4 ml syringe, ???). The chromatographic method for SEC was created and adjusted in ChromLab operating software, both pump inlets (A and B) were in the degassed HKMG150 buffer (**Tab.3**). After the fraction elution, the column was washed with 48 ml of degassed and filtered dH_2O and consequently with 36 ml of 20% ethanol (degassed, filtered).
7) **SDS-PAGE of the collected SEC fractions.** Each fraction was analyzed by the standard SDS-PAGE protocol (see "*SDS-PAGE*" section above). 10 μl of each fraction was mixed with 2.5 μl of the SDS-PAGE sample buffer, heated and spinned down.
8) **Protein storage.** The chosen fractions were united, the concentration was spectrophotometrically measured and the protein mixture was aliquoted (1.5 ml) into eppendorf tubes. The ready aliquots were quickly frozen in liquid nitrogen and stored at -80 ^{o}C.

Tab.3. The elution method for the SEC procedure.

Step	Total volume, ml	Step Description	Phase Name
1	2	Isocratic flow	Equilibration
2	2	Zero Baseline (of the UV detector)	Equilibration
3	6	Load Inject Sample: 1) Inject Sample (4 ml) 2) Change Valve (Sample Inject Valve)	Sample application
4	46	Fraction collection (fraction size 1 ml; start on intensity 20 mAU, end on intensity 20 mAU)	Elution

SDS-PAGE

SDS-PAGE was used to verify the induction of protein expression in the production strain, at all steps of protein purification and after the XL reaction. The stacking gel (5%) and the resolving gel (10%) was prepared according to **Tab.1.** The stacking gel was prepared first and poured between the glasses up to approx. 2 cm below the comb level; immediately after, 100 μl of isopropanol was placed to avoid the bubbles and to align the gel level. After 20-30 min, isopropanol was removed and the resolving gel was poured to the top and the well comb was placed. The polymerized gel was placed into the SDS-PAGE container, which was filled with 1x Tris-glycine SDS-PAGE running buffer. The samples were prepared by mixing in the 1:5 (v/v) ratio with 5x SDS-PAGE sample buffer, heated in a thermostat at 95 ^{o}C for 5 min and quickly spinned down with a table centrifuge (2 min, 5000g). The 10-15 μl of a sample was placed into a gel well by pipetting; as the protein marker, Pierce™ Unstained Protein MW Marker was used. The electrophoretic separation was running for 45-60 min at the constant current (50-60 mA per 1 gel). Then, the gel was stained by the staining solution for 6-7 min and then placed into the destaining solution for several hours; during the destaining, the solution was exchanged after 1 h. The pictures of the gels were taken in ChemiDoc MP.

Tab.1. The SDS-PAGE gels composition.

Chemicals	Resolving gel (10%) - 20 ml	Stacking gel (5%) - 5 ml
dH_2O	7.9 ml	3.4 ml
30% acrylamide mix	6.7 ml	830 μl
1.5 M Tris-HCl, pH = 8.8	5 ml	—
1 M Tris-HCl, pH = 6.8	—	630 μl
10% SDS	200 μl	50 μl
10% APS	200 μl	50 μl
TEMED	15 μl	7 μl

Mass photometry

Mass photometry works on the basis of the interference between scattered and reflected light incident on a solution and on the basis of ratiometric imaging. The scattering pattern varies for the species of different molecular mass, which makes mass photometry a robust and simple tool to analyze the heterogeneity of large biomolecules, i.e., proteins or nucleic acids (*Sonn-Segev et al. 2020*). This technique was applied to indicate and estimate the HSC70 dimer in the molecular population. In order to verify the role of ATP in chaperone dimerization, the ATP-time-dependent mass photometry experiments were performed on the Refeyn Two^{MP} (placed on the Accurion vibration-isolation bench); the instrument was operated by the Refeyn AcquireMP software, the data were processed in Refeyn DiscoverMP software. Prior to the experiments, the objective was carefully cleaned by 40% isopropanol solution. The calibration was done on the BSA solution and IgG solution (final concentrations 10 nM). The instrument was calibrated each 2-3 h during the measurements.

All the HSC70 samples were thawed on ice prior to the measurements and concentrated with the Amicon UltraCentrifugal filter (30 kDa MWCO). The buffer was exchanged to HKM buffer by ZebaSpin 7 kDa MWCO columns. All the protein mixtures were diluted to 20 μM total final concentration; the 6 mixtures were prepared in total: only HSC70 WT, only SBD mutant, only linker mutant, 1:1 equimolar mix of WT+SBD mutant, 1:1 equimolar mix of

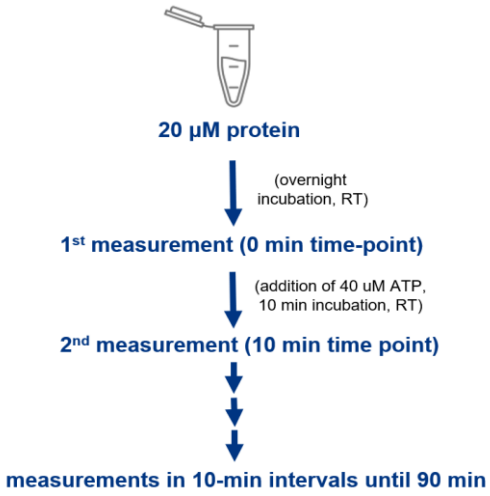


Fig.15. The HSC70 mass photometry experiment design. The initial mixture (20 μM of HSC70 in HKM buffer) was incubated overnight at room temperature (RT) and serially diluted to 10 nM prior to the measurement (0 min time-point). After, the ATP was added (final concentration 40 μM in 20 μM of HSC70), incubated at RT for 10 min and measured. Consequently, the measurements were done in 10 min intervals (20, 30, 40, 50, 60 min time-points), the last measurement was done at 90 min time-point, when the dimer level stabilized .

WT:linker mutant, 1:1 equimolar mix of SBD-mutant+linker-mutant. All the protein solutions were incubated at room temperature overnight in the HKM buffer to achieve the monomer/dimer equilibrium in a liquid phase. The experiment design is depicted in **Fig. 15.** It is practically possible to measure the 10-20 nM final protein concentration; higher concentrations ultimately lead to the molecular crowding effect and inability to distinguish between the particles of different molecular weight in the solution due to the noisy signal and decreased resolution. Thus, it was necessary to perform the serial dilution of a sample prior to the mass photometry measurement. During the data analysis, the dimer/monomer ratio was calculated by dividing the dimer signal intensity (or "counts") by the monomer signal intensity; then, the normalized dimer/monomer ratio values were plotted against time (done in Microsoft Office Excel).

Cross-linking and sample preparation for LC-MS

1. **Incubation.** ^{14}N- and ^{15}N-versions of HSC70 were thawed on ice, concentrated and desalted in the same way as was described in the section "*Mass photometry*" (see above). During the desalting/buffer exchange, the final buffer for HSC70 incubation was HKM supplemented with 1 mM TCEP (to avoid the emergence of the artificial dimers through the oxidation of cysteine thiol groups and formation of disulfide bonds). Then, the ^{14}N- and ^{15}N-proteins were mixed in 1:1 molar ratio and digested by trypsin/Lys-C (final

concentration 0.2 mg/ml) for 30 min at 37 ºC. The mixture of digested peptides was then analyzed by MALDI-TOF to adjust the real molar ratio of $^{14}N/^{15}N$-proteins for the final incubation. The matrix solution was prepared by mixing the equal volumes of acetonitrile, 0.1% TFA solution and saturated α-cyano-4-hydroxycinnamic acid solution in methanol. After, the sample solution (digested peptides) and the matrix solution were mixed together in 1:1 ratio (v/v); 1 μl of the new solution was pipetted onto the well of MTP 384 target plate ground steel BC and let to completely dry (10-15 min). In the same fashion, the standard sample (Peptide Calibration Standard II) was prepared and placed onto the MALDI plate. The samples were analyzed at the laser intensity 35-45%, the ionization spots on the well were picked up manually by the MALDI-TOF operating software (FlexControl 3.4). For the calibration and measurements, the preset method for the peptides of 700-4500 kDa molecular weight range was used. According to the ratio between the signal intensities of ^{4}N- and ^{15}N-peptides originating from the tryptically digested HSC70 mixture, the proper adjustments of the protein concentration for the final incubation were made. Unlabeled and labeled HSC70 were then mixed in 1:1 molar ratio (final total protein concentration 20 μM) in HKM buffer with 1 mM TCEP and incubated overnight at 25 ºC. Each sample was prepared in triplicate.

2. **Cross-linking.** For all XL experiments, the DSS cross-linker was used (see section "*Cross-linking mass spectrometry".* The final concentration of the cross-linker was calculated based on the number of lysine residues in the HSC70 WT sequence (54 Lys, therefore approximated molar ratio HSC70:DSS = 1:50). Prior to the XL reaction, 40 μM ATP (final concentration) was added to the chaperone solution and incubated for 90 min; this cross-linking time was determined by the mass photometry measurements described earlier. After 90 min incubation, DSS powder aliquote was initially dissolved in DMSO, added to HSC70 and carefully mixed by pipetting (DSS final concentration 1 mM). The cross-linking reaction was going for 60 min at room temperature.
3. **Reduction, carbamidomethylation and digestion.** At this step, it's important to get rid of all probable disulfide bonds and block the thiol groups, so in the cross-linked sample no possible cysteine oxidation and disulfide bond formation can occur. The reducing agent used is TCEP; for the carbamidomethylation, 2-chloroacetamide is used. The latter reagent reacts with thiol groups via nucleophilic attack, leaving $-CH_2CONH_2$ moiety covalently attached to the thiol sulfur atom. The reaction mixture content is listed in **Tab.4**. The addition of the basic cleavage buffer is necessary due to the following

digestion step with trypsin/Lys-C (pH optimum 7.8-8.7). The reduction-carbamidomethylation reaction was carried out in an incubator at 70 °C for 5 min. For the digestion, the final trypsin/Lys-C concentration 0.045 mg/ml was used (1 μg per sample). The digestion was carried out overnight at 37 °C in an incubator. The following morning, the digestion was quenched by the addition of TFA (final concentration 0.1% (v/v), to the final volume 40 μl.

Tab.4. The reduction/carbamidomethylation reaction composition.

Component	Volume added, μl
XL-sample (HSC70)	10
cleavage buffer	10
100 mM TCEP (in cleavage buffer)	2
1 M 2-chloroacetamide (in cleavage buffer)	0.8
Total volume	22.8

LC-MS measurements

For the chromatographic separation of the tryptically digested peptides and mass spectrometry analysis, the UHPLC system Agilent 1290 Infinity II connected to the inlet of 15T solariX FT-ICR was used. 5 μl of the sample was injected by the Agilent 1290 autosampler into the desalting precolumn and desalted at flow rate 20 μl/min for 5 min. Then, the desalted peptide mixture was injected into the analytical column separated at flow rate 10 μl/min and 50 °C in acetonitrile gradient (5-35%) for 35 min (maximal pressure 1000 bar). Then the column was washed with 95% acetonitrile for 5 min and equilibrated with 5% acetonitrile for 15 min, ready for the next injection. Mass spectra were acquired in positive mode over the m/z range 250–2500 with 10^6 data points transient and 0.2 s ion accumulation with two averaged scans per spectrum. The MS measurements were performed by the FT-ICR operating software, ftmsControl 2.1.0. As a standard, 1 pmol/μl BSA solution was used.

Data analysis of identified XL-peptides

The acquired MS-spectra were analyzed with Bruker Data Analysis 4.4 (LC-MS data export, manual validation of "raw" spectra) and LinX 2.0 softwares (XL-peptide identification, the spectra assignment to the theoretically possible cross-link library). LinX requires 3 inputs for the assignment operation (*Kukačka et al. 2021*) :

1. **Input 1** includes protein sequences, cleavage sites and possible residue modifications: the FASTA sequences of HSC70 WT, SBD-mutant and linker-mutant were used as input sequences; the cleavage sites were defined by trypsin/Lys-C cleavage specificity (C-terminal side of lysine and arginine) with 3 miss-cleavages allowed and the forbidden cleavage on modified sites (i.e., on the cross-linked Lys residues). The cross-linker, which has reacted with the residues and now covalently attaches 2 different amino acid residues, sterically hinders the docking of a protease on the corresponding cleavage site. In our experiments, the Lys-specific DSS cross-linker was used, which eventually blocks the action of trypsin/LysC on the DSS-cross-linked site. For modifications, variable oxidation of methionines (possible during the purification and sample preparation; +15.9949 Da) and fixed carbamidomethylation on cysteines (see "*Cross-linking and sample preparation for LC-MS"* section; +57.0215 Da) were defined.
2. **Input 2** requires the definition of a used cross-linker and cross-linkable residues. In our case, the DSS cross-linker was used; Lys and N-terminus were defined as cross-linkable residues.
3. The LC-MS data were exported in Data Analysis into .txt format, i.e., the suitable input for LinX, by the method script written in our laboratory. The .txt data files generated in this fashion served as an **input 3** for LinX. To improve the data analysis, the XL-peptides identified over the range of chromatographic scans were joined by the internal LinX tool PeakJoiner (mass tolerance - 2 ppm, maximum missing scans - 4).

After providing all the inputs, the assignment process was run with 1 ppm precision and no peptide length/mass filtering; the input 3 format was defined as MSe. The output was then checked for ^{15}N-labeling (all the settings were kept default). The resulting table includes: an XL-peptide sequence, the residues numbers from the original protein sequence, the possible modifications, bonds created by a cross-linker molecule, number of chromatographic scans where the given XL-peptide was identified (optimized by PeakJoiner tool), the real peptide

charge state, how many forms were found (among 4 theoretically possible - $^{14}N/^{14}N$, $^{14}N/^{15}N$, $^{15}N/^{14}N$, $^{15}N/^{15}N$) and the m/z values for all the forms in a mono charged and the m/z values of the actual corresponding charge state. If for a given XL-peptide all 4 forms were assigned by LinX, the inter/intra ratio is calculated as well according to the equation from **Fig. 13**. Due to possible false-negative hits, every relevant identification was manually validated in DataAnalysis software by checking the raw mass spectra for the presence of the corresponding m/z peaks. In **Fig.16,** the example of LinX identification (**Fig.16A)** and the mass spectra displayed in DataAnalysis (**Fig.16B**) are shown. For the majority of identification, the inter/intra ratio was calculated manually from the base peak intensities from XL-peptide form isotopic shells; for the given example, the ratio was calculated in **Fig.16C.**

It's noteworthy that not in each experiment the identifications with all 4 forms present were analyzed. In case of the mixed experiments (i.e., WT and SBD-mutant, WT and linker-mutant, SBD-mutant and linker-mutant) the emergence of a symmetrical homodimer isn't guaranteed. Taking into consideration dimer models, which *aren't symmetrical* (see **Fig.10**), it is possible that in some cases the given cross-link might appear only either in $^{14}N/^{15}N$ or $^{15}N/^{14}N$ orientation. Thus, for the aforementioned experiments, both the hits with 4 forms and 3 forms found were analyzed.

The relevant identifications were then exported to an Excel table, where the inter/intra ratio was calculated for every replicate. The ratios among replicates were then averaged and the standard deviation was calculated. Then, the Venn diagrams were generated to filter out the cross-links found in all 3 replicates. The refined data was later used to perform *in silico* docking with HADDOCK 2.4 server.

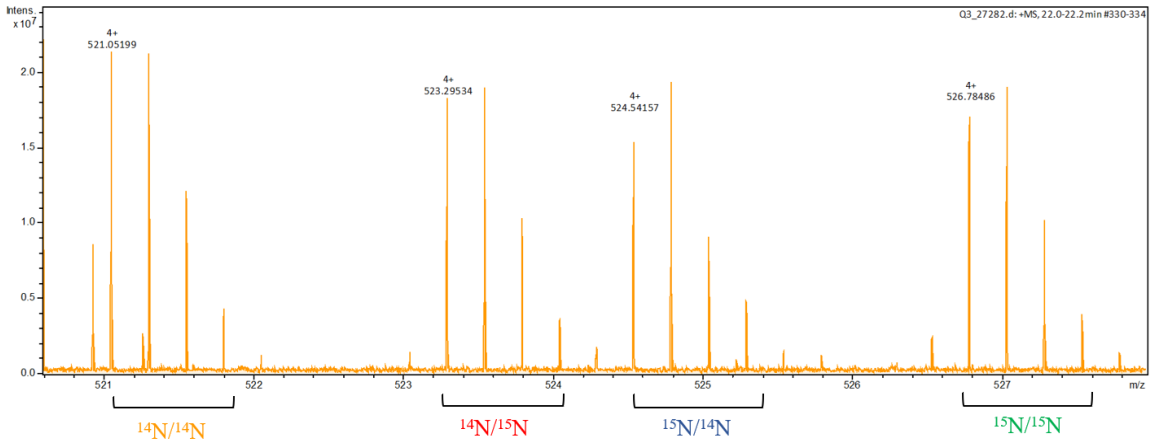
Homology modeling of the ADP-bound HSC70

Due to the absence of a 3D model of the ADP-bound human HSC70 in any available repository, it was decided to create a homology model based on the ADP-bound structure of the Hsp70 homologs. The procedure was carried out in the Modeller 10.4. Firstly, an appropriate template was found in the RCSB-PDB database. The most suitable structure including both NBD and SBD domains with a bound ATP molecule was the Hsp70 chaperone from *E. coli*, DnaK, the model of which was determined by the solution NMR (PDB code 2KHO, **Fig. 5B**).

A

Cross-link	Protein	Bonds	Charge	Scan	Forms found	LL m/z (1+)	LL m/z (X+)	LH m/z (1+)	LH m/z (X+)	HL m/z (1+)	HL m/z (X+)	HH m/z (1+)	HH m/z (X+)
IPKIQK - ATVEDEKLQGK	HSC70 WT (343, 348) [A] - HSC70 WT (551, 561) [B]	DSSd0 (A.345; B.557)	4	358-363	4	2081.185	521.0518	2095.144	524.5414	2090.159	523.2951	2104.117	526.7847

B



C

$$\text{inter/intra ratio} = \frac{I_{1415}+I_{1514}}{I_{1414}+I_{1515}} = \frac{I_{LH}+I_{HL}}{I_{LL}+I_{HH}} = 0.975$$ => comes from the dimer

Fig.16. The example of an identified XL-peptide. **A.** The output from LinX. The columns are commented above (section *"Data analysis of identified XL-peptides"*); LL - light-light form ($^{14}N/^{14}N$); LH - light-heavy form ($^{14}N/^{15}N$); HL - heavy-light form ($^{15}N/^{14}N$); HH - heavy-heavy for ($^{15}N/^{15}N$). **B.** The corresponding mass spectra of all 4 identified forms. The m/z values of each form are listed in the table above the spectrum. **C.** The example of the inter/intra ratio calculation for the identified cross-link: the sum of base peak intensities of the intermediate forms (i.e.,$^{14}N/^{15}N$ and $^{15}N/^{14}N$) is divided by the sum of base peak intensities of the completely light ($^{14}N/^{14}N$) and completely heavy form ($^{15}N/^{15}N$). This value provides a rough estimate of how often the given cross-link emerged between different homodimer subunits. In this case, the calculated value is very close to 1, which suggests that it is most probably of the dimer origin

The sequence pairwise alignment of human HSC70 and *E. coli*, DnaK was done in the EMBOSS Needle web server; the human HSC70 sequence was taken from the Uniprot database (entry P11142), the template DnaK sequence was taken from the RCSB-PDB structure web page. The output alignment file was adjusted to the .ali format suitable for Modeller, The models were generated by running the installed Modeller 10.4 script through the command line tool according to the developer's manual. Ten models were generated in total. The best model was chosen by the lowest superposition RMSD value: each model was superimposed to the DnaK template structure and evaluated in PyMOL.

Mapping of the identified cross-links onto the protein model

For this step, UCSF ChimeraX 1.5 and the XMAS ChimeraX package were applied. The ADP-HSC70 model was opened in ChimeraX twice to map the identified cross-links on a putative HSC70 dimer and analyze the mutual subunit orientation. The input file for XL-mapping, i.e., "the evidence file", was created manually in .csv format (an Excel table), similar to the output of the supported XL-MS search engines, e.g., XlinkX for Proteome Discoverer. The table consists of columns: 1) Checked (empty); 2) Cross-link sequence (a XL-peptide sequence divided by a vertical slash, e.g. IPKIQK|ATVEDEKLQGK); 3) Max. XlinkX Score (random numbers); 4) Crosslinker (e.g., DSS); 5) Crosslink Type (in our case, Interlink, between the subunits); 7) # CSMs (default 1); 8) # Proteins (empty), 9) Sequence A (first peptide from the XL-pair with square brackets at the XL-site, e.g., IP[K]IQK); 10,11,12) Modifications A; Accession A; Position A (3 empty columns); 13) Sequence B (second peptide from the XL-pair with square brackets at the XL-site, e.g., ATVEDE[K]LQGK); 14,15,16) Modifications B; Accession B; Position B (3 empty columns); 17,18) Protein description A; Protein description B (empty, can be filled at discretion); 19) Is Decoy (set to FALSE); 20) Q-value (set to 0); 21) merged (empty). The evidence file in .csv format, designed manually in the aforementioned fashion, was uploaded to ChimeraX to the field "Evidence files"; consequently, the pseudobond file (.pb) was generated by using the "Map cross-links" function. With the "Integrate" function, the distance restraints files in .tbl format were generated for HSC70 protein-protein docking in the HADDOCK 2.4 server. During this procedure, the chain A and chain B were selected (the 1st and the 2nd HSC70 molecules, respectively), as well as the corresponding pseudobond file. The distance restraints were set in the following manner (according to the properties of the used cross-linker - DSS): the minimum - 5 Å, the median – 24 Å, the maximum – 30 Å. These distances are between the Cα atoms of the cross-linked residues, rather than the cross-linker spacer arm length.

Additionally, the identified XL-sites were mapped on the HSC70 cartoon sequence by the xiVIEW webtool. As the inputs, HSC70 FASTA sequences and the identification file with the XL-sites in .csv format were used. The identification file contains the columns: 1) Type (in our case, Inter-protein); 2) Protein_1 (e.g., first HSC70 subunit, HSC70_1); 3) AbsPos1 (the number of a corresponding XL-site on the first subunit); 4) Protein_2 (e.g., second HSC70 subunit, HSC70_2); 5)AbsPos2 (the number of a corresponding XL-site on the second subunit).

The names of the proteins defined in “Protein_1” and “Protein_2” columns should be exactly the same as the used ones in the input FASTA files.

Docking and validation of the HSC70 WT dimer model

Due to the highly flexible nature of the ADP-bound HSC70 conformation (**Fig. 5B)**, where NBD and SBD can adopt various mutual positions enabled by the disordered linker, the specific approach for docking was chosen. In the developed docking workflow for multidomain flexible proteins, such proteins are treated as a certain number of individual subunits: in putative disordered or hinged regions, a multidomain protein is cut and the emerged subunits are docked as independent entities (*Karaca and Bonvin 2011*). For the human HSC70 WT, the separated domains were created manually by dividing NBD (residues from 1 to 381) and SBD (residues from 399 to 646) into 2 .pdb files and erasing the atom coordinates for the HSC70 linker (residues between the domains, i.e., 382-398). Since no XL-sites were identified in this linker region, it was ensured that no structural information was lost during the dimer docking; nonetheless, to include the necessary distance constraint, naturally provided by the linker, the XL-site detected between NBD and SBD was also used for the docking. This XL-site was therefore mimicking the linker distance constraint.

As a docking tool, the HADDOCK 2.4 web server was applied. The created .pdb files for NBD and SBD were uploaded as input structures; it’s worth noting that each domain was used twice and the according chain names were assigned (for the 1st subunit, the NBD was assigned as “Chain A”, the SBD - as “Chain B”; for the 2nd subunit, the NBD - as “Chain C”, the SBD - as chain “Chain D”). Thus, 4 polypeptide molecules were used as input data in total. In the “Input parameters”, all the settings were set default without defining active and passive residues (the first are directly involved in interactions; the second are less involved in interactions and not present in an interaction interface). In the “Docking parameters”, the previously generated distance constraint .tbl file was used as unambiguous restraints (see “*Mapping of the identified cross-links onto the protein model*” section). The center of mass restraints was defined to enforce contact between the molecules (the forced constant for surface contact restraints - 1). In addition, the changes were made in the scoring function - the energy of distance restraints for all iterations and water refinement (Eair1, Eair2, Eair3) were set to 1. The output dimer models were then validated manually in PyMOL2.

Results

Molar extinction coefficient calculation

The molar extinction coefficients of the studied proteins were calculated by Expasy ProtParam web tool as was stated previously (see "*Spectrophotometric measurements of protein concentration"*). The output molar weight and the coefficients (**Tab.5)** were then applied to the DeNovix DS-11 FX+ spectrophotometer for the correct protein concentration measurements. However, since the mutations haven't introduced any tryptophan residues, the value of the molar extinction coefficient is the same for all 3 proteins. For the ^{15}N-HSC70 proteins, the specific MW and coefficients weren't calculated, because the molar ratio adjustment was done later by MALDI-TOF analysis (see "*Cross-linking and sample preparation for LC-MS*" section).

Tab. 5. The calculated values of MW and molar extinction coefficients for the studied proteins.

Protein	MW, Da	Molar ext. coefficient, M^{-1} cm^{-1}
HSC70 WT	70898.09	33350
HSC70 SBD-mutant (V438F)	70946.14	33350
HSC70 linker-mutant (L392E, L394E)	70930.00	33350

Protein overexpression and purification

Before and after the addition of IPTG for the protein expression induction from the plasmid, the aliquots of bacterial culture were taken to verify the protein overexpression by SDS-PAGE (**Fig. 17**). The visible band of 70 kDa protein can be seen after induction, both for ^{14}N- and ^{15}N-labeled HSC70 (in the **Fig.17** only ^{15}N-HSC70 are shown).

For the first purification carried out (for HSC70 WT), the 1st IMAC fractions collected by the imidazole gradient elution were checked by SDS-PAGE: the very prominent bands corresponding to the uncleaved HSC70-SUMO-His_6 can be seen (**Fig.18B**). It's noteworthy

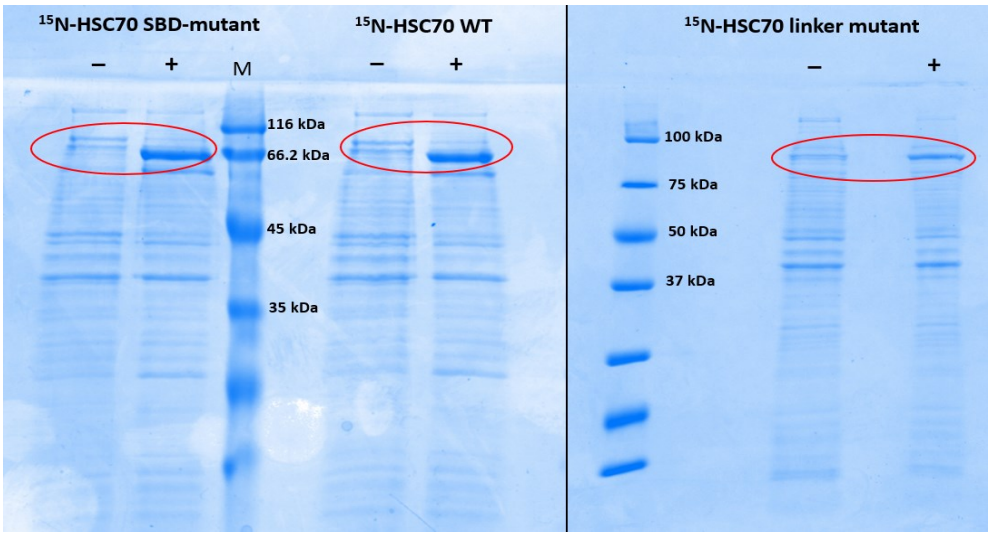


Fig.18. The SDS-PAGE gels for the ^{15}N-HSC70 overexpression. The effect of IPTG induction (- before induction, + after induction. The HSC70-SUMO-His$_6$ bands emerged after the addition of IPTG to the bacterial cultures. The results for all the overexpressed proteins were identical.

that much less intense bands just below the uncleaved protein are also visible; these bands are believed to be the cleaved HSC70, without the tag. The tag cleavage might have partially occurred naturally, within the expression strain cells. For the purifications of the remaining HSC70 proteins, the 1st IMAC fractions weren't checked by SDS-PAGE due to time saving. The band in 35-45 kDa range was present in all the fractions and has an unknown nature; probably, it can be a histidine-rich protein, which coeluted with the tagged chaperone.

The activity of the Ulp1 protease was verified on the small aliquot of HSC70-SUMO-His$_6$ (**Fig.19**). After the overnight incubation with 0.25 mg/ml Ulp1-His$_6$, the tagged HSC70 band shifted lower on the SDS-PAGE gel and the very low-intensity SUMO-His$_6$ band (approximately 12 kDa) appeared. In fact, SUMO-His$_6$ has a retarded electrophoretic mobility and is usually visible at 17-18 kDa level. In addition, the noticeable band corresponding to Ulp1-His$_6$ (27 kDa) obviously appeared after the cleavage. The same unknown band within the 35-45 kDa range was detected as well. For the following routine purification, only HSC70 samples before cleavage (B.C.) and after cleavage (A.C.) were verified by SDS-PAGE. The 2nd, or reverse, IMAC elution hasn't been performed with fraction collection: the flow-through was just collected into a 1 falcon tube. Therefore, only 1 sample from the overall 2nd IMAC eluate was analyzed by SDS-PAGE to see the removal of Ulp1-His$_6$ and SUMO-His$_6$, which have to be captured onto Ni-NTA resin and thus separated from the cleaved HSC70. It can be perfectly seen in **Fig.19B** (A.C. and 2.Ni lines). For the last chromatographic procedure, SEC, all the fractions collected were analyzed by SDS-PAGE for all the purification; the example of SEC chromatograms and the SDS-PAGE fraction analysis for ^{14}N- and ^{15}N-HSC70 WT is

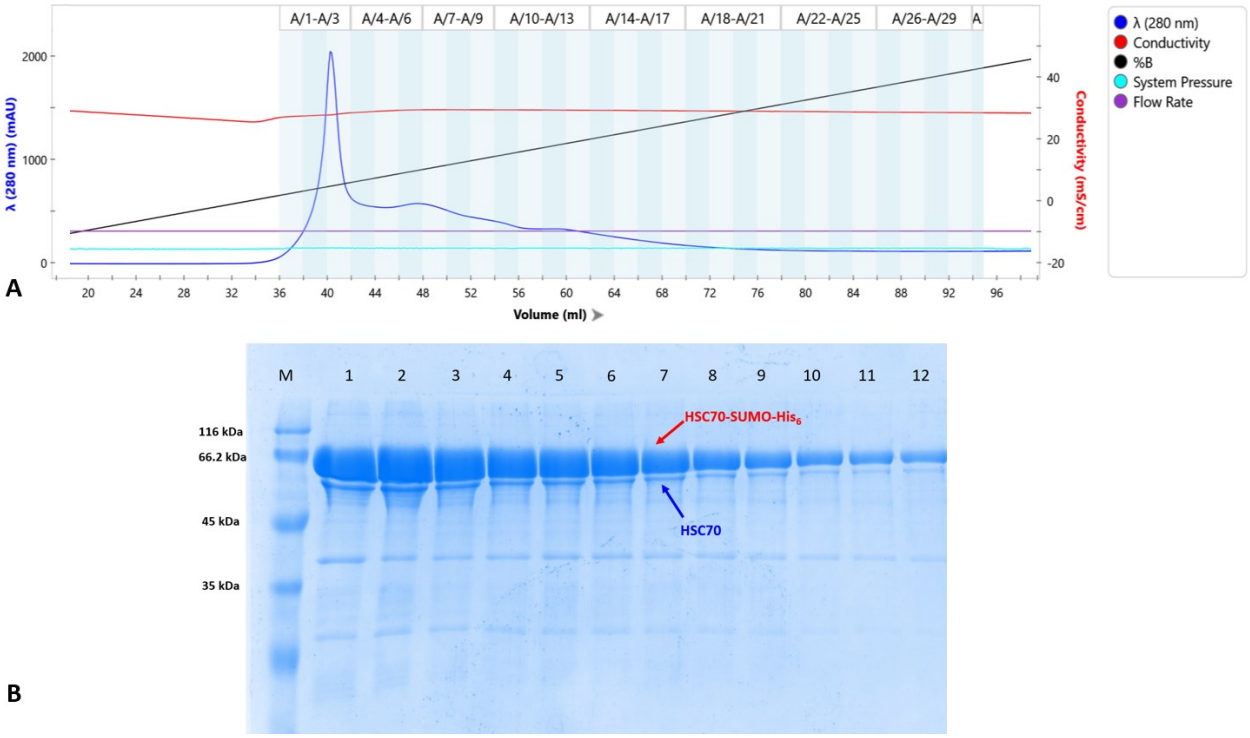


Fig.18. The chromatogram (**A**) and the fraction analysis by SDS-PAGE (**B**) for the 1st IMAC procedure during the ^{14}N-HSC70 WT purification. The numbers of the collected fractions from the chromatogram correspond to the numbers of SDS-PAGE samples. Towards the later fractions, the band intensity decreases; eventually, the fractions from 1 to 12 were joined for the further purification process. The cleaved HSC70 is present at a small amount in each fractions, which is explained by the *in vivo* occurring SUMO-His$_6$ cleavage.

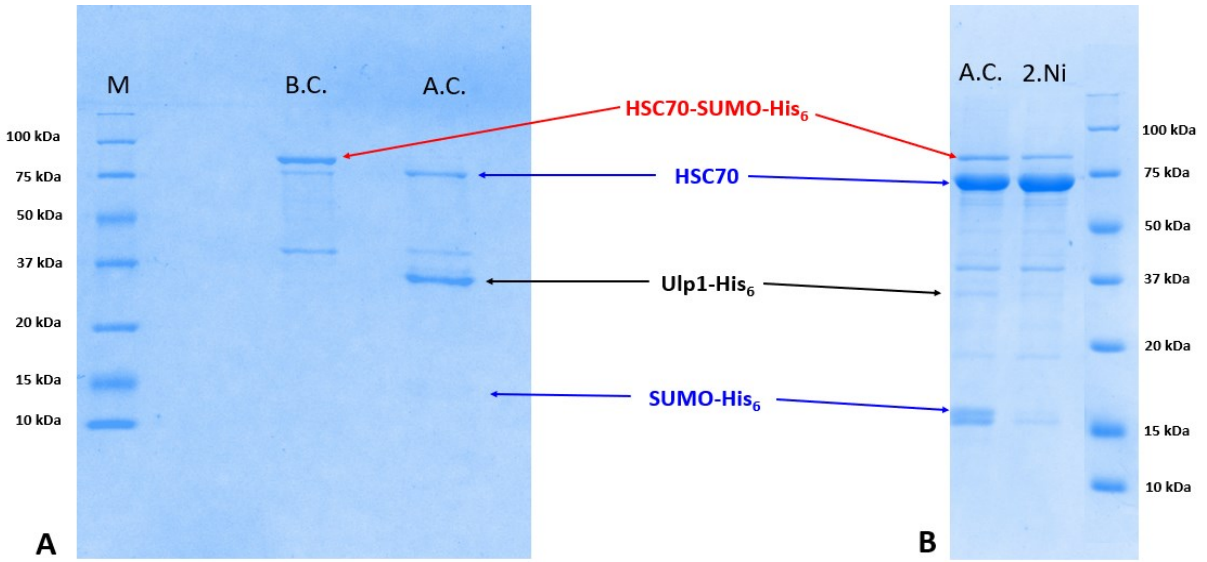


Fig.19. Ulp1 activity test (**A**) and the SUMO-His$_6$ removal by the 2nd (reverse) IMAC (**B**). B.C. - before cleavage, A.C. - after cleavage, M - marker, 2.Ni - after the reverse IMAC process. Ulp1 ensured the effective cleavage of the vast majority of tagged HSC70, which ultimately led to the electrophoretic shift of the initial tagged HSC70 band downwards. During the purification process, the band corresponding to the cleaved SUMO-His$_6$ is more pronounced (A.C.); however, it's significantly diminished after performing the reverse IMAC (2.Ni). The Ulp1-His$_6$ band has the lower intensity after the 2nd IMAC as well, since it's captured by the Ni-NTA stationary phase in the column.

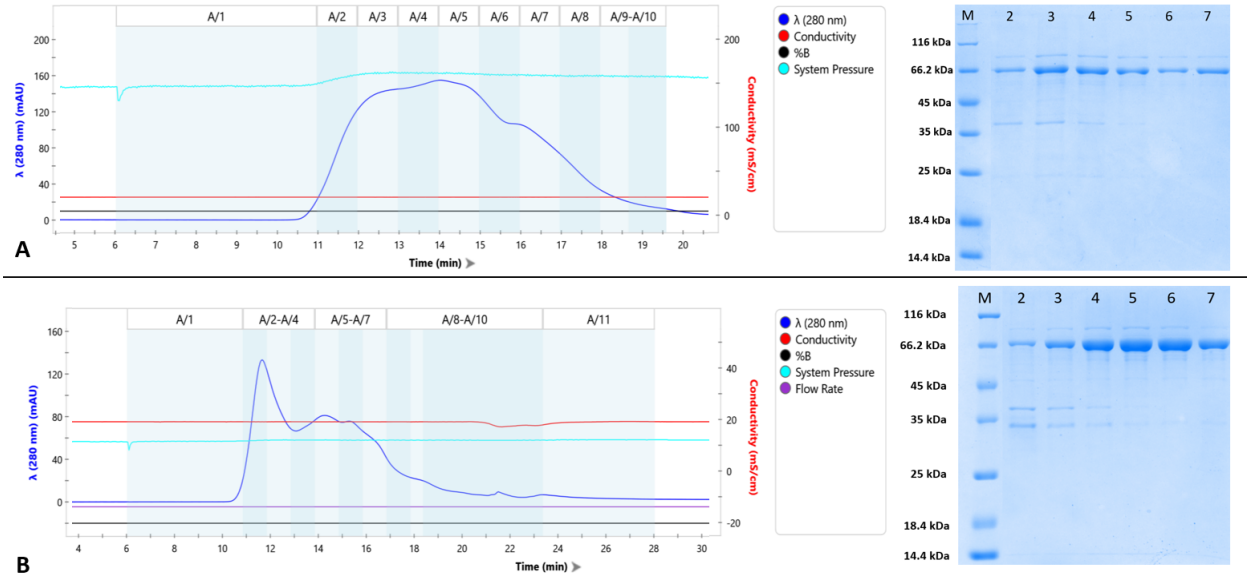


Fig. 20. The SEC chromatogram and the corresponding SDS-PAGE fraction analysis for ^{14}N-HSC70 WT (**A**) and ^{15}N-HSC70 WT (**B**). The initial fractions contain the unknown 35-45 kDa band and apparently the unremoved Ulp1 protease, as well as lower amount of HSC70; towards the later fractions, the HSC70-band intensity significantly increases and the other bands disappears; in **A,** the fraction 3-7 were merged and stored, in **B** - fractions 4-7. The similar pattern of separation was observed for the HSC70 mutants (**Appendix 2**).

depicted in **Fig.20**. For the purified ^{14}N- and ^{15}N-HSC70 SBD mutant and linker mutant proteins, the similar data is shown in **Appendix 2.**

Even after the successful cleavage by Ulp1, some HSC70-SUMO-His$_6$ bands are still visible on the SEC fractions, together with the unknown 35-45 kDa band. However, this unknown band isn't very conspicuous and slowly disappears towards the later SEC fractions. The choice of the fractions to merge was dependent on the relative concentration of HSC70 and the contaminants (Ulp1 etc.). Because no more than 4-5 mg of the protein can be injected into the SEC column used in this process (see **Methods**), during the purification of one protein several injections and SEC runs were performed, producing the same outcome. In total, per 600 ml of the bacterial culture grown on the minimal M9 medium, approximately 8-10 mg of the purified protein was obtained. The overall quality and purity of the purified HSC70 WT and mutants in both ^{14}N- and ^{15}N-forms were estimated as satisfactory and sufficient for the following mass photometry and mass spectrometry experiments.

Mass photometry

The mass photometry measurements allowed us to compare the ATP-dependent dimerization abilities of HSC70 WT, SBD-mutant and linker-mutant in time.Nonetheless, the overnight incubation at room temperature didn't lead to protein denaturation or aggregation in none of the studied samples. It's worth noting that the representation of a chosen molecular species is calculated by the data analysis software (i.e., AcquireMP) to the total amount of the detected counts and is therefore irrelevant for us; instead, the dimer/monomer relative abundance is of a greater interest. During the initial measurements at the 0 min time point, i.e., before the addition of ATP, the dimer population level corresponded to the previous estimations. In relation to the monomer, HSC70 WT exhibited 0.40-0.45 relative dimer abundance. On the contrary, the SBD-mutant and the linker-mutant achieve lower levels of the dimer, being 0.06-0.07 for the SBD and 0.19-0.20 for the linker-mutant (**Fig. 21A-C**). For the mixed experiments, 1:1 equimolar mix of HSC70 WT and SBD-mtuant, HSC70 WT and linker-mutant, SBD-mutant and linker-mutant, the very different initial dimer populations were observed (**Fig.21D-F)**. The highest dimer abundance value among the mixed incubations was detected for the WT:SBD-mutant experiment (0.37-0.38), almost reaching the levels indicated by WT alone. For the WT:linker-mutant case, the calculated dimer/monomer ratio (0.32) falls into the interval of values between the WT alone and the linker-mutant alone (from 0.20 to 0.45). The same can be concluded from the SBD-mutant:linker-mutant measurement, where the dimer abundance value (0.13-0.14) is accommodated within the interval of values corresponding to the proteins alone (from 0.06-0.07 to 0.20).

The necessary ATP-dependent conformational shift must occur to break down the old dimer populations and allow the heterodimer species to form. In order to verify this hypothesis, the ATP-dependent dimerization was tackled for the aforementioned protein mixtures by the mass photometry in time.

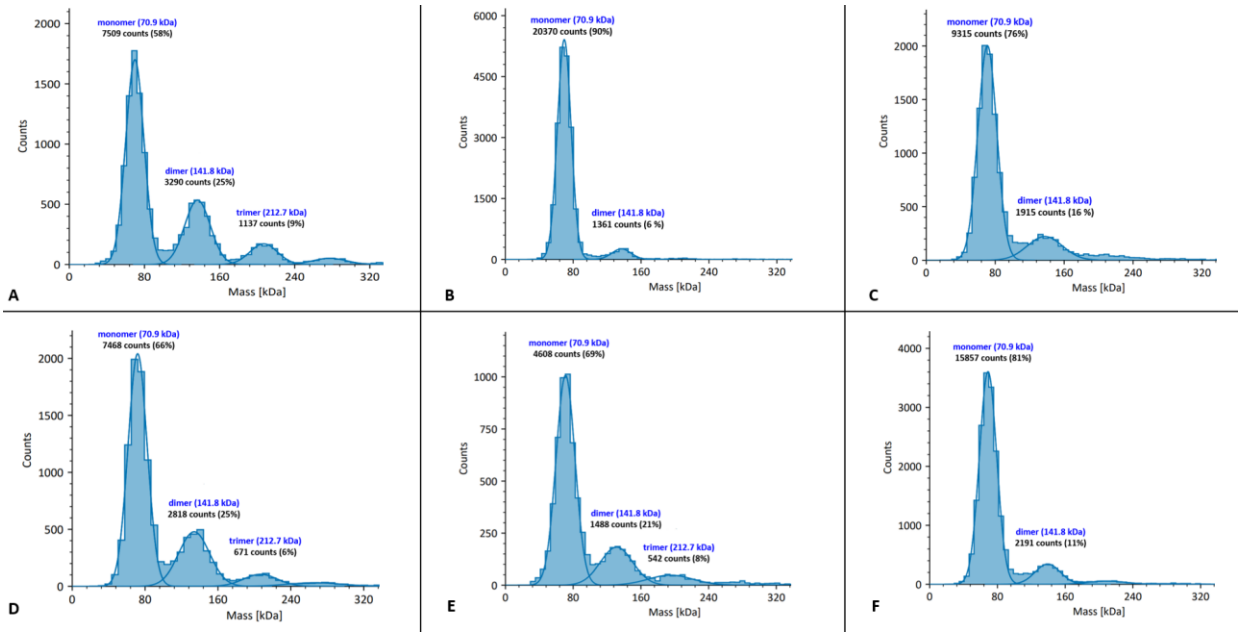


Fig.21. The mass photometry measurements of the HSC70 WT alone (**A**), SBD-mutant alone (**B**), linker-mutant alone(**C**), equimolar WT:SBD-mutant (**D**), equimolar WT:linker-mutant (**E**) and equimolar SBD-mutant:linker-mutant (**F**) at 20 μM total concentration in HKM buffer after overnight incubation at room temperature. In case of WT, the dimer population reaches around 43% relative to the monomer; even trimer and higher oligomeric species were detected. However, for the mutants the dimer level is significantly lower (6-7% for SBD-mutant and 19-20% for linker mutant) and higher oligomeric species don't reach beyond the noise level. Among the equimolar mixtures of the HSC70 variants, the highest dimer proportion is reached by the WT:SBD-mutant (37-38%). The WT:linker-mutant mixture and the mixture of mutants exhibited 32% and 14% dimer abundance respectively. The trimer population was detected for the WT:SBD-mutant and WT:linker-mutant mixtures, but not for the SBD:linker-mutant mixtures.

To quickly achieve the reformation of the HSC70 dimer population, double molar amount of ATP was used (20 μM protein and 40 μM ATP in our experiment setup). The rough estimation of the ATP-hydrolysis reaction half-time ($t_{1/2}$) was 30-40 min, meaning the majority of ATP will be hydrolyzed after 1-2 hours of incubation (for bacterial DnaK, $t_{1/2} \cong$ 20 min – *Laufen et al. 1999*). For the WT, the effect of ATP binding on the dimer population was noticeable after 10 min of incubation (**Fig. 22**). The dimer peak immediately decreased about 3 times (from 0.43 to 0.14 ratio) and remained at the same level for the following 30 min after the ATP addition. When the ATP pool is consumed, approximately after 30-40 min of incubation, the dimer species abundance starts to increase, reaching even the higher dimer/monomer ratio after 90 min (0.49-0.50). For the rest of the samples, the experiments were conducted in the same manner (**Appendix 3**) and the dimer/monomer ratio values were plotted against the time (**Fig.23**). Interestingly, the dimer level after 90 min achieved in the WT:SBD-mutant experiment (0.5) is the same as was observed in the WT only sample (0.5).

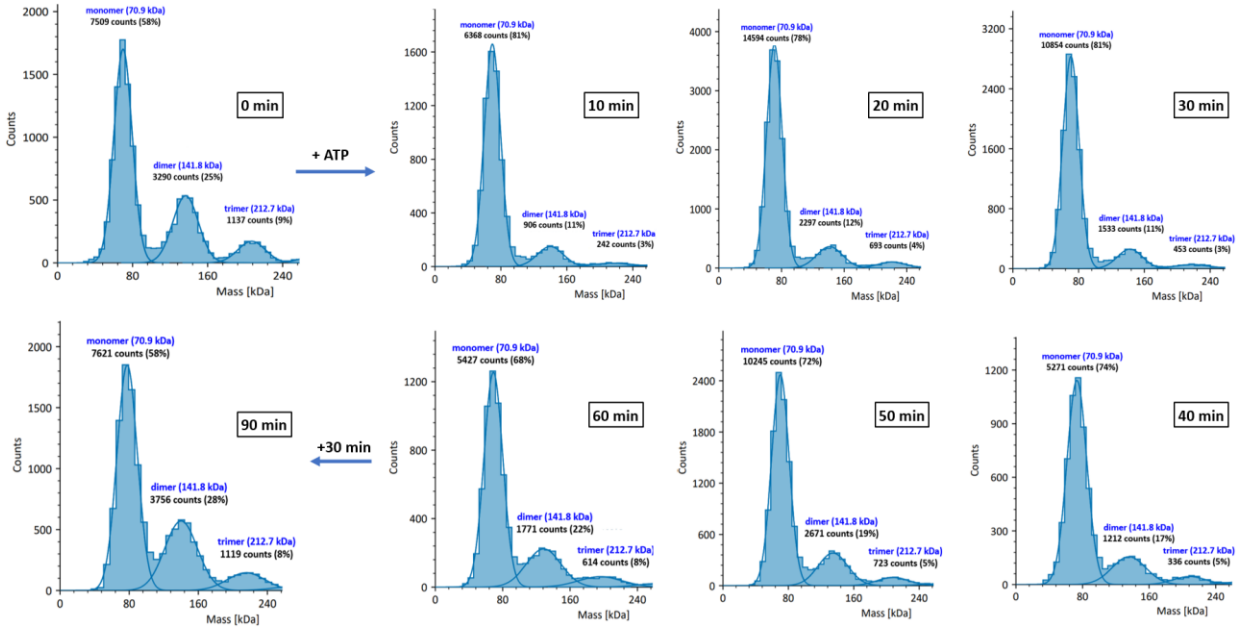


Fig.22. The ATP-dependent dimerization of HSC70 WT in time. Ten minutes after the ATP addition, the dimer level significantly decreases and gradually restores, reaching even higher levels than before the ATP was added (0 min - 0.43 vs 90 min - 0.49). The gradual increase of the dimer population abundance begins 30-40 min after the ATP addition. For the ATP-dependent mass photometry measurements of other protein mixtures, see **Appendix 3.**

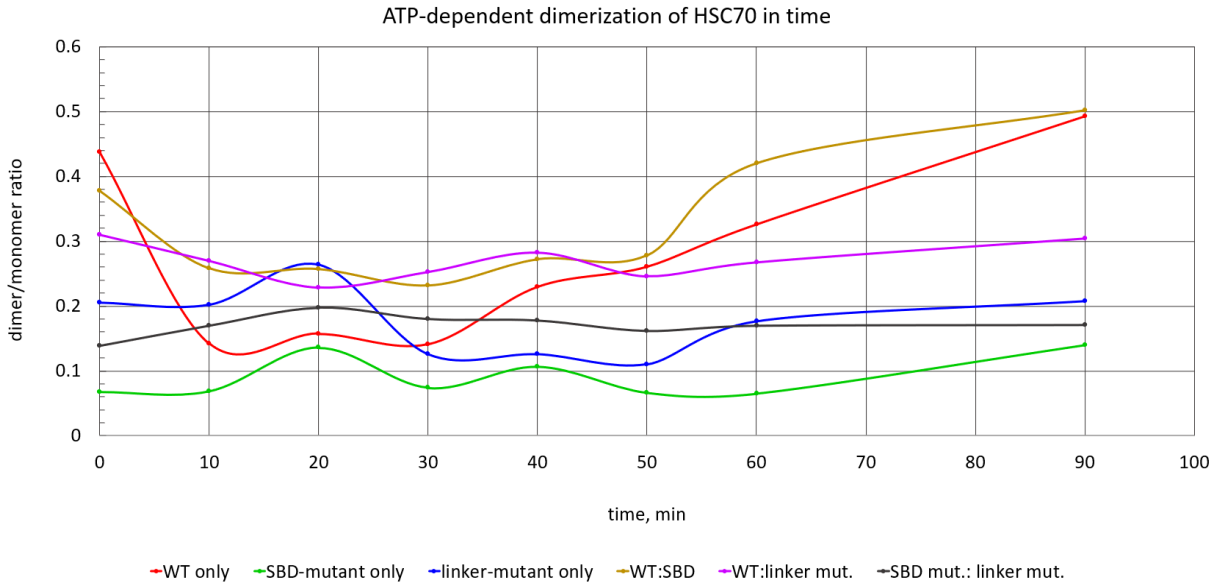


Fig.23. The dimer/monomer ratio changes in time after the addition of ATP. The most drastic changes of the dimer abundance was observed for HSC70 WT only experiment; on the contrary, the both mutants (green and blue curves) exhibited no reasonable pattern in dimer level changes. Nevertheless, for the non-mixed experiments (WT only, SBD-mutant only, and linker-mutant only), the local maximum at the 20 min time-point can be visible. For the mixed experiments, the highest dimer level at the final time-point was observed for the WT:SBD-mutant experiment - the dimer/monomer ratio eventually reached the similar value as the WT-alone (red and gold curves). The WT:linker-mutant and SBD-mutant:linker-mutant samples didn't respond to the ATP addition, since only minor fluctuations of the dimer abundance were detected.

The pattern of the dimer level changes is also similar to the WT only, but the initial dimer decrease after the ATP addition is less pronounced (for the WT only, the dimer level drop is from 0.43 to 0.14, i.e., about 3x; for the WT:SBD-only, the drop is from 0.37-0.38 to 0.25, i.e. about 1.5x). In the case of the SBD-mutant only sample, the lowest dimer/monomer ratio is observed among all other samples; moreover, no ATP-dependent dimerization pattern similar to WT can be seen.

For the linker-mutant only measurement, the overall dimer/monomer ratio is higher than in the case of SBD-mutant, but still 2.5x times lower than in the WT case. After 90 min, the initial dimer level is achieved. The moderate increase 20 min after the ATP addition can be seen for WT only, SBD only and linker only. However, in the 20-30 min time range, the dimer level decrease occurs for all of these non-mixed experiments:. In the case of WT:linker mutant mixture, no drastic fluctuations in dimer/monomer ratio were observed; eventually, the initial ratio value was restored at 90 min time-point. It suggests that less interaction between WT and linker-mutant occurs, indicating that the polar residues introduced to the HSC70 linker area affect the subunit interaction in the dimer. The similar situation is seen in the two mutants mix - the overall pattern of the dimerization for the WT:linker-mutant and the SBD-mutant:linker-mutant experiments is quite similar, with the exception of even lower interaction happening between the mutants.

Based on the mass photometry measurements, it was discovered that 90 min after the addition of ATP in double-molar equivalent, an HSC70 dimer population undergoes the dissociation and the subsequent reassembly, reaching or even overcoming the initial dimer/monomer ratio. By the incubation with ATP, the artificial heterogeneity in the dimer population is eliminated and the new "monomer-dimer" equilibrium is reached. These conditions were found optimal to perform the cross-linking reaction and analyze the HSC70 dimer structure by structural mass spectrometry.

XL-MS data analysis and processing

In total, 9 XL-MS experiments were conducted: 1) ^{14}N-WT + ^{15}N-WT; 2) ^{14}N-SBD-mutant + ^{15}N-SBD-mutant; 3) ^{14}N-WT + ^{15}N-SBD-mutant; 4) ^{15}N-WT + ^{14}N-SBD-mutant; 5) ^{14}N-linker-mutant + ^{15}N-linker-mutant; 6) ^{14}N-WT + ^{15}N-linker-mutant; 7) ^{15}N-WT + ^{14}N-linker-mutant; 8) ^{14}N-SBD-mutant + ^{15}N-linker-mutant; 9) ^{15}N-SBD-mutant + ^{14}N-linker-mutant. For each of

the mixed experiments (WT+SBD-mutant, WT+linker-mutant, SBD-mutant+linker-mutant), both direct and reverse labeling experiments were conducted to ensure the persistence of identified XL-sites orientation. After the XL-reaction and prior to any further processing, the protein mixtures were analyzed by SDS-PAGE (**Fig.24**). Surprisingly, no visible differences between the different experiments were observed. The band distribution is quite similar - as well as in the control, the pronounced monomer band can be seen; the most retarded bands corresponding to the dimer (above 116 kDa) are visible too in each replicate, but not in the control. It's noteworthy that in each XL-replicate the multiple intermediate bands between the monomeric and dimeric bands are also clearly distinguishable.

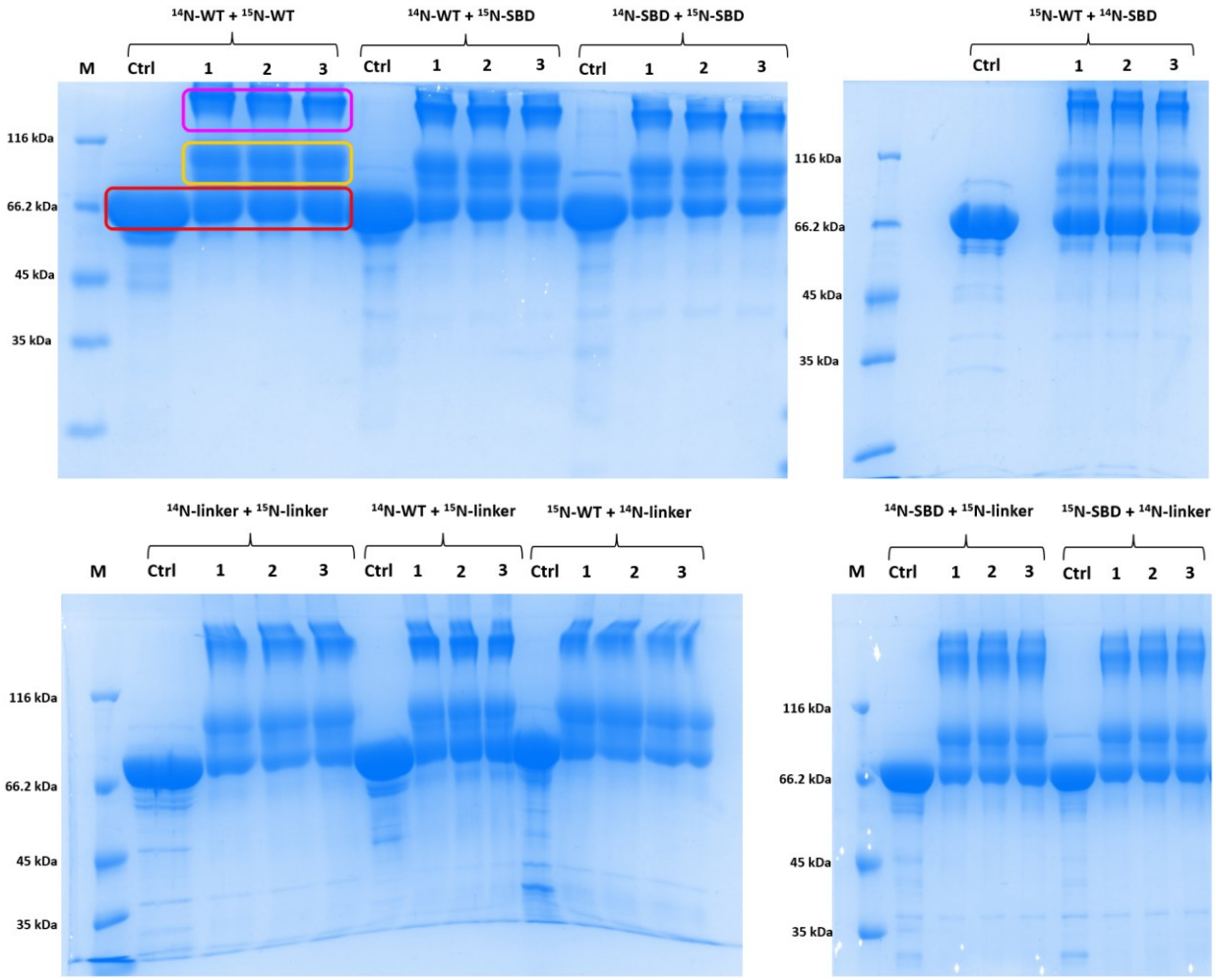


Fig.24. The SDS-PAGE of the cross-linked protein mixtures.. M - marker, Ctrl - control (non-cross-linked sample), 1/2/3 - number of a replicate. The distinguishable monomer (red) and dimer (magenta) bands are seen in each replicate and experiment; however, the dimeric band isn't visible in the controls. During the XL-reaction, the putative intra-cross-linked monomeric species with the delayed electrophoretic mobility emerge (yellow).

After the cross-linking reaction, the subsequent modifications (i.e., thiol group reduction and carbamidomethylation) and tryptic digestion were performed. The samples prepared in this manner were then analyzed by LC-MS (see *LC-MS measurements*). The collected LC-MS data were then processed by LinX software; the raw mass spectra were analyzed in Compass Data Analysis software. Each replicate was checked for the sequence coverage: in all cases, the very satisfactory HSC70 sequence coverage was achieved, with the most hits from the N-terminal NBD region (residues 1-129) and from the C-terminal SBD region and C-tail (residues 515-646) (**Fig.25**). However, no significant differences in sequence coverage were observed among different experiments, meaning that the mutations introduced into HSC70 WT don't affect the digestion process.

The number of the XL-sequences identified by LinX was variable across the different experiments (**Tab.6**). The largest numbers of identifications were observed for the samples containing linker-mutant; on the other hand, the lowest number of hits was obtained for the ^{14}N-WT + ^{15}N-WT experiment. These results were manually validated by checking the raw

Tab.6. The number of identified XL-sequences by the LinX software.

Experiment	Number of identified XL-sequences
^{14}N-WT + ^{15}N-WT	973
^{14}N-SBD-mutant + ^{15}N-SBD-mutant	1174
^{14}N-WT + ^{15}N-SBD-mutant	1222
^{15}N-WT + ^{14}N-SBD-mutant	1434
^{14}N-linker-mutant + ^{15}N-linker-mutant	1836
^{14}N-WT + ^{15}N-linker-mutant	1760
^{15}N-WT + ^{14}N-linker-mutant	1634
^{14}N-SBD-mutant + ^{15}N-linker-mutant	1885
^{15}N-SBD-mutant + ^{14}N-linker-mutant	2001

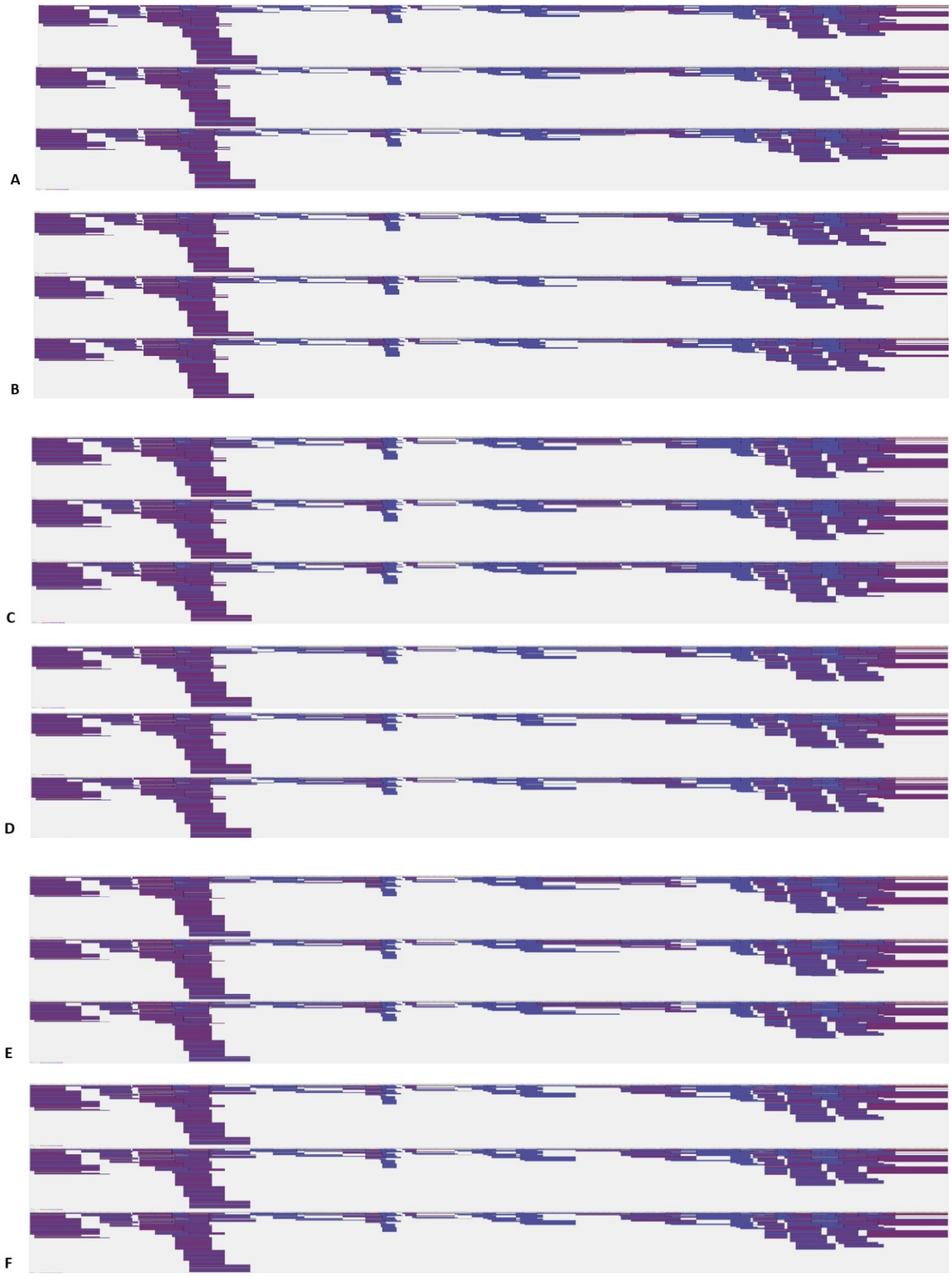
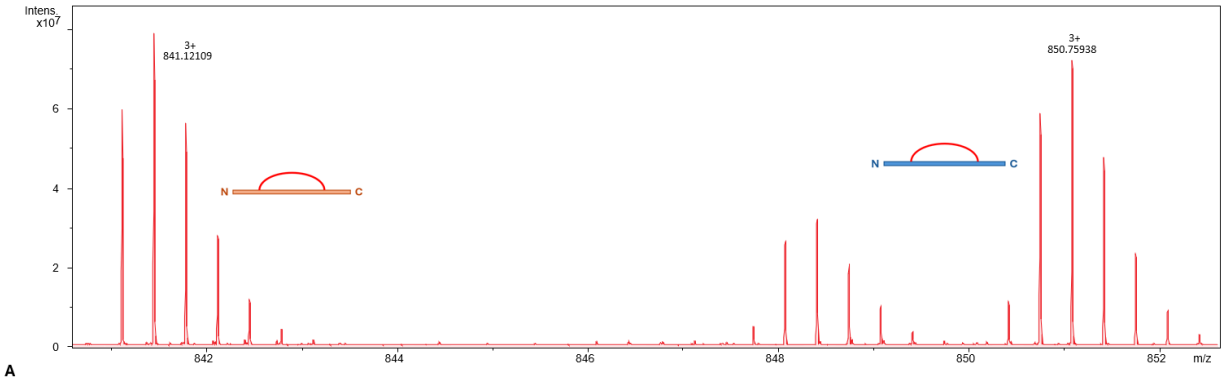

A
B
C
D
E
F

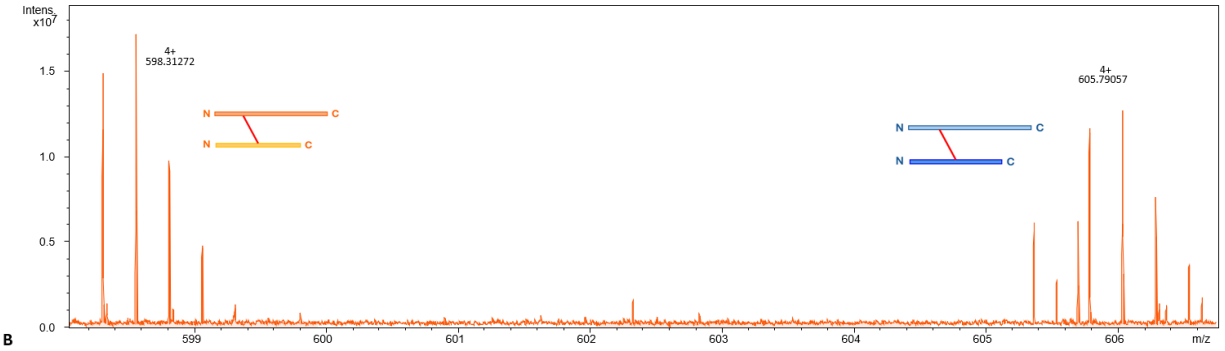
Fig.25. Sequence coverage of the individual XL-MS experiments. **A -** ^{14}N-WT + ^{15}N-WT; **B -** ^{14}N-SBD-mutant + ^{15}N-SBD-mutant; **C -** ^{14}N-WT + ^{15}N-SBD-mutant; **D -** ^{15}N-WT + ^{14}N-SBD-mutant; **E -** ^{14}N-linker-mutant + ^{15}N-linker-mutant; **F -** ^{14}N-WT + ^{15}N-linker-mutant; **G -** ^{15}N-WT + ^{14}N-linker-mutant; **H -** ^{14}N-SBD-mutant + ^{15}N-linker-mutant; **I -** ^{15}N-SBD-mutant + ^{14}N-linker-mutant. White - unmodified peptide, red - peptide with any modification including hanging cross-linker (dead-end cross-linker); blue - intra- and inter-cross-linked peptides; purple - intra- and inter-cross-linked peptides with modification.

mass spectra for the corresponding hits. The detected XL-peptides can be classified into several categories. For each of the types, an example spectrum and an output from LinX software is provided (**Fig.26**). Type 1 is basically an intra-peptide cross-link, where the closely localized lysine residues get attached by the reagent (**Fig.26A**). In this case, a cross-link can be identified only in ^{14}N/^{14}N (LL) and ^{15}N/^{15}N (HH) forms and 100% comes from a monomer (0% inter XL). Type 2 has the similar spectral pattern as type 1, however, here 2 peptides are cross-linked (**Fig.26B**). The cross-linked lysine residues are also located in proximity in the native polypeptide, but within the cross-linked region a proteolytic cleavage has occurred. This XL is also 100% coming from a monomer, so only LL and HH forms are detected in spectra. Type 3 is the most valuable cross-links, because all the 4 forms (LL, LH, HL, HH) are detected at more or less the same intensity (**Fig.26C**). The inter/intra ratio (see equation from **Fig.14**) in this case is within 0.9-1.0 range and therefore 90-100% comes from a dimer. The type 3 cross-links were eventually used to build the HSC70 dimer model.



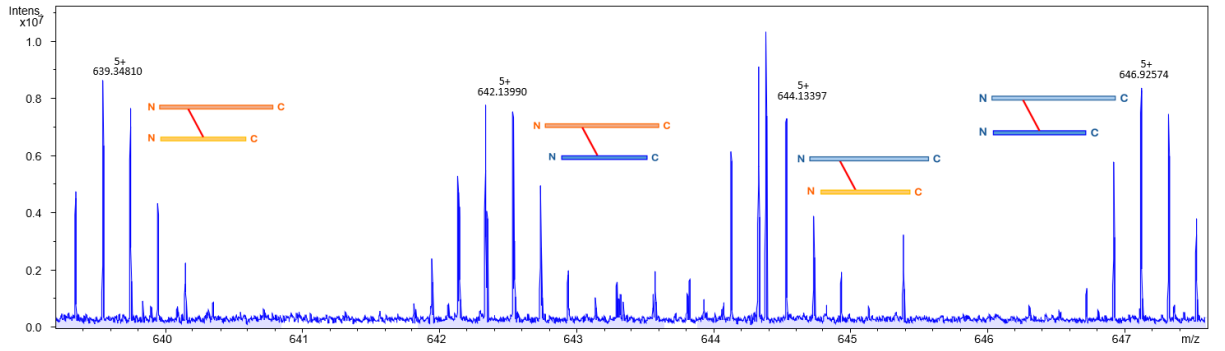
A

Cross-link	Protein	Modification	Bonds	Charge (X)	Scan	Forms found	LL m/z (1+)	LL m/z (X+)	LH m/z (1+)	LH m/z (X+)	HL m/z (1+)	HL m/z (X+)	HH m/z (1+)	HH m/z (X+)	Inter XL (%)	Intra XL (%)
AMTKDNNLLGKFELTGIPPAPR	HSC70 WT (448, 469)		DSSd0 (451; 458)	3	688-726	2	2521.348	841.121	NaN	NaN	NaN	NaN	2550.262	850.759	0	100



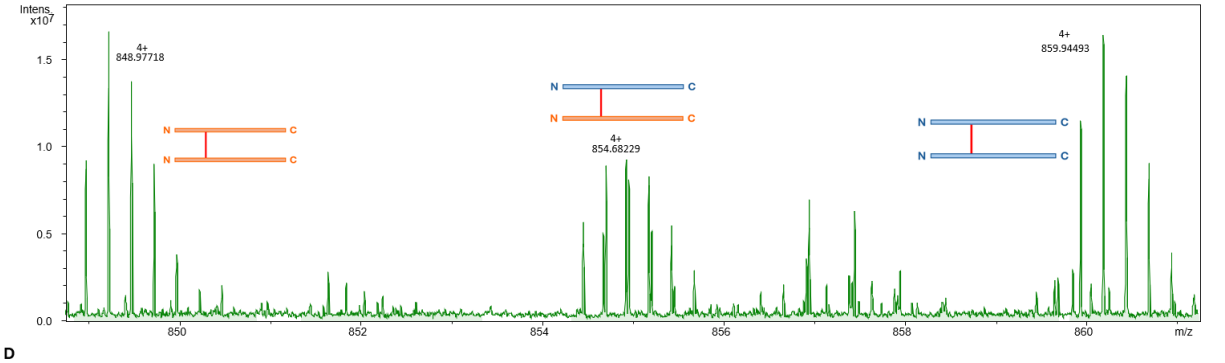
B

Cross-link	Protein	Modification	Bonds	Charge (X)	Scan	Forms found	LL m/z (1+)	LL m/z (X+)	LH m/z (1+)	LH m/z (X+)	HL m/z (1+)	HL m/z (X+)	HH m/z (1+)	HH m/z (X+)	Inter XL (%)	Intra XL (%)
MVNHFIAEFKR - DISENKR	HSC70 WT (237, 247) [A] - HSC70 WT (252, 258) [B]		DSSd0 (A.246; B.257)	4	440-446	2	2390.229	598.3126	2402.193	601.3037	2408.175	602.7992	2420.14	605.7904	0	100



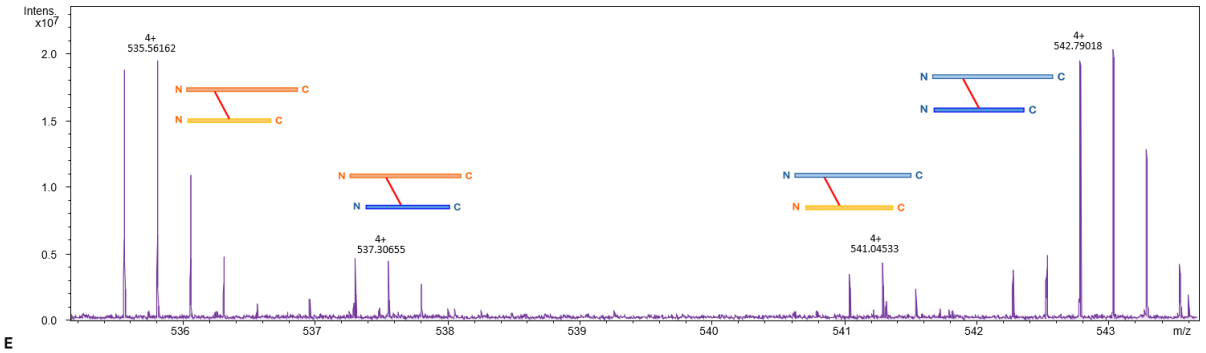
C

Cross-link	Protein	Modification	Bonds	Charge (X)	Scan	Forms found	LL m/z (1+)	LL m/z (X+)	LH m/z (1+)	LH m/z (X+)	HL m/z (1+)	HL m/z (X+)	HH m/z (1+)	HH m/z (X+)	Inter XL (%)	Intra XL (%)
LDKSQIHDIVLVGGSTR - ATVEDEKLQGK	HSC70 WT (326, 342) [A] - HSC70 WT (551, 561) [B]		DSSd0 (A.328; B.557)	5	482-488	4	3192.711	639.348	3206.67	642.1397	3216.64	644.1338	3230.598	646.9255	100	0



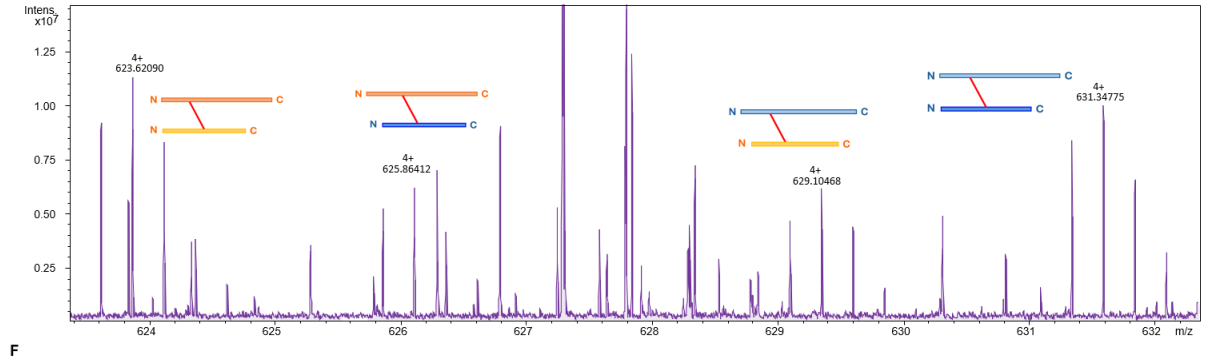
D

Cross-link	Protein	Modification	Bonds	Charge (X)	Scan	Forms found	LL m/z (1+)	LL m/z (X+)	LH/HL m/z (1+)	LH/HL m/z (X+)	HH m/z (1+)	HH m/z (X+)	Inter XL (%)	Intra XL (%)
QATKDAGTIAGLNVLR - QATKDAGTIAGLNVLR	HSC70 WT (156, 171) [A] - HSC70 linker (156, 171) [B]		DSSd0 (A.159; B.159)	4	486-491	4	3392.886	848.977	3414.8208	854.4606	3436.756	859.9443	28	72



E

Cross-link	Protein	Modification	Bonds	Charge (X)	Scan	Forms found	LL m/z (1+)	LL m/z (X+)	LH m/z (1+)	LH m/z (X+)	HL m/z (1+)	HL m/z (X+)	HH m/z (1+)	HH m/z (X+)	Inter XL (%)	Intra XL (%)
AKR - QATKDAGTIAGLNVLR	HSC70 WT (270, 272) [A] - HSC70 linker (156, 171) [B]		DSSd0 (A.271; B.159)	4	343-349	4	2139.224	535.5616	2161.159	541.0453	2146.204	537.3064	2168.138	542.7901	20	80



F

Cross-link	Protein	Modification	Bonds	Charge (X)	Scan	Forms found	LL m/z (1+)	LL m/z (X+)	LH m/z (1+)	LH m/z (X+)	HL m/z (1+)	HL m/z (X+)	HH m/z (1+)	HH m/z (X+)	Inter XL (%)	Intra XL (%)
QATKDAGTIAGLNVLR - IPKIQK	HSC70 WT (156, 171) [A] – HSC70 WT (343, 348) [B]		DSSd0 (A.159; B.345)	4	471-476	4	2491.461	623.6206	2500.434	625.8639	2513.395	629.1043	2522.369	631.3476	50	50

Fig.26. Types of the detected XL-peptides.

The type 4 cross-link occurs, when the same residues on the 2 same peptides get attached; only 3 possible forms (LL, LH/HL, HH) can be detected, because $^{14}N/^{15}N$ and $^{15}N/^{14}N$ forms have identical m/z value (**Fig.26D**). For type 4, the equation for the inter/intra ratio is therefore slightly modified - the numerator contains only LH/HL intensity instead of 2 intensities of LH and HL forms. Both type 5 and 6 belong to the certain "intermediates" between type 2 (0% inter) and type 3 (100% inter): the calculated ratio is below 0.9-1.0, but higher than 0 (**Fig.26E,F**). Thus, it's challenging to unambiguously assign such cross-link to a dimer or to a monomer. In this situation we can speculate if a given cross-link originates from a dimer or a monomer based on the calculated ratio, which is converted to the probability. In the **Fig.26E,** the ratio is 0.2, therefore it's 20% probability that this cross-link appears on the interface between subunits of a dimer; so, it's rather intra than inter. In the **Fig.26F,** the ratio is 0.5, so the probability of the dimer origin is 50%. In the later steps of data analysis, the threshold of 0.5 was introduced to filter out rather intra- and rather inter-cross-links: if the ratio is below 0.5, a cross-link is assigned to a monomer; if the ratio is above or equal 0.5, a cross-link is assigned to a dimer. In the ^{14}N-WT + ^{15}N-WT, ^{14}N-SBD-mutant + ^{15}N-SBD-mutant and ^{14}N-linker-mutant + ^{15}N-linker-mutant experiments, only XL-peptides with all possible forms detected

(LL, LH, HL, HH) were taken for further manual validation. However, in the mixed experiments (WT + SBD-mutant etc.), both hits with all 4 forms detected and the hits with 3 forms detected were used for the analysis. The reason for it lies in the fact that in the mixed experiments, non-symmetrical dimer configurations were considered. For instance, if the "substrate-like fashion" dimer model is considered, it becomes clear that certain cross-links cannot occur in both directions.

During the manual validation of cross-links, types 3, 4, 5 and 6 (i.e., with the inter/intra ratio above 0) were selected across all the obtained datasets. The validated cross-links were then compared between the replicates to verify the reproducibility of the experiments (**Fig.27A**). It can be seen on the Venn diagrams that a significant amount of the chosen XL-sites is overlapping between the replicates, which indicates the good reproducibility. In the case of the mixed experiments, the cross-links identified in all 3 replicates were compared between the direct and reverse labeling experiments (**Fig.27B**) - only those detected in 6 replicates for a given sample were further analyzed. Interestingly, much lesser amount of cross-links with 4 forms detected was observed in the case of ^{14}N-WT + ^{15}N-WT and ^{14}N-SBD-mutant + ^{15}N-SBD-mutant samples (16 and 15, respectively); the highest amount of hits was observed for the experiments containing linker-mutant. This is also in consent with the total amount of identified XL-sequences in all samples (**Tab.6**). The list of the manually validated cross-links is attached to **Appendix 4.**

The cross-links identified in all the replicates for the corresponding experiments were then mapped onto the HSC70 sequences (**Fig.28**). In the non-mixed samples (^{14}N-WT + ^{15}N-WT, ^{14}N-SBD-mutant + ^{15}N-SBD-mutant, ^{14}N-linker-mutant + ^{15}N-linker-mutant), the symmetry of the mapped cross-links indicates the symmetrical homooligomeric HSC70 dimer; however, the number of the identified XL-sites is significantly higher in the ^{14}N-linker-mutant + ^{15}N-linker-mutant experiment. In the mixed WT + SBD-mutant sample, the cross-link distribution isn't completely symmetrical; these oriented XL-sites are in agreement with the presumption that non-symmetrical dimer species might be present in the mixed experiments. It's less visible in the WT + linker-mutant and SBD-mutant + linker-mutant experiments, since the cross-ink maps are extremely overcrowded with many XL-bonds.

Despite the evident differences in the number of identified XL-sites among the different samples, in all cases it's still complicated to derive an unambiguous subunit orientation in the putative HSC70 dimer. Therefore, the XL-sites inter/intra ratio values were plotted in a form

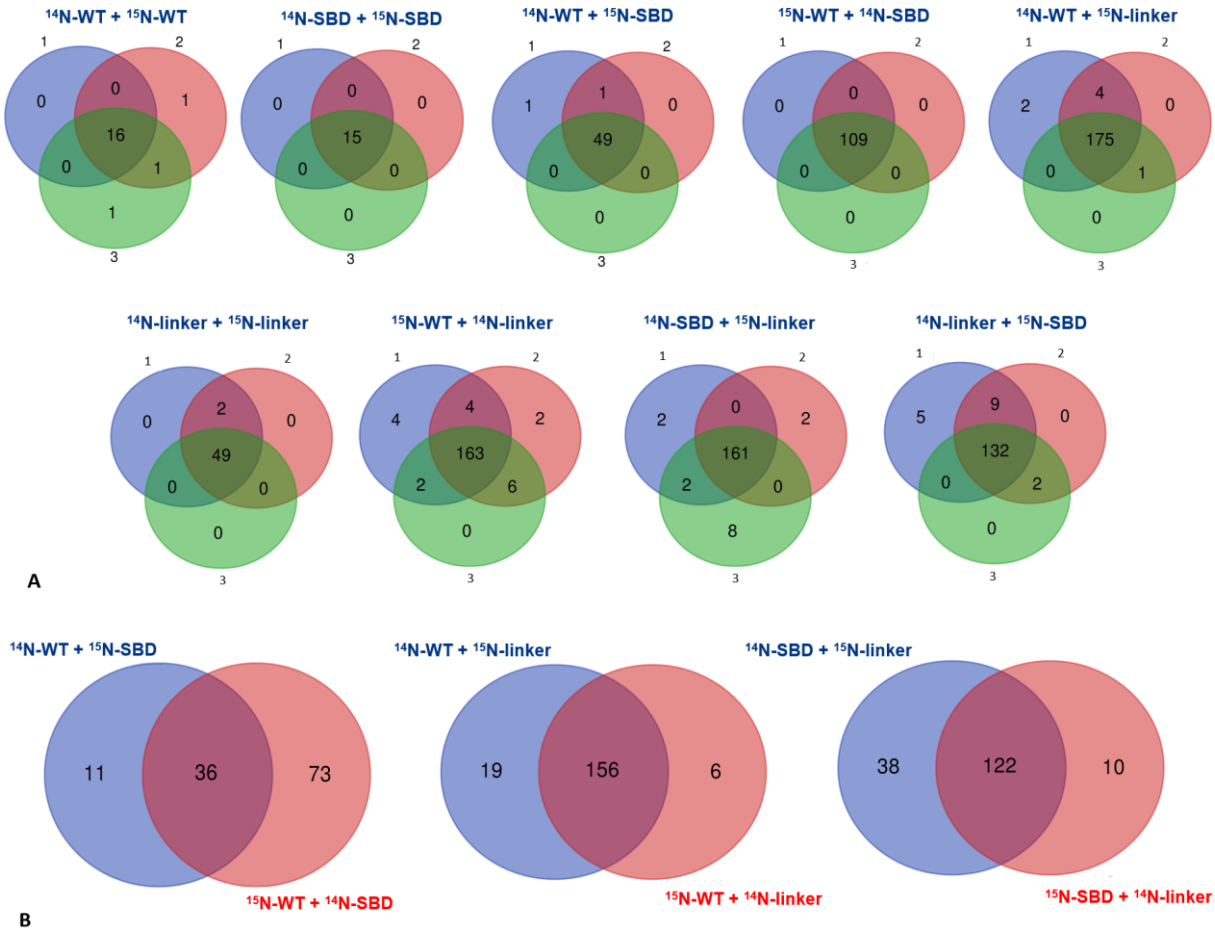


Fig.27. Venn diagrams representing the reproducibility of the XL-MS experiments across the replicates. **A.** The Venn diagrams showing the reproducibility of the individual experiments across the triplicates. **B.** The replicability of the mixed experiments with direct and reverse isotope labeling. In these diagrams, only the XL-sites identified in all 3 replicates of corresponding experiments were compared.

of histogram (**Fig.29**). The histogram plots allow us to see the inter/intra ratio distribution in the analyzed XL-MS datasets. According to the cross-link types described above, the intermediate types (**Fig.26E,F**) can be encountered as well - as can be seen in **Fig.29,** they represent the majority in some samples. In order to distinguish the “rather inter " and “rather intra" cross-links, the ratio threshold 0.5 was applied. In the first 3 cases (i.e., ^{14}N-WT + ^{15}N-WT, ^{14}N-SBD-mutant + ^{15}N-SBD-mutant, WT + SBD), the noticeable difference in the ratio distribution is seen: the certain XL-sites possess the ratio value well above 0.5 (i.e., 0.8-1.2) and the rest have these values below the 0.5 threshold. However, it isn’t the case for the linker-mutant experiments (**Fig.29D,E,F**) - the vast majority of the XL-sites are of the ratio below 0.5, which suggests their rather monomeric, or intra-protein origin.

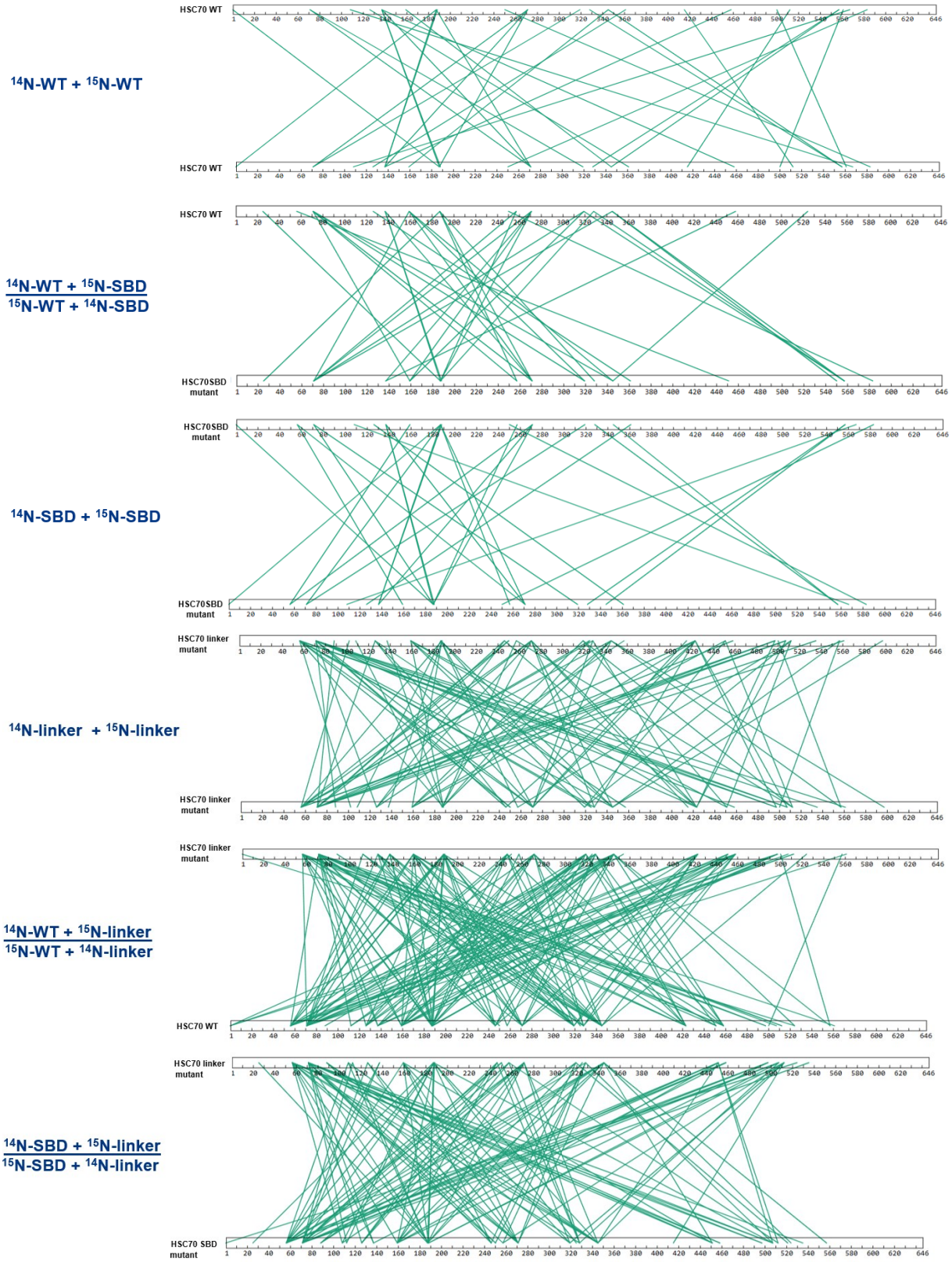
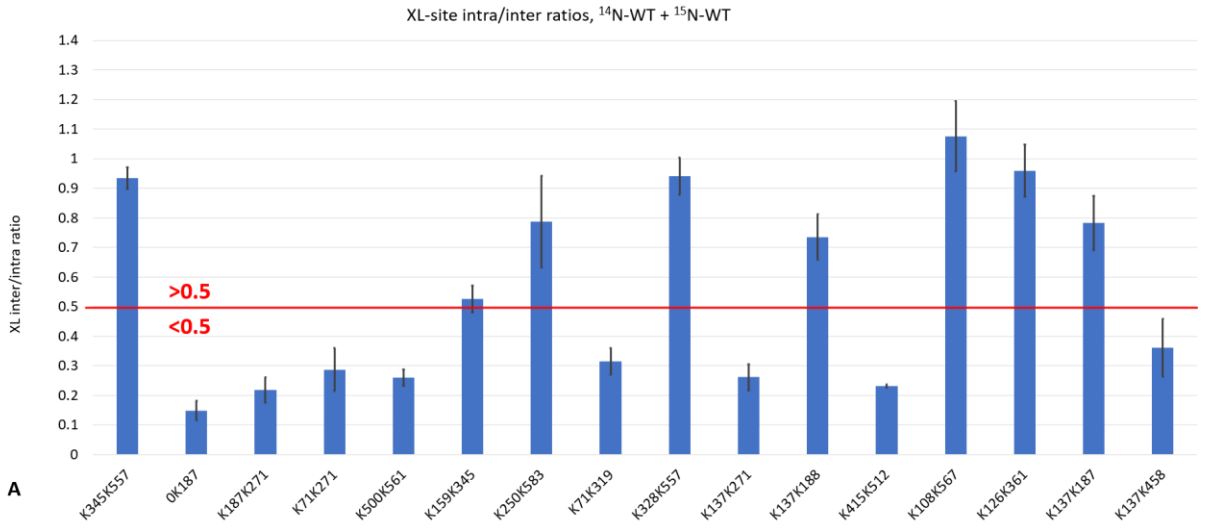
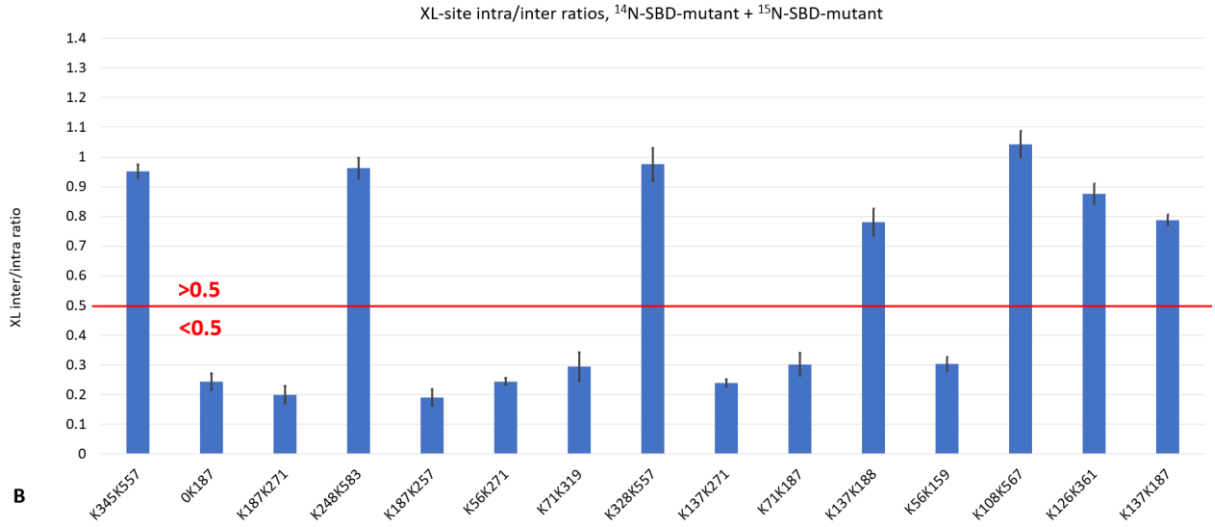


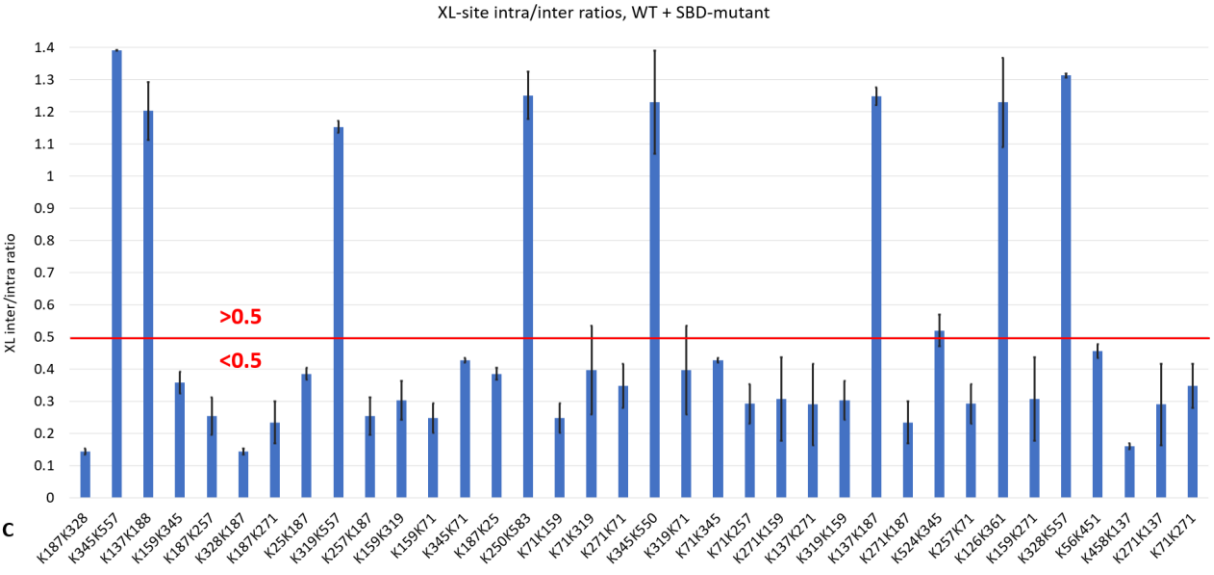
Fig.28. Mapping of the identified XL-sites onto the HSC70 sequences.


XL-site intra/inter ratios, ^{14}N-WT + ^{15}N-WT
XL inter/intra ratio
>0.5
<0.5
K345K557
0K187
K187K271
K71K271
K500K561
K159K345
K250K583
K71K319
K328K557
K137K271
K137K188
K415K512
K108K567
K126K361
K137K187
K137K458

A


XL-site intra/inter ratios, ^{14}N-SBD-mutant + ^{15}N-SBD-mutant
XL inter/intra ratio
>0.5
<0.5
K345K557
0K187
K187K271
K248K583
K187K257
K56K271
K71K319
K328K557
K137K271
K71K187
K137K188
K56K159
K108K567
K126K361
K137K187

B


XL-site intra/inter ratios, WT + SBD-mutant
XL inter/intra ratio
>0.5
<0.5
K187K328
K345K557
K137K188
K159K345
K187K257
K328K187
K187K271
K25K187
K319K557
K257K187
K159K319
K159K71
K345K71
K187K25
K250K583
K71K159
K71K319
K271K71
K345K550
K319K71
K71K345
K71K257
K271K159
K137K271
K319K159
K137K187
K271K187
K524K345
K257K71
K126K361
K159K271
K328K557
K56K451
K458K137
K271K137
K71K271

C

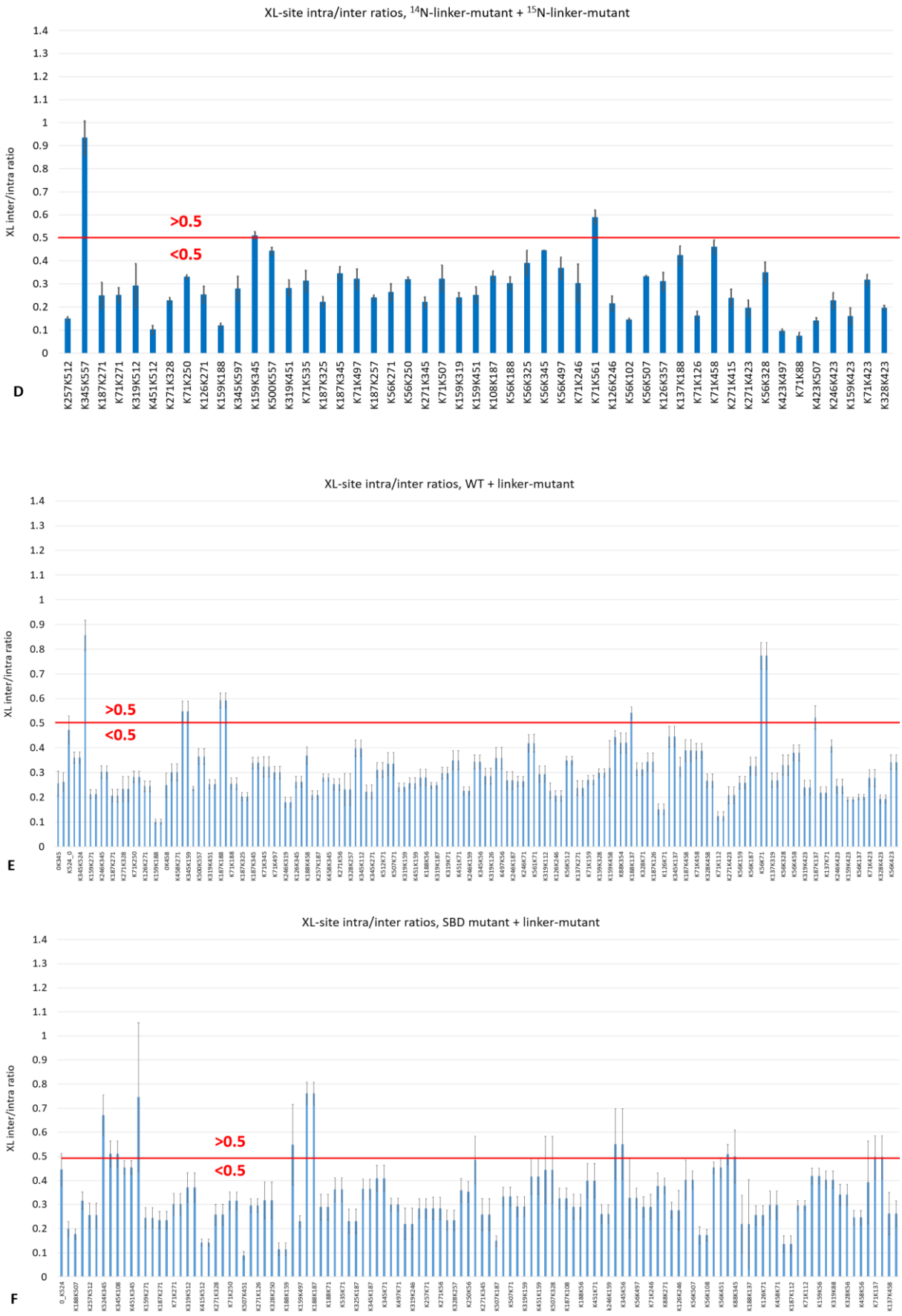


Fig.29. The histogram plots representing the inter/intra ratio distribution across the XL-MS datasets.

Homology modeling of HSC70 in ADP-state

The output of the pairwise sequence alignment of the target human HSC70 WT and the bacterial DnaK is depicted in **Fig.30.** Even though the C-terminus of the DnaK template hasn't been included to the structural NMR model due to its disordered nature, the main domains and the linker were successfully aligned to the target HSC70 sequence. The alignment was performed with 45.6 % sequence identity and 59.6 % sequence similarity, which clearly indicates the homology of the aligned proteins. Therefore, *E.coli* DnaK was considered as an appropriate template for the HSC70 homology modeling.

The obtained HSC70 homology models and their superposition with template DnaK is shown in **Fig.31.** According to the RMSD superposition values, the best model is model 9 (0.789 Å). This model was thus used for the XL-mapping and docking.

Fig.30. The pairwise sequence alignment of human HSC70 and *E.coli* DnaK. Vertical lines - conserved/identical residue, single dot - semi-conservative substitution, double dot - conservative substitution, dashed line - gaps.

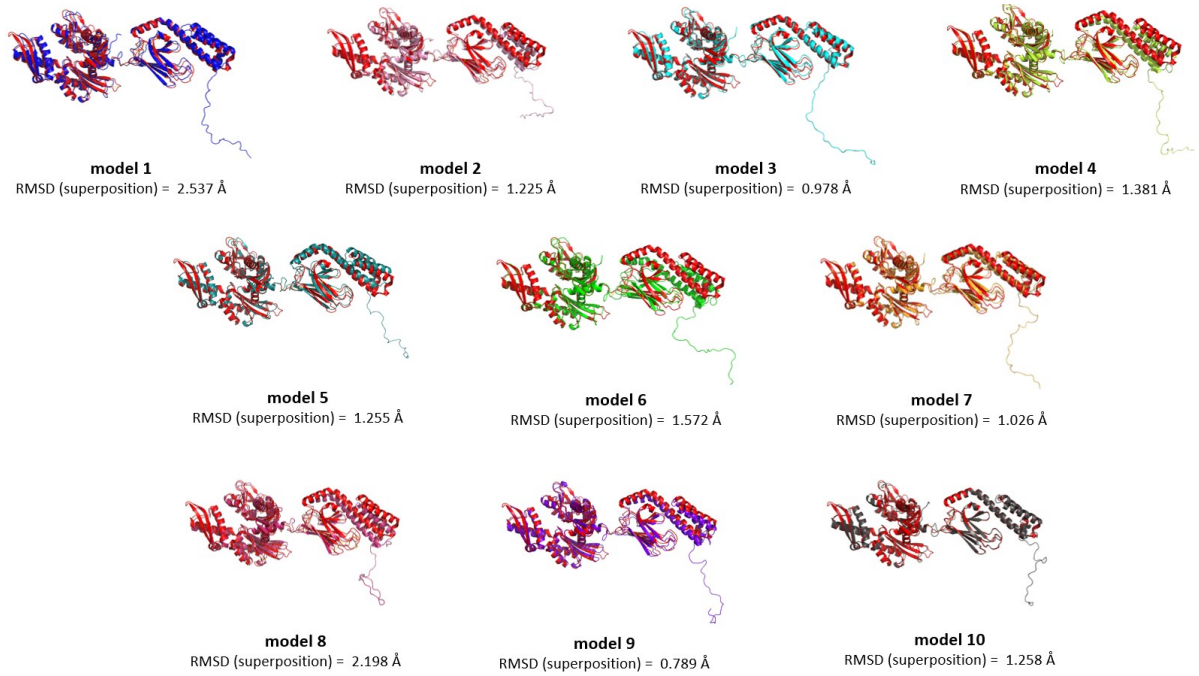


Fig.31. The homology models of human HSC70 in the ADP-bound conformation superimposed onto the *E.coli* DnaK template structure (red).

Docking of the HSC70 WT dimer

After the visual inspection and mapping of the identified cross-links from the $^{14}N/^{15}N$-HSC70 WT experiment, only those with the inter/intra ratio value above 0.5 were used for the docking (**Fig.29A**). Nevertheless, because we applied the aforementioned approach for flexible multidomain proteins (see "*Docking and validation of the HSC70 WT dimer model"* section), it was necessary to establish a mutual position of the separated HSC70 domains. To imitate the distance constraints imposed by the linker in the native protein, an identified intra-cross-link was introduced for the docking – namely, the one connecting the N-terminal methionine amino-group (located on NBD) with the lysine 512 side chain amino-group (located on SBD). This cross-link was defined therefore only between the domains originating
from the same monomer (i.e., between "Chain A" and "Chain B" and between "Chain C" and "Chain D"). After generating the distance restraints TBL file with XMAS ChimeraX package, it was manually edited in Windows Notepad and adjusted to the separated domain docking. It was necessary to perform since the input names for HADDOCK didn't match the actual object names. For example, 1 identified cross-link, K345K557, connects the NBD and SBD coming from 2 subunits; so, in this case, it was defined that this cross-link occurred between Chain A

(1st NBD) and Chain D (2nd SBD) and between Chain B (2nd NBD) and Chain C (1st SBD). In this manner, all cross-links were defined manually. After using the initial distance restraints file containing the cross-links with the ratio value above 0.5, the docking output wasn't satisfactory. It happened apparently due to the cross-link conflict: the chosen set of dimer cross-links cannot be accommodated on the one dimer model, which represents a major problem in a cross-link-guided docking. Therefore, it was decided to modify the initial set of cross-links and discard some of them – namely, K159K345 (also due to its relatively low average ratio, 0.52), and K108K567. The final set of cross-links used as the distance restraints used for HSC70 WT docking is enlisted in **Tab.7.**

HADDOCK created 5 structure clusters, each containing 2 best structures (10 in total). The output scores are depicted in **Tab.8.** In cluster 3, 4 and 5, the best HADDOCK score and Z-score values were achieved; in addition, the restraints violation energy is close to 0, meaning that the input cross-linking distance restraints were satisfied at the highest extent. Nonetheless, after closer visual inspection of all the created dimer models in PyMOL, the cluster 1 was chosen as the most plausible in terms of the mutual domain positions.

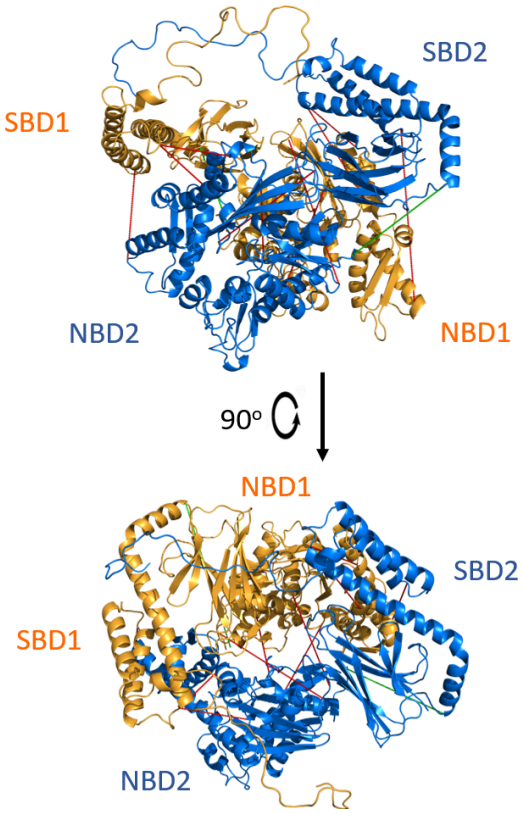
Tab. 7. The final set of the identified inter-cross-links used for the docking as distance restraints.

Cross-link site	Inter/intra ratio
K345K557	0.935 ± 0.036
K250K583	0.788 ± 0.155
K328K557	0.941 ± 0.062
K137K187/188	0.736 ± 0.077
K126K361	0.960 ± 0.089
0_K512	0 (the linker mimicker)

Tab. 8. The HADDOCK output scores for the created HSC70 WT dimer models.

Cluster number	1	2	3	4	5
HADDOCK score	-201.8 ± 44.9	-187.2 ± 24.9	-212.9 ±12.3	-203.6 ± 1.0	-244.5 ± 5.0
RMSD from the overall lowest-energy structure	39.8 ± 0.2	35.7 ± 0.3	26.0 ± 0.1	30.7 ± 0.1	28.7 ± 0.0
Van der Waals energy	-95.5 ± 14.2	-103.7 ± 19.2	-83.6 ± 11.5	-89.5 ± 4.6	-106.2 ± 13.5
Electrostatic energy	-813.7 ± 125.3	-829.0 ± 143.0	-899.6 ± 24.2	-872.3 ± 48.6	-1036.7 ± 66.3
Desolvation energy	44.0 ± 5.0	61.5 ± 2.1	50.6 ± 4.0	60.2 ± 4.1	69.1 ± 5.3
Restraints violation energy	12.4 ± 17.3	20.8 ± 20.8	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0
Buried Surface Area, Å²	4922.7 ± 796.1	5162.9 ± 962.4	5631.9 ± 689.6	5219.2 ± 38.7	6105.4 ± 325.3
Z-score	0.4	1.2	-0.2	0.3	-1.8

The input cross-links were mapped onto the cluster 1 model and the distances between the Cα atoms corresponding to the detected XL-sites were measured to find out whether the distance restraints were violated in the model or not (**Fig.32**). As can be seen, not all the distance restraints fulfill the 30 Å threshold; however, the mutual orientation of NBD and SBD coming from the 1 monomer is satisfactory, since the linker can be easily accommodated between them. The introduced intra-protein XL discussed above (0_K512) was used only to maintain this mutual domain orientation, and even though the distance restraint isn't fulfilled in this case, the length of the flexible interdomain linker fits in the given domain distance (the linker length is approximately 58 Å in the extended conformation). The cluster 1 model exhibits the perpendicular orientation of the HSC70 subunits with the several interfaces: 1) between the NBD1 and NBD2; 2) two interfaces between the given SBD and NBDs. In the latter case, an SBD is docked onto the surface of the cleft emerging from the NBD-NBD interface.



XL site	Distance, Å
K137K188	44.3
K188K137	44.0
K345K557	35.8
K557K345	20.1
K250K583	52.9
K583K250	25.0
K126K361	33.4
K361K126	31.9
K328K557	19.0
K557K328	18.9
0_K512_AB	43.3
0_K512_CD	37.5

Fig.32. The HSC70 WT dimer model created by HADDOCK 2.4 (cluster 1, model 1) and the measured Cα-Cα distance restraints (inter-XL – red, intra-XL – green). The domains originating from the same subunit are colored similarly (orange and blue, respectively). The mutual subunit position is perpendicular, with the interfaces between the NBDs and between the SBD and NBDs, where the corresponding SBD is accommodated in the cleft between the NBDs.

Discussion

During this project, several important goals have been achieved: expression and purification of the recombinant and isotopically labeled HSC70 on the minimal medium; tackling the ATP-dependent dimerization of HSC70 by mass photometry in time; performing the XL-MS experiments and the analysis of the acquired data; building the homology model of the ADP-bound HSC70 conformation; applying the distance restraints obtained from the XL-MS experiments to build a putative HSC70 dimer model by docking. At each step of this work, a lot of optimization was done and a number of important discoveries regarding the experiment conduction were made.

One important finding was that one of the storage buffer components, glycerol, negatively affects the cross-linking reactions involving NHS esters (e.g., DSS). Before optimization, the proteins were directly incubated and cross-linked in the storage buffer, which contains 10% glycerol (i.e., HKMG150 buffer). It led to the different number of identified XL-sites and non-

reproducibility of the replicates. It was speculated that the reason for this is the side reactivity of NHS esters with hydroxyl groups - the added cross-linker simply reacts with the more abundant glycerol and depletes, before reacting with the target proteins. It leads to t lower amounts of detected cross-linked peptides in mass spectra. It is shown for the initial HSC70 ^{14}N-WT + ^{15}N-WT XL-MS experiment (**Fig.33**). On the plot it's clearly visible that the presence of glycerol reduces the amount of detected XL-sites 2-3 fold. This discovery helped us to optimize the sample processing prior to the XL-reaction - namely, the addition of the buffer exchange step and removal of glycerol from the incubation buffer.

Then, it was definitely concluded that the presence of the reducing agent throughout the protein purification and the sample preparation is necessary. The artificial disulfide dimers can emerge during the purification process due to cysteine thiol group oxidation. TCEP, which was utilized in this work, has a several advantages over the standard reducing agents, β-mercaptoethanol or dithiothreitol: it doesn't contain thiol groups, which allows to avoid unwanted side reactions and conjugations with thiol side chains; it is irreversible and more effective than DTT or β-mercaptoethanol; it is very selective and reduces only cysteine thiol side groups (*Getz et al. 1999*). It suggests that TCEP is the perfect choice as a reducing agent, especially for the mass spectrometry samples. However, it was empirically discovered that TCEP severely deteriorates the quality of mass photometry measurements, namely increasing background noise levels. Thus, prior to mass photometry measurements, the sample storage buffer containing TCEP was exchanged, which allowed to successfully conduct the mass photometry experiments.

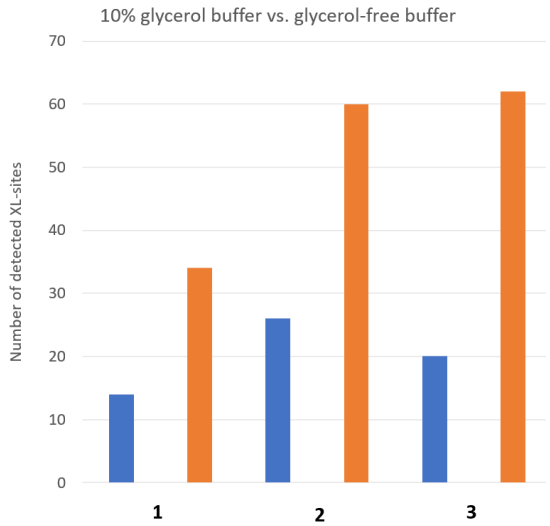


Fig.33. The negative effect of glycerol on the NHS-cross-linking reaction. Blue - replicates containing 10% glycerol, orange - the same replicates without glycerol.

The mass photometry results have clearly shown the differences in dimerization capabilities of the HSC70 mutants and WT, as well as unsimilar patterns of the dimer association and dissociation in time **(Fig.23)**. For the mixture containing only HSC70 WT, the observations perfectly fit the previous findings about the effect of ATP on Hsp70 homologs, which were described in the theoretical part of this work. Specifically, the ATP binding causes the conformational changes of HSC70, which ultimately leads to the dimer dissociation; however, as an ATP pool is running out due to the intrinsic HSC70 ATPase activity, the conformational equilibrium shifts towards the ADP-bound form and the dimer reassembles back. However, neither SBD-mutant, nor linker-mutant aren't responsive to the ATP addition - certain fluctuations in dimer/monomer ratio can be seen, but not the certain pattern shared by HSC70 WT. Therefore, the mutations in SBD and the linker disrupt the characteristic allosteric cycle of HSC70, which normally allows the reversible and ATP-dependent conformational changes. At the same time, these perturbations in the allosteric cycle negatively affect the dimerization abilities, which explains the lower initial dimer level at 0 min time point. In the case of the mixed experiments at 0 min time point, the simple explanation for the acquired mass photometry data lies within the fact that the majority of protein molecules in solution resides in the ADP-bound conformation within the preformed dimer population. In this situation, when no conformational changes occur, the different protein molecules won't begin to interact with each other to form a "heterodimer". For instance, the WT:linker-mutant solution contains the population of WT dimer and the separate population of linker-mutant dimer, but almost no WT+linker-mutant heterodimer. Nevertheless, in the equimolar mixture of WT and SBD-mutant, it can be concluded that the heterodimer can still be formed at the level comparable to the WT alone sample. The same can't be described for the WT+linker-mutant and SBD-mutant+linker mutant experiments, where the dimer/monomer ratio at the final time point is 2-3x lower than in the case of WT alone or WT+SBD-mutant samples. Therefore, WT is still able to interact with the SBD-mutant at the same extent, but not with the linker-mutant. Neither SBD-mutant can't interact with linker-mutant so extensively. It was suggested that polar residues introduced to the linker may reside at the subunit interface in HSC70 dimer and negatively affect the native non-covalent interactions.

In order to confirm the hypotheses derived from the mass photometry data, numerous XL-MS experiments were performed. During the analysis of the cross-linked species by SDS-PAGE, the multiple intermediate bands between the monomeric and the dimeric bands were detected (**Fig.24)**. These intermediate species are considered to be intra-XL HSC70 molecules, whose conformations in the cross-linked state provide a diminished electrophoretic mobility (e.g., less

compact). Then, after the XL-MS data acquisition and processing, the similarly uneven sequence coverage was detected by LinX (**Fig.25**). Such irregularity in sequence coverage can be attributed to sterical hindrance for the trypsin and LysC binding in these less covered regions within the cross-linked HSC70 molecules. To make a definitive conclusion, the control sample of the non-cross-linked and digested HSC70 has to be measured by LC-MS; unfortunately, it wasn't done. However, the complete HSC70 sequence is covered in each replicate.

In the XL-MS datasets, it was observed that in the certain experiments (i.e.,^{14}N-WT + ^{15}N-WT, WT+SBD-mutant, ^{14}N-SBD-mutant + ^{15}N-SBD-mutant), the intra/inter ratio distribution across the datasets aren't uniform and it's easy to notice the apparent discrete distribution of intra- and intermolecular restraints (**Fig.29A,B,C**). Therefore, it was decided to apply a 0.5 ratio threshold, to attempt to categorize the XL-sites according to their ratio value. As was noticed earlier, for the ^{14}N-WT + ^{15}N-WT, WT+SBD-mutant, ^{14}N-SBD-mutant + ^{15}N-SBD-mutant mixtures, some XL-sites have the ratio value above 0.5 (from 0.8 to 1.2) and some have it below 0.5 (0.0-0.4). This categorization was made based on the fact that multiple dimer species can be present in the solution at the same time. It is possible that these 2 categories of XL-sites (above 0.5 ratio - "rather inter/dimer", and below 0.5 - "rather intra/monomeric") can be assigned to 2 different populations of dimer (e.g., the one more abundant and stable, and the one less abundant and stable). Then, for the discussed experiments, the categorized XL-sites were mapped onto the HSC70 sequences and ADP-bound HSC70 structural model created by homology modeling (**Fig.34**). In the ^{14}N-WT + ^{15}N-WT and ^{14}N-SBD-mutant + ^{15}N-SBD-mutant samples, the same outcome was obtained: if the XL-sites with the ratio above 0.5 are mapped onto the structures, the mutual orientation of the subunits is antiparallel; for the mapped XL-sites with the ratio below 0.5, it's parallel. Therefore, it can be concluded that HSC70 WT can adopt the same dimer configuration as SBD-mutant; in addition, the antiparallel model reported by Morgner et al. was confirmed. In the WT+SBD-mutant sample, in the above 0.5 category, some oriented XL-sites appear, indicating the dimer configuration is non-symmetrical. The mapping on the HSC70 and SBD-mutant model provided an orientation close to a "substrate-like" fashion dimer model, suggested earlier by Chang et al. and C.-C. Wu et al.; the mapping of below 0.5 XL-sites provided a parallel model, as in the previous cases. For the linker-mutant samples, we have detected a large number of XL-sites, where the majority has a ratio value below 0.5. Therefore, the 0.5 threshold isn't applicable in these experiments, and perhaps other categorization should be taken into consideration.

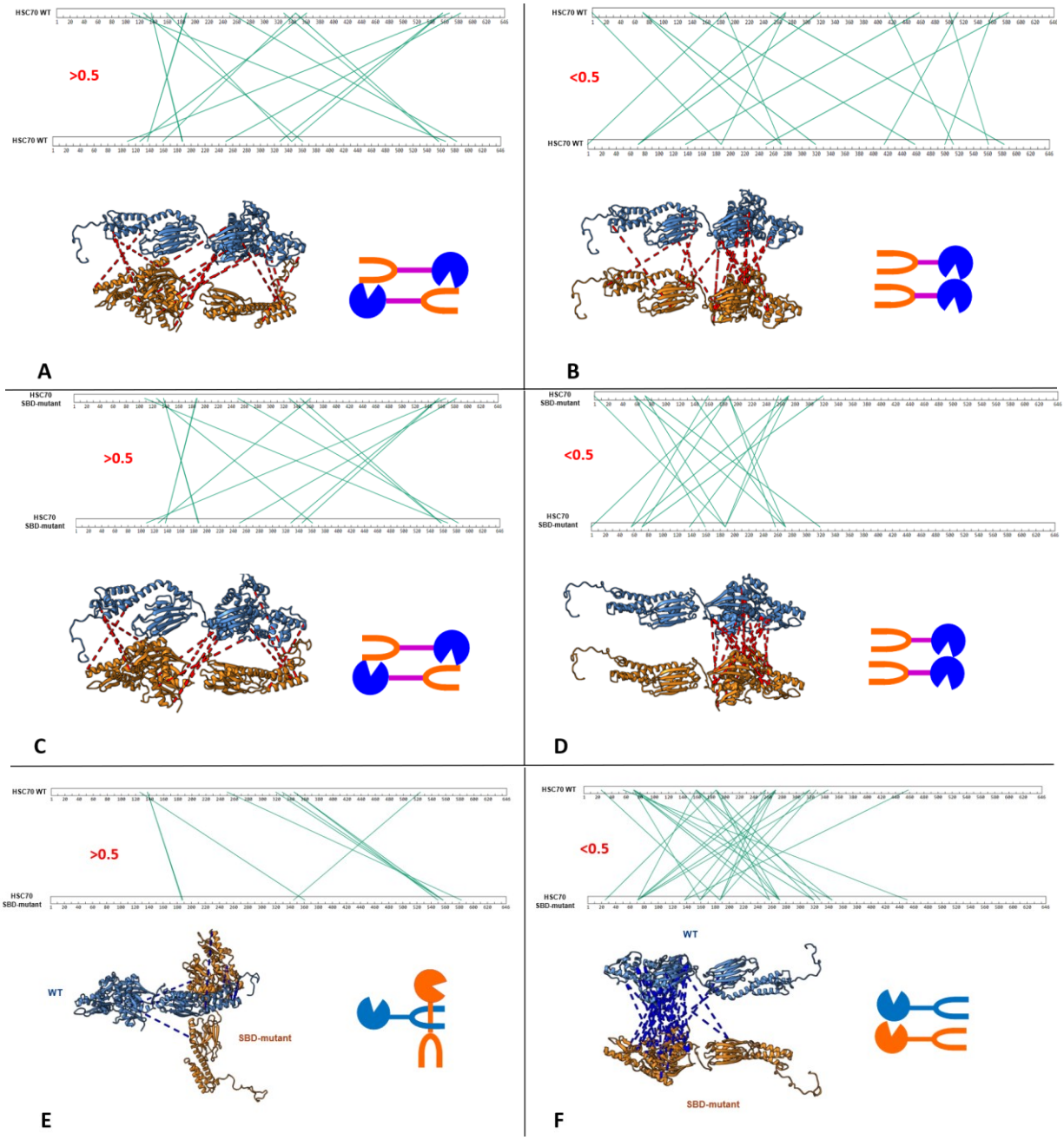
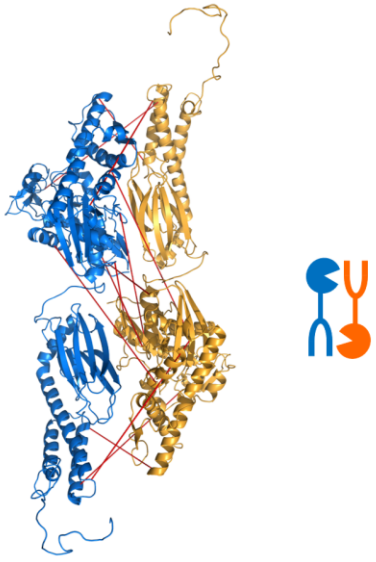


Fig.34. The mapping of the separated groups of XL-sites (above and below 0.5 inter/intra ratio). **A,B.** ^{14}N-WT + ^{15}N-WT; **C,D.** ^{14}N-SBD-mutant + ^{15}N-SBD-mutant; **E,F.** WT+SBD-mutant.

These “parallel dimer” models are hypothetical and are based on the mapping of “rather intra” XL-sites onto the HSC70 molecules, so it’s possible that it represents only a minor fraction of an HSC70 dimer population. It is worth to note that a lesser amount of the high-inter/intra ratio cross-link found in these experiments advocates for the existence of a stable dimer in the case of WT and SBD-mutant. On the other hand, in the linker-mutant experiments, the large number of the low-ratio XL-sites may be an evidence of many weak and transient interactions occurring between the linker-mutant, WT and SBD-mutant molecules. Nevertheless, further investigation

and the complementary data are required (e.g., native ion mobility MS, NMR, cryoEM) to come to the definitive conclusion.

The XL-sites identified in ^{14}N-WT + ^{15}N-WT were used for the HSC70 WT dimer docking; only the XL-sites with the ratio above 0.5 were taken (**Fig.34A**). The initial approach for the docking included 2 HSC70 WT monomers, defining the linker region as fully flexible. Unfortunately, all the attempts utilizing this approach were failed, since the HADDOCK script kept the linker rigid and treated an HSC70 monomer as a rigid body. The initial model didn't satisfy the XL-distance restraints at all (**Fig.35**). Even though the mutual antiparallel position of the subunits was maintained, the huge violation of the distance restraints led us to the approach for docking of flexible multidomain protein described above. In addition, several changes in the used distance restraints were made. Arbitrarily, a couple of XL-sites with higher distance restraint violation were removed (namely, K159K345, K345K159, K108K567), also due to their lower intensity in the mass spectra. In addition, the intra XL-site (connecting N-terminus and K512) was added to keep the relative position of the separated domains. All these adjustments allowed to successfully utilize the aforementioned docking approach and build a dimer model with partially fulfilled distance restraints (**Fig.32.**). Again, complementary structural data are necessary to make a certain conclusion about the HSC70 dimer structure and configuration. However, the created model is a starting point for further refinement and adjustments. The upcoming native ion mobility MS measurements will allow to define an ion mobility profile and the collision cross-section of the HSC70 dimer. It will also indicate the possible presence of the multiple HSC70 dimer species in the solution.



XL site	Distance, Å
K137K188	38.4
K188K137	38.0
K345K557	38
K557K345	38.6
K159K345	63.4
K345K159	62.4
K250K583	37.8
K583K250	34.9
K126K361	54
K361K126	54
K328K557	57.6
K557K328	58
K108K567	51.6
K567K108	54

Fig.35. The initial HSC70 WT dimer model created by docking. None of the used distance restraints derived from the XL-MS data wasn't fulfilled. The red XL-sites were removed in the final distance restraint file.

Conclusion

1. I successfully overexpressed and purified the recombinant human HSC70 WT, SBD-mutant and linker-mutant in *E.coli* on the minimal medium, both in the ^{14}N- and ^{15}N-forms;

2. I tackled the kinetics of ATP-dependent dimerization for HSC70 WT and mutants by mass photometry in time and optimized the conditions for the cross-linking reaction;

3. I cross-linked, processed and analyzed the HSC70 samples by LC-MS; in addition, I managed to analyze and classify the acquired LC-MS data and visualize it on the HSC70 model and sequence.

4. I created a homology model of ADP-bound HSC70;

5. I used the distance restraints obtained from the XL-MS experiments to create the HSC70 WT homodimer model by molecular docking.

References

1. Agarraberes, Fernando A, and J Fred Dice. “Lysosomal Membrane Chaperone Complex,” 2001. https://doi.org/10.1242/jcs.114.13.2491.

2. Ahn, Sang-Gun, Soo-A Kim, Jung-Hoon Yoon, and Panayiotis Vacratsis. “Heat-Shock Cognate 70 Is Required for the Activation of Heat-Shock Factor 1 in Mammalian Cells.” *Biochemical Journal* 392, no. 1 (November 15, 2005): 145–52. https://doi.org/10.1042/BJ20050412.

3. Arispe, Nelson, Michael Doh, and Antonio De Maio. “Lipid Interaction Differentiates the Constitutive and Stress-Induced Heat Shock Proteins Hsc70 and Hsp70.” *Cell Stress & Chaperones* 7, no. 4 (2002): 330. https://doi.org/10.1379/1466-1268(2002)007<0330:LIDTCA>2.0.CO;2.

4. Arrigo, André-Patrick, and W. E. G. Müller, eds. *Small Stress Proteins*. Vol. 28. Progress in Molecular and Subcellular Biology. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. https://doi.org/10.1007/978-3-642-56348-5.

5. Balchin, David, Manajit Hayer-Hartl, and F. Ulrich Hartl. “In Vivo Aspects of Protein Folding and Quality Control.” *Science* 353, no. 6294 (July 2016): aac4354. https://doi.org/10.1126/science.aac4354.

6. Benaroudj, Nadia, GCrard Batelier, Franqoise Triniolles, and Moncef M Ladjimi. “Self-Association of the Molecular Chaperone HSC70t,” n.d.

7. Bercovich, Beatrice, Ilana Stancovski, Arie Mayer, Nava Blumenfeld, Andrei Laszlo, Alan L. Schwartz, and Aaron Ciechanover. “Ubiquitin-Dependent Degradation of Certain Protein Substrates in Vitro Requires the Molecular Chaperone Hsc70.” *Journal of Biological Chemistry* 272, no. 14 (April 1997): 9002–10. https://doi.org/10.1074/jbc.272.14.9002.

8. Blond-Elguindi, S., A.M. Fourie, J.F. Sambrook, and M.J. Gething. “Peptide-Dependent Stimulation of the ATPase Activity of the Molecular Chaperone BiP Is the Result of Conversion of Oligomers to Active Monomers.” *Journal of Biological Chemistry* 268, no. 17 (June 1993): 12730–35. https://doi.org/10.1016/S0021-9258(18)31449-2.

9. Blumberg, Shmaryahu, and Bert L. Vallee. “Superactivation of Thermolysin by Acylation with Amino Acid N-Hydroxysuccinimide Esters.” *Biochemistry* 14, no. 11 (June 1, 1975): 2410–19. https://doi.org/10.1021/bi00682a022.

10. Brehmer, Dirk, Stefan Rüdiger, Claudia S Gässler, Dagmar Klostermeier, Lars Packschies, Jochen Reinstein, Matthias P Mayer, and Bernd Bukau. “Tuning of Chaperone Activity of Hsp70 Proteins by Modulation of Nucleotide Exchange.” *Nature Structural Biology* 8, no. 5 (2001). https://doi.org/10.1038/87588.

11. Bryngelson, Joseph D., José Nelson Onuchic, Nicholas D. Socci, and Peter G. Wolynes. “Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis.” *Proteins: Structure, Function, and Genetics* 21, no. 3 (March 1995): 167–95. https://doi.org/10.1002/prot.340210302.

12. Buczynski, Greg, Sergey V. Slepenkov, Michael G. Sehorn, and Stephan N. Witt. “Characterization of a Lidless Form of the Molecular Chaperone DnaK.” *Journal of Biological Chemistry* 276, no. 29 (July 2001): 27231–36. https://doi.org/10.1074/jbc.M100237200.

13. Chang, Yi-Wei, Yuh-Ju Sun, Chung Wang, and Chwan-Deng Hsiao. “Crystal Structures of the 70-kDa Heat Shock Proteins in Domain Disjoining Conformation.” *Journal of Biological Chemistry* 283, no. 22 (May 2008): 15502–11. https://doi.org/10.1074/jbc.M708992200.

14. Chappell, Thomas G., William J. Welch, David M. Schlossman, Karen B. Palter, Milton J. Schlesinger, and James E. Rothman. “Uncoating ATPase Is a Member of the 70 Kilodalton Family of Stress Proteins.” *Cell* 45, no. 1 (April 1986): 3–13. https://doi.org/10.1016/0092-8674(86)90532-5.

15. Cheetham, Michael E., and Avrom J. Caplan. “Structure, Function and Evolution of DnaJ: Conservation and Adaptation of Chaperone Function.” *Cell Stress & Chaperones* 3, no. 1 (1998): 28. https://doi.org/10.1379/1466-1268(1998)003<0028:SFAEOD>2.3.CO;2.

16. Chen, Bin, William H. Piel, Liming Gui, Elspeth Bruford, and Antónia Monteiro. “The HSP90 Family of Genes in the Human Genome: Insights into Their Divergence and Evolution.” *Genomics* 86, no. 6 (December 2005): 627–37. https://doi.org/10.1016/j.ygeno.2005.08.012.

17. Chen, Zhuo Angel, Anass Jawhari, Lutz Fischer, Claudia Buchen, Salman Tahir, Tomislav Kamenski, Morten Rasmussen, et al. “Architecture of the RNA Polymerase II–TFIIF Complex Revealed by Cross-Linking and Mass Spectrometry.” *The EMBO Journal* 29, no. 4 (February 17, 2010): 717–26. https://doi.org/10.1038/emboj.2009.401.

18. Clark, John I, and Paul J Muchowski. “Small Heat-Shock Proteins and Their Potential Role in Human Disease.” *Current Opinion in Structural Biology* 10, no. 1 (February 2000): 52–59. https://doi.org/10.1016/S0959-440X(99)00048-2.

19. Cuervo, Ana Maria, and J. Fred Dice. “A Receptor for the Selective Uptake and Degradation of Proteins by Lysosomes.” *Science* 273, no. 5274 (July 26, 1996): 501–3. https://doi.org/10.1126/science.273.5274.501.

20. De Hoffmann, Edmond, Vincent Stroobant, and Edmond De Hoffmann. *Mass Spectrometry: Principles and Applications*. 3. ed., Reprinted. Chichester Weinheim: Wiley, 2011.

21. Deshaies, Raymond J., Bruce D. Koch, Margaret Werner-Washburne, Elizabeth A. Craig, and Randy Schekman. “A Subfamily of Stress Proteins Facilitates Translocation of Secretory and Mitochondrial Precursor Polypeptides.” *Nature* 332, no. 6167 (April 1988): 800–805. https://doi.org/10.1038/332800a0.

22. Deuerling, Elke, Agnes Schulze-Specking, Toshifumi Tomoyasu, Axel Mogk, and Bernd Bukau. “Trigger Factor and DnaK Cooperate in Folding of Newly Synthesized Proteins.” *Nature* 400, no. 6745 (August 1999): 693–96. https://doi.org/10.1038/23301.

23. Deville, Célia, Marta Carroni, Kamila B. Franke, Maya Topf, Bernd Bukau, Axel Mogk, and Helen R. Saibil. “Structural Pathway of Regulated Substrate Transfer and Threading through an Hsp100 Disaggregase.” *Science Advances* 3, no. 8 (August 4, 2017): e1701726. https://doi.org/10.1126/sciadv.1701726.

24. Diehl, J. Alan, Wensheng Yang, Ronald A. Rimerman, Hua Xiao, and Andrew Emili. “Hsc70 Regulates Accumulation of Cyclin D1 and Cyclin D1-Dependent Protein Kinase.” *Molecular and Cellular Biology* 23, no. 5 (March 1, 2003): 1764–74. https://doi.org/10.1128/MCB.23.5.1764-1774.2003.

25. Dill, Ken A., Sarina Bromberg, Kaizhi Yue, Hue Sun Chan, Klaus M. Ftebig, David P. Yee, and Paul D. Thomas. “Principles of Protein Folding - A Perspective from Simple Exact Models.” *Protein Science* 4, no. 4 (December 31, 2008): 561–602. https://doi.org/10.1002/pro.5560040401.

26. Ditzel, Lars, Jan Löwe, Daniela Stock, Karl-Otto Stetter, Harald Huber, Robert Huber, and Stefan Steinbacher. “Crystal Structure of the Thermosome, the Archaeal Chaperonin and Homolog of CCT.” *Cell* 93, no. 1 (April 1998): 125–38. https://doi.org/10.1016/S0092-8674(00)81152-6.

27. Dobson, Christopher M., Andrej Šali, and Martin Karplus. “Protein Folding: A Perspective from Theory and Experiment.” *Angewandte Chemie International Edition* 37, no. 7 (April 20, 1998): 868–93. https://doi.org/10.1002/(SICI)1521-3773(19980420)37:7<868::AID-ANIE868>3.0.CO;2-H.

28. Doyle, Shannon M., and Sue Wickner. “Hsp104 and ClpB: Protein Disaggregating Machines.” *Trends in Biochemical Sciences* 34, no. 1 (January 2009): 40–48. https://doi.org/10.1016/j.tibs.2008.09.010.

29. Dubaquie, Y. “Identification of Invivo Substrates of the Yeast Mitochondrial Chaperonins Reveals Overlapping but Non-Identical Requirement for Hsp60 and Hsp10.” *The EMBO Journal* 17, no. 20 (October 15, 1998): 5868–76. https://doi.org/10.1093/emboj/17.20.5868.

30. Dwornczak, Bernd, and Marc-Edouard Mirault. “Structure and Expression of a Human Gene Coding for a 71 Kd Heat Shock ‘Cognate’ Protein.” *Nucleic Acids Research* 15, no. 13 (1987): 5181–97. https://doi.org/10.1093/nar/15.13.5181.

31. Dyachenko, Andrey, Ranit Gruber, Liat Shimon, Amnon Horovitz, and Michal Sharon. “Allosteric Mechanisms Can Be Distinguished Using Structural Mass Spectrometry.” *Proceedings of the National Academy of Sciences* 110, no. 18 (April 30, 2013): 7235–39. https://doi.org/10.1073/pnas.1302395110.

32. Ellis, R J. “Molecular Chaperones.” *MOLECULAR CHAPERONES* (July 1991). https://doi.org/10.1146/annurev.bi.60.070191.001541.

33. Englander, Joan J., Jose R. Rogero, and S.Walter Englander. “Protein Hydrogen Exchange Studied by the Fragment Separation Method.” *Analytical Biochemistry* 147, no. 1 (May 1985): 234–44. https://doi.org/10.1016/0003-2697(85)90033-8.

34. Erzberger, Jan P., Florian Stengel, Riccardo Pellarin, Suyang Zhang, Tanja Schaefer, Christopher H.S. Aylett, Peter Cimermančič, et al. “Molecular Architecture of the 40S·eIF1·eIF3 Translation Initiation Complex.” *Cell* 158, no. 5 (August 2014): 1123–35. https://doi.org/10.1016/j.cell.2014.07.044.

35. Esser, Claudia, Martin Scheffner, and Jörg Höhfeld. “The Chaperone-Associated Ubiquitin Ligase CHIP Is Able to Target P53 for Proteasomal Degradation.” *Journal of Biological Chemistry* 280, no. 29 (July 2005): 27443–48. https://doi.org/10.1074/jbc.M501574200.

36. Fayet, O, T Ziegelhoffer, and C Georgopoulos. “The groES and groEL Heat Shock Gene Products of Escherichia Coli Are Essential for Bacterial Growth at All Temperatures.” *Journal of Bacteriology* 171, no. 3 (March 1989): 1379–85. https://doi.org/10.1128/jb.171.3.1379-1385.1989.

37. Feder, J H, J M Rossi, J Solomon, N Solomon, and S Lindquist. “The Consequences of Expressing Hsp70 in Drosophila Cells at Normal Temperatures.” *Genes & Development* 6, no. 8 (August 1992): 1402–13. https://doi.org/10.1101/gad.6.8.1402.

38. Ferbitz, Lars, Timm Maier, Holger Patzelt, Bernd Bukau, Elke Deuerling, and Nenad Ban. “Trigger Factor in Complex with the Ribosome Forms a Molecular Cradle for Nascent Proteins” 431 (2004). https://doi.org/10.1038/nature02899.

39. Ferreiro, Diego U., Joseph A. Hegler, Elizabeth A. Komives, and Peter G. Wolynes. “Localizing Frustration in Native Proteins and Protein Assemblies.” *Proceedings of the National Academy of Sciences* 104, no. 50 (December 11, 2007): 19819–24. https://doi.org/10.1073/pnas.0709915104.

40. Flaherty, K M, D B McKay, W Kabsch, and K C Holmes. “Similarity of the Three-Dimensional Structures of Actin and the ATPase Fragment of a 70-kDa Heat Shock Cognate Protein.” *Proceedings of the National Academy of Sciences* 88, no. 11 (June 1991): 5041–45. https://doi.org/10.1073/pnas.88.11.5041.

41. Flaherty, Kevin M., Camilla DeLuca-Flaherty, and David B. McKay. “Three-Dimensional Structure of the ATPase Fragment of a 70K Heat-Shock Cognate Protein.” *Nature* 346, no. 6285 (August 1990): 623–28. https://doi.org/10.1038/346623a0.

42. Flynn, Gregory C., Jan Pohl, Mark T. Flocco, and James E. Rothman. “Peptide-Binding Specificity of the Molecular Chaperone BiP.” *Nature* 353, no. 6346 (October 1991): 726–30. https://doi.org/10.1038/353726a0.

43. Fojtík, Lukáš, Jan Fiala, Petr Pompach, Josef Chmelík, Václav Matoušek, Petr Beier, Zdeněk Kukačka, and Petr Novák. “Fast Fluoroalkylation of Proteins Uncovers the Structure and Dynamics of Biological Macromolecules.” *Journal of the American Chemical Society* 143, no. 49 (December 15, 2021): 20670–79. https://doi.org/10.1021/jacs.1c07771.

44. Freiden, P.J., J.R. Gaut, and L.M. Hendershot. “Interconversion of Three Differentially Modified and Assembled Forms of BiP.” *The EMBO Journal* 11, no. 1 (January 1992): 63–70. https://doi.org/10.1002/j.1460-2075.1992.tb05028.x.

45. Frydman, Judith, and F. Ulrich Hartl. “Principles of Chaperone-Assisted Protein Folding: Differences Between in Vitro and in Vivo Mechanisms.” *Science* 272, no. 5267 (June 7, 1996): 1497–1502. https://doi.org/10.1126/science.272.5267.1497.

46. Frydman, Judith, Elmar Nimmesgern, Kenzo Ohtsuka, and F. Ulrich Hartl. “Folding of Nascent Polypeptide Chains in a High Molecular Mass Assembly with Molecular Chaperones.” *Nature* 370, no. 6485 (July 1994): 111–17. https://doi.org/10.1038/370111a0.

47. Gehrmann, M, J Marienhagen, H Eichholtz-Wirth, E Fritz, J Ellwart, M Jäättelä, T Zilch, and G Multhoff. “Dual Function of Membrane-Bound Heat Shock Protein 70 (Hsp70), Bag-4, and Hsp40: Protection against Radiation-Induced Effects and Target Structure for Natural Killer Cells.” *Cell Death & Differentiation* 12, no. 1 (January 2005): 38–51. https://doi.org/10.1038/sj.cdd.4401510.

48. Genevaux, Pierre, Costa Georgopoulos, and William L. Kelley. “The Hsp70 Chaperone Machines of *Escherichia Coli*: A Paradigm for the Repartition of Chaperone Functions.” *Molecular Microbiology* 66, no. 4 (November 2007): 840–57. https://doi.org/10.1111/j.1365-2958.2007.05961.x.

49. Gething, Mary-Jane, and Joseph Sambrook. “Protein Folding in the Cell.” *Nature* 355, no. 6355 (January 1992): 33–45. https://doi.org/10.1038/355033a0.

50. Getz, Elise Burmeister, Ming Xiao, Tania Chakrabarty, Roger Cooke, and Paul R. Selvin. “A Comparison between the Sulfhydryl Reductants Tris(2-Carboxyethyl)Phosphine and Dithiothreitol for Use in Protein Biochemistry.” *Analytical Biochemistry* 273, no. 1 (August 1999): 73–80. https://doi.org/10.1006/abio.1999.4203.

51. Goffin, Laurence, and Costa Georgopoulos. “Genetic and Biochemical Characterization of Mutations Affecting the Carboxy-terminal Domain of the *Escherichia Coli* Molecular Chaperone DnaJ.” *Molecular Microbiology* 30, no. 2 (October 1998): 329–40. https://doi.org/10.1046/j.1365-2958.1998.01067.x.

52. Gong, Weibin, Wanhui Hu, Linan Xu, Huiwen Wu, Si Wu, Hong Zhang, Jinfeng Wang, Gary W. Jones, and Sarah Perrett. “The C-Terminal GGAP Motif of Hsp70 Mediates Substrate Recognition and Stress Response in Yeast.” *Journal of Biological Chemistry* 293, no. 46 (November 2018): 17663–75. https://doi.org/10.1074/jbc.RA118.002691.

53. Graham, Martin, Colin Combe, Lars Kolbowski, and Juri Rappsilber. “xiView: A Common Platform for the Downstream Analysis of Crosslinking Mass Spectrometry Data.” Preprint. Molecular Biology, February 26, 2019. https://doi.org/10.1101/561829.

54. Gregory, John D. “The Stability of N-Ethylmaleimide and Its Reaction with Sulfhydryl Groups.” *Journal of the American Chemical Society* 77, no. 14 (July 1955): 3922–23. https://doi.org/10.1021/ja01619a073.

55. Gribaldo, Simonetta, Valentina Lumia, Roberta Creti, Everly Conway De Macario, Annamaria Sanangelantoni, and Piero Cammarano. “Discontinuous Occurrence of the *Hsp70* (*dnaK*) Gene among *Archaea* and Sequence Features of HSP70 Suggest a Novel Outlook on Phylogenies Inferred from This Protein.” *Journal of Bacteriology* 181, no. 2 (January 15, 1999): 434–43. https://doi.org/10.1128/JB.181.2.434-443.1999.

56. Guan, Jing-Qu, and Mark R. Chance. “Footprinting Methods to Examine the Structure and Dynamics of Proteins.” In *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, edited by Robert A. Meyers, mcb.200300025. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2006. https://doi.org/10.1002/3527600906.mcb.200300025.

57. Guan, Jing-Qu, and Mark R. Chance. “Structural Proteomics of Macromolecular Assemblies Using Oxidative Footprinting and Mass Spectrometry.” *Trends in Biochemical Sciences* 30, no. 10 (October 2005): 583–92. https://doi.org/10.1016/j.tibs.2005.08.007.

58. Haas, Ingrid G., and Matthias Wabl. “Immunoglobulin Heavy Chain Binding Protein.” *Nature* 306, no. 5941 (November 1983): 387–89. https://doi.org/10.1038/306387a0.

59. Harrison, Celia J, Manajit Hayer-Hartl, and Maurizio Di Liberto. “Crystal Structure of the Nucleotide Exchange Factor GrpE Bound to the ATPase Domain of the Molecular Chaperone DnaK,” 1997. https://doi.org/10.1126/science.276.5311.431.

60. Hartl, F. Ulrich, Andreas Bracher, and Manajit Hayer-Hartl. “Molecular Chaperones in Protein Folding and Proteostasis.” *Nature* 475, no. 7356 (July 2011): 324–32. https://doi.org/10.1038/nature10317.

61. Haslbeck, Martin, Nathalie Braun, Thusnelda Stromer, Bettina Richter, Natascha Model, Sevil Weinkauf, and Johannes Buchner. “Hsp42 Is the General Small Heat Shock Protein in the Cytosol of Saccharomyces Cerevisiae.” *The EMBO Journal* 23, no. 3 (February 11, 2004): 638–49. https://doi.org/10.1038/sj.emboj.7600080.

62. Haslbeck, Martin, Titus Franzmann, Daniel Weinfurtner, and Johannes Buchner. “Some like It Hot: The Structure and Function of Small Heat-Shock Proteins.” *Nature Structural & Molecular Biology* 12, no. 10 (October 2005): 842–46. https://doi.org/10.1038/nsmb993.

63. Heck, Albert J R. “Native Mass Spectrometry: A Bridge between Interactomics and Structural Biology.” *Nature Methods* 5, no. 11 (November 2008): 927–33. https://doi.org/10.1038/nmeth.1265.

64. Hendrick, J P. “Molecular Chaperone Functions of Heat-Shock Proteins.” *E. Coli*, 1993. https://doi.org/10.1146/annurev.bi.62.070193.002025.

65. Heuvel, Robert Hh Van Den, and Albert Jr Heck. “Native Protein Mass Spectrometry: From Intact Oligomers to Functional Machineries.” *Current Opinion in Chemical Biology* 8, no. 5 (October 2004): 519–26. https://doi.org/10.1016/j.cbpa.2004.08.006.

66. Hoffmann, Anja, Frieder Merz, Anna Rutkowska, Beate Zachmann-Brand, Elke Deuerling, and Bernd Bukau. “Trigger Factor Forms a Protective Shield for Nascent Polypeptides at the Ribosome.” *Journal of Biological Chemistry* 281, no. 10 (March 2006): 6539–45. https://doi.org/10.1074/jbc.M512345200.

67. Honorato, Rodrigo V., Panagiotis I. Koukos, Brian Jiménez-García, Andrei Tsaregorodtsev, Marco Verlato, Andrea Giachetti, Antonio Rosato, and Alexandre M. J. J. Bonvin. “Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem.” *Frontiers in Molecular Biosciences* 8 (July 28, 2021): 729513. https://doi.org/10.3389/fmolb.2021.729513.

68. Horwich, Arthur L., Wayne A. Fenton, Eli Chapman, and George W. Farr. “Two Families of Chaperonin: Physiology and Mechanism.” *Annual Review of Cell and Developmental Biology* 23, no. 1 (November 1, 2007): 115–45. https://doi.org/10.1146/annurev.cellbio.23.090506.123555.

69. Huang, Zhong, Lei Nie, Min Xu, and Xiao-Hong Sun. “Notch-Induced E2A Degradation Requires CHIP and Hsc70 as Novel Facilitators of Ubiquitination.” *Molecular and Cellular Biology* 24, no. 20 (October 1, 2004): 8951–62. https://doi.org/10.1128/MCB.24.20.8951-8962.2004.

70. Hvidt, Aase, and K. Linderstrøm-Lang. “The Kinetics of the Deuterium Exchange of Insulin with D2O. An Amendment.” *Biochimica et Biophysica Acta* 16 (January 1955): 168–69. https://doi.org/10.1016/0006-3002(55)90200-6.

71. Jahn, Thomas R., and Sheena E. Radford. “The Yin and Yang of Protein Folding: The Yin and Yang of Protein Folding.” *FEBS Journal* 272, no. 23 (November 10, 2005): 5962–70. https://doi.org/10.1111/j.1742-4658.2005.05021.x.

72. Jakob, U., M. Gaestel, K. Engel, and J. Buchner. “Small Heat Shock Proteins Are Molecular Chaperones.” *Journal of Biological Chemistry* 268, no. 3 (January 1993): 1517–20. https://doi.org/10.1016/S0021-9258(18)53882-5.

73. Jennebach, Stefan, Franz Herzog, Ruedi Aebersold, and Patrick Cramer. “Crosslinking-MS Analysis Reveals RNA Polymerase I Domain Architecture and Basis of rRNA Cleavage.” *Nucleic Acids Research* 40, no. 12 (July 2012): 5591–5601. https://doi.org/10.1093/nar/gks220.

74. Jiang, Yajun, Paolo Rossi, and Charalampos G Kalodimos. “Structural Basis for Client Recognition and Activity of Hsp40 Chaperones,” 2019. https://doi.org/10.1126/science.aax1280.

75. Kao, Athit, Arlo Randall, Yingying Yang, Vishal R Patel, Wynne Kandur, Shenheng Guan, Scott D Rychnovsky, Pierre Baldi, and Lan Huang. “Mapping the Structural Topology of the Yeast 19S Proteasomal

Regulatory Particle Using Chemical Cross-Linking and Probabilistic Modeling*", 2012. https://doi.org/10.1074/mcp.M112.018374.

76. Karaca, Ezgi, and Alexandre M.J.J. Bonvin. "A Multidomain Flexible Docking Approach to Deal with Large Conformational Changes in the Modeling of Biomolecular Complexes." *Structure* 19, no. 4 (April 2011): 555–65. https://doi.org/10.1016/j.str.2011.01.014.

77. Kelley, William L. "The J-Domain Family and the Recruitment of Chaperone Power." *Trends in Biochemical Sciences* 23, no. 6 (June 1998): 222–27. https://doi.org/10.1016/S0968-0004(98)01215-8.

78. Kerner, Michael J., Dean J. Naylor, Yasushi Ishihama, Tobias Maier, Hung-Chun Chang, Anna P. Stines, Costa Georgopoulos, et al. "Proteome-Wide Analysis of Chaperonin-Dependent Protein Folding in Escherichia Coli." *Cell* 122, no. 2 (July 2005): 209–20. https://doi.org/10.1016/j.cell.2005.05.028.

79. Kim, Kyeong Kyu, Rosalind Kim, and Sung-Hou Kim. "Crystal Structure of a Small Heat-Shock Protein" 394 (1998). https://doi.org/10.1038/29106.

80. King, Camille, Evan Eisenberg, and Lois Greene. "Polymerization of 70-kDa Heat Shock Protein by Yeast DnaJ in ATP." *Journal of Biological Chemistry* 270, no. 38 (September 1995): 22535–40. https://doi.org/10.1074/jbc.270.38.22535.

81. Kityk, Roman, Jürgen Kopp, and Matthias P. Mayer. "Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones." *Molecular Cell* 69, no. 2 (January 2018): 227-237.e4. https://doi.org/10.1016/j.molcel.2017.12.003.

82. Kityk, Roman, Jürgen Kopp, Irmgard Sinning, and Matthias P. Mayer. "Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones." *Molecular Cell* 48, no. 6 (December 2012): 863–74. https://doi.org/10.1016/j.molcel.2012.09.023.

83. Kityk, Roman, Markus Vogel, Rainer Schlecht, Bernd Bukau, and Matthias P. Mayer. "Pathways of Allosteric Regulation in Hsp70 Chaperones." *Nature Communications* 6, no. 1 (September 18, 2015): 8308. https://doi.org/10.1038/ncomms9308.

84. Konermann, Lars, Jingxi Pan, and Yu-Hong Liu. "Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics." *Chem. Soc. Rev.* 40, no. 3 (2011): 1224–34. https://doi.org/10.1039/C0CS00113A.

85. Konijnenberg, A., A. Butterer, and F. Sobott. "Native Ion Mobility-Mass Spectrometry and Related Methods in Structural Biology." *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1834, no. 6 (June 2013): 1239–56. https://doi.org/10.1016/j.bbapap.2012.11.013.

86. Kose, Shingo, Maiko Furuta, Makiko Koike, Yoshihiro Yoneda, and Naoko Imamoto. "The 70-kD Heat Shock Cognate Protein (Hsc70) Facilitates the Nuclear Export of the Import Receptors." *The Journal of Cell Biology* 171, no. 1 (October 10, 2005): 19–25. https://doi.org/10.1083/jcb.200506074.

87. Kukačka, Zdeněk, Michal Rosůlek, Jan Jelínek, Lukáš Slavata, Daniel Kavan, and Petr Novák. "LinX: A Software Tool for Uncommon Cross-Linking Chemistry." *Journal of Proteome Research* 20, no. 4 (April 2, 2021): 2021–27. https://doi.org/10.1021/acs.jproteome.0c00858.

88. Lagerwaard, Ilse M., Pascal Albanese, Andris Jankevics, and Richard A. Scheltema. "Xlink Mapping and AnalySis (XMAS) - Smooth Integrative Modeling in ChimeraX." Preprint. Bioinformatics, April 22, 2022. https://doi.org/10.1101/2022.04.21.489026.

89. Lasker, Keren, Friedrich Förster, Stefan Bohn, Thomas Walzthoeni, Elizabeth Villa, Pia Unverdorben, Florian Beck, Ruedi Aebersold, Andrej Sali, and Wolfgang Baumeister. "Molecular Architecture of the 26S Proteasome Holocomplex Determined by an Integrative Approach." *Proceedings of the National Academy of Sciences* 109, no. 5 (January 31, 2012): 1380–87. https://doi.org/10.1073/pnas.1120559109.

90. Lauber, Matthew A., and James P. Reilly. “Structural Analysis of a Prokaryotic Ribosome Using a Novel Amidinating Cross-Linker and Mass Spectrometry.” *Journal of Proteome Research* 10, no. 8 (August 5, 2011): 3604–16. https://doi.org/10.1021/pr200260n.

91. Laufen, Thomas, Matthias P. Mayer, Christian Beisel, Dagmar Klostermeier, Axel Mogk, Jochen Reinstein, and Bernd Bukau. “Mechanism of Regulation of Hsp70 Chaperones by DnaJ Cochaperones.” *Proceedings of the National Academy of Sciences* 96, no. 10 (May 11, 1999): 5452–57. https://doi.org/10.1073/pnas.96.10.5452.

92. Lee, G. J. “A Small Heat Shock Protein Stably Binds Heat-Denatured Model Substrates and Can Maintain a Substrate in a Folding-Competent State.” *The EMBO Journal* 16, no. 3 (February 1, 1997): 659–71. https://doi.org/10.1093/emboj/16.3.659.

93. Leitner, Alexander, Alexandre M.J.J. Bonvin, Christoph H. Borchers, Robert J. Chalkley, Julia Chamot-Rooke, Colin W. Combe, Jürgen Cox, et al. “Toward Increased Reliability, Transparency, and Accessibility in Cross-Linking Mass Spectrometry.” *Structure* 28, no. 11 (November 2020): 1259–68. https://doi.org/10.1016/j.str.2020.09.011.

94. Lewis, Victoria A., Gillian M. Hynes, Dong Zheng, Helen Saibil, and Keith Willison. “T-Complex Polypeptide-1 Is a Subunit of a Heteromeric Particle in the Eukaryotic Cytosol.” *Nature* 358, no. 6383 (July 16, 1992): 249–52. https://doi.org/10.1038/358249a0.

95. Liberek, K, J Marszalek, D Ang, C Georgopoulos, and M Zylicz. “Escherichia Coli DnaJ and GrpE Heat Shock Proteins Jointly Stimulate ATPase Activity of DnaK.” *Proceedings of the National Academy of Sciences* 88, no. 7 (April 1991): 2874–78. https://doi.org/10.1073/pnas.88.7.2874.

96. Lindquist, S. “The Heat-Shock Proteins”, 1988. https://doi.org/10.1146/annurev.ge.22.120188.003215.

97. Lindquist, S. “The Heat-Shock Response”, 1986. https://doi.org/10.1146/annurev.bi.55.070186.005443.

98. Linke, Katrin, Tobias Wolfram, Johanna Bussemer, and Ursula Jakob. “The Roles of the Two Zinc Binding Sites in DnaJ.” *Journal of Biological Chemistry* 278, no. 45 (November 2003): 44457–66. https://doi.org/10.1074/jbc.M307491200.

99. Liu, Xiaoran Roger, Mengru Mira Zhang, and Michael L. Gross. “Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications.” *Chemical Reviews* 120, no. 10 (May 27, 2020): 4355–4454. https://doi.org/10.1021/acs.chemrev.9b00815.

100. Llorca, O. “Eukaryotic Chaperonin CCT Stabilizes Actin and Tubulin Folding Intermediates in Open Quasi-Native Conformations.” *The EMBO Journal* 19, no. 22 (November 15, 2000): 5971–79. https://doi.org/10.1093/emboj/19.22.5971.

101. Lu, Zhen, and Douglas M. Cyr. “Protein Folding Activity of Hsp70 Is Modified Differentially by the Hsp40 Co-Chaperones Sis1 and Ydj1.” *Journal of Biological Chemistry* 273, no. 43 (October 1998): 27824–30. https://doi.org/10.1074/jbc.273.43.27824.

102. Marley, Jonathan, Min Lu, and Clay Bracken. “A Method for Efficient Isotopic Labeling of Recombinant Proteins,” 2001. https://doi.org/10.1023/A:1011254402785.

103. Masson, Glenn R., John E. Burke, Natalie G. Ahn, Ganesh S. Anand, Christoph Borchers, Sébastien Brier, George M. Bou-Assaf, et al. “Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments.” *Nature Methods* 16, no. 7 (July 2019): 595–602. https://doi.org/10.1038/s41592-019-0459-y.

104. Mayer, M. P., and B. Bukau. “Hsp70 Chaperones: Cellular Functions and Molecular Mechanism.” *Cellular and Molecular Life Sciences* 62, no. 6 (March 2005): 670–84. https://doi.org/10.1007/s00018-004-4464-6.

105. Mayer, Matthias P. “The Hsp70-Chaperone Machines in Bacteria.” *Frontiers in Molecular Biosciences* 8 (June 7, 2021): 694012. https://doi.org/10.3389/fmolb.2021.694012.

106. Mayhew, Mark, Ana C. R. Da Silva, Jörg Martin, Hediye Erdjument-Bromage, Paul Tempst, and F. Ulrich Hartl. “Protein Folding in the Central Cavity of the GroEL–GroES Chaperonin Complex.” *Nature* 379, no. 6564 (February 1996): 420–26. https://doi.org/10.1038/379420a0.

107. Meier, Iris, Xiao Zhou, Jelena Brkljacić, Annkatrin Rose, Qiao Zhao, and Xianfeng Morgan Xu. “Targeting Proteins to the Plant Nuclear Envelope.” *Biochemical Society Transactions* 38, no. 3 (June 1, 2010): 733–40. https://doi.org/10.1042/BST0380733.

108. Mendoza, Vanessa Leah, and Richard W. Vachet. “Probing Protein Structure by Amino Acid-specific Covalent Labeling and Mass Spectrometry.” *Mass Spectrometry Reviews* 28, no. 5 (September 2009): 785–815. https://doi.org/10.1002/mas.20203.

109. Michels, Annemieke A., Bart Kanon, Antonius W.T. Konings, Kenzo Ohtsuka, Olivier Bensaude, and Harm H. Kampinga. “Hsp70 and Hsp40 Chaperone Activities in the Cytoplasm and the Nucleus of Mammalian Cells.” *Journal of Biological Chemistry* 272, no. 52 (December 1997): 33283–89. https://doi.org/10.1074/jbc.272.52.33283.

110. Mitra, Rishav, Kevin Wu, Changhan Lee, and James C.A. Bardwell. “ATP-Independent Chaperones.” *Annual Review of Biophysics* 51, no. 1 (May 9, 2022): 409–29. https://doi.org/10.1146/annurev-biophys-090121-082906.

111. Mogk, Axel, Bernd Bukau, and Harm H. Kampinga. “Cellular Handling of Protein Aggregates by Disaggregation Machines.” *Molecular Cell* 69, no. 2 (January 2018): 214–26. https://doi.org/10.1016/j.molcel.2018.01.004.

112. Mogk, Axel, Elke Deuerling, Sonja Vorderwülbecke, Elizabeth Vierling, and Bernd Bukau. “Small Heat Shock Proteins, ClpB and the DnaK System Form a Functional Triade in Reversing Protein Aggregation.” *Molecular Microbiology* 50, no. 2 (October 2003): 585–95. https://doi.org/10.1046/j.1365-2958.2003.03710.x.

113. Morán Luengo, Tania, Roman Kityk, Matthias P. Mayer, and Stefan G.D. Rüdiger. “Hsp90 Breaks the Deadlock of the Hsp70 Chaperone System.” *Molecular Cell* 70, no. 3 (May 2018): 545-552.e9. https://doi.org/10.1016/j.molcel.2018.03.028.

114. Morgner, Nina, Carla Schmidt, Victoria Beilsten-Edmands, Ima-obong Ebong, Nisha A. Patel, Eugenia M. Clerico, Elaine Kirschke, et al. “Hsp70 Forms Antiparallel Dimers Stabilized by Post-Translational Modifications to Position Clients for Transfer to Hsp90.” *Cell Reports* 11, no. 5 (May 2015): 759–69. https://doi.org/10.1016/j.celrep.2015.03.063.

115. Morshauser, Robert C., Hong Wang, Gregory C. Flynn, and Erik R. P. Zuiderweg. “The Peptide-Binding Domain of the Chaperone Protein Hsc70 Has an Unusual Secondary Structure Topology.” *Biochemistry* 34, no. 19 (May 16, 1995): 6261–66. https://doi.org/10.1021/bi00019a001.

116. Munro, Sean, and Hugh R B Pelham. “An Hsp70-like Protein in the ER: Identity with the 78 kd Glucose-Regulated Protein and Immunoglobulin Heavy Chain Binding Protein,” 1986. https://doi.org/10.1016/0092-8674(86)90746-4.

117. Newmyer, Sherri L., and Sandra L. Schmid. “Dominant-Interfering Hsc70 Mutants Disrupt Multiple Stages of the Clathrin-Coated Vesicle Cycle in Vivo.” *The Journal of Cell Biology* 152, no. 3 (February 5, 2001): 607–20. https://doi.org/10.1083/jcb.152.3.607.

118. Nover, Lutz, and Klaus-Dieter Scharf. “Synthesis, Modification and Structural Binding of Heat-Shock Proteins in Tomato Cell Cultures.” *European Journal of Biochemistry* 139, no. 2 (March 1984): 303–13. https://doi.org/10.1111/j.1432-1033.1984.tb08008.x.

119. Orbán-Németh, Zsuzsanna, Rebecca Beveridge, David M Hollenstein, Evelyn Rampler, Thomas Stranzl, Otto Hudecz, Johannes Doblmann, Peter Schlögelhofer, and Karl Mechtler. “Structural Prediction of Protein Models Using Distance Restraints Derived from Cross-Linking Mass Spectrometry Data.” *Nature Protocols* 13, no. 3 (March 2018): 478–94. https://doi.org/10.1038/nprot.2017.146.

120. Packschies, Lars, Holger Theyssen, Alexander Buchberger, Bernd Bukau, Roger S. Goody, and Jochen Reinstein. “GrpE Accelerates Nucleotide Exchange of the Molecular Chaperone DnaK with an Associative Displacement Mechanism.” *Biochemistry* 36, no. 12 (March 1, 1997): 3417–22. https://doi.org/10.1021/bi9628351.

121. Panaretou, B. “ATP Binding and Hydrolysis Are Essential to the Function of the Hsp90 Molecular Chaperone Invivo.” *The EMBO Journal* 17, no. 16 (August 17, 1998): 4829–36. https://doi.org/10.1093/emboj/17.16.4829.

122. Pappenberger, Günter, Julie A Wilsher, S Mark Roe, Damian J Counsell, Keith R Willison, and Laurence H Pearl. “Crystal Structure of the CCTγ Apical Domain: Implications for Substrate Binding to the Eukaryotic Cytosolic Chaperonin.” *Journal of Molecular Biology* 318, no. 5 (May 2002): 1367–79. https://doi.org/10.1016/S0022-2836(02)00190-0.

123. Parsell, D.A., A.S. Kowal, and S. Lindquist. “Saccharomyces Cerevisiae Hsp104 Protein. Purification and Characterization of ATP-Induced Structural Changes.” *Journal of Biological Chemistry* 269, no. 6 (February 1994): 4480–87. https://doi.org/10.1016/S0021-9258(17)41804-7.

124. Peroutka Iii, Raymond J., Steven J. Orcutt, James E. Strickler, and Tauseef R. Butt. “SUMO Fusion Technology for Enhanced Protein Expression and Purification in Prokaryotes and Eukaryotes.” In *Heterologous Gene Expression in E.Coli*, edited by Thomas C. Evans, and Ming-Qun Xu, 705:15–30. Methods in Molecular Biology. Totowa, NJ: Humana Press, 2011. https://doi.org/10.1007/978-1-61737-967-3_2.

125. Peterson, James J, Malin M Young, and Larry J Takemoto. “Probing α-Crystallin Structure Using Chemical Cross-Linkers and Mass Spectrometry.” *Molecular Vision*, n.d.

126. Picard, D. “Heat-Shock Protein 90, a Chaperone for Folding and Regulation.” *Cellular and Molecular Life Sciences* 59, no. 10 (October 2002): 1640–48. https://doi.org/10.1007/PL00012491.

127. Preissler, Steffen, Joseph E Chambers, Ana Crespillo-Casado, Edward Avezov, Elena Miranda, Juan Perez, Linda M Hendershot, Heather P Harding, and David Ron. “Physiological Modulation of BiP Activity by Trans-Protomer Engagement of the Interdomain Linker.” *eLife* 4 (October 16, 2015): e08961. https://doi.org/10.7554/eLife.08961.

128. Preissler, Steffen, and Elke Deuerling. “Ribosome-Associated Chaperones as Key Players in Proteostasis.” *Trends in Biochemical Sciences* 37, no. 7 (July 2012): 274–83. https://doi.org/10.1016/j.tibs.2012.03.002.

129. Rabu, Catherine, Peter Wipf, Jeffrey L. Brodsky, and Stephen High. “A Precursor-Specific Role for Hsp40/Hsc70 during Tail-Anchored Protein Integration at the Endoplasmic Reticulum.” *Journal of Biological Chemistry* 283, no. 41 (October 2008): 27504–13. https://doi.org/10.1074/jbc.M804591200.

130. Radons, Jürgen. “The Human HSP70 Family of Chaperones: Where Do We Stand?” *Cell Stress and Chaperones* 21, no. 3 (May 2016): 379–404. https://doi.org/10.1007/s12192-016-0676-6.

131. Rahimidashaghoul, Kheironnesae, Iveta Klimánková, Martin Hubálek, Václav Matoušek, Josef Filgas, Petr Slavíček, Tomáš Slanina, and Petr Beier. “Visible-Light-Driven Fluoroalkylation of Tryptophan Residues in Peptides.” *ChemPhotoChem* 5, no. 1 (January 2021): 43–50. https://doi.org/10.1002/cptc.202000214.

132. Rohde, Mikkel, Mads Daugaard, Mette Hartvig Jensen, Kristian Helin, Jesper Nylandsted, and Marja Jäättelä. “Members of the Heat-Shock Protein 70 Family Promote Cancer Cell Growth by Distinct Mechanisms.” *Genes & Development* 19, no. 5 (March 1, 2005): 570–82. https://doi.org/10.1101/gad.305405.

133. Rosenzweig, Rina, Nadinath B. Nillegoda, Matthias P. Mayer, and Bernd Bukau. “The Hsp70 Chaperone Network.” *Nature Reviews Molecular Cell Biology* 20, no. 11 (November 2019): 665–80. https://doi.org/10.1038/s41580-019-0133-3.

134. Rudiger, S. “Its Substrate Specificity Characterizes the DnaJ Co-Chaperone as a Scanning Factor for the DnaK Chaperone.” *The EMBO Journal* 20, no. 5 (March 1, 2001): 1042–50. https://doi.org/10.1093/emboj/20.5.1042.

135. Rye, Hays S., Steven G. Burston, Wayne A. Fenton, Joseph M. Beechem, Zhaohui Xu, Paul B. Sigler, and Arthur L. Horwich. "Distinct Actions of Cis and Trans ATP within the Double Ring of the Chaperonin GroEL." *Nature* 388, no. 6644 (August 1997): 792–98. https://doi.org/10.1038/42047.

136. Šali, Andrej, and Tom L. Blundell. "Comparative Protein Modelling by Satisfaction of Spatial Restraints." *Journal of Molecular Biology* 234, no. 3 (December 1993): 779–815. https://doi.org/10.1006/jmbi.1993.1626.

137. Schirmer, Eric C., John R. Glover, Mike A. Singer, and Susan Lindquist. "HSP100/Clp Proteins: A Common Mechanism Explains Diverse Functions." *Trends in Biochemical Sciences* 21, no. 8 (August 1996): 289–96. https://doi.org/10.1016/S0968-0004(96)10038-4.

138. Schmid, S L, W A Braell, and J E Rothman. "ATP Catalyzes the Sequestration of Clathrin during Enzymatic Uncoating." *Journal of Biological Chemistry* 260, no. 18 (August 1985): 10057–62. https://doi.org/10.1016/S0021-9258(17)39211-6.

139. Sharon, Michal. "How Far Can We Go with Structural Mass Spectrometry of Protein Complexes?" *Journal of the American Society for Mass Spectrometry* 21, no. 4 (April 1, 2010): 487–500. https://doi.org/10.1016/j.jasms.2009.12.017.

140. Sheffield, W P, G C Shore, and S K Randall. "Mitochondrial Precursor Protein. Effects of 70-Kilodalton Heat Shock Protein on Polypeptide Folding, Aggregation, and Import Competence." *Journal of Biological Chemistry* 265, no. 19 (July 1990): 11069–76. https://doi.org/10.1016/S0021-9258(19)38558-8.

141. Shi, Yanggu, and John O. Thomas. "The Transport of Proteins into the Nucleus Requires the 70-Kilodalton Heat Shock Protein or Its Cytosolic Cognate." *Molecular and Cellular Biology* 12, no. 5 (May 1, 1992): 2186–92. https://doi.org/10.1128/mcb.12.5.2186-2192.1992.

142. Sinz, Andrea. "Chemical Cross-linking and Mass Spectrometry for Mapping Three-dimensional Structures of Proteins and Protein Complexes." *Journal of Mass Spectrometry* 38, no. 12 (December 2003): 1225–37. https://doi.org/10.1002/jms.559.

143. Sinz, Andrea. "Chemical Cross-linking and Mass Spectrometry to Map Three-dimensional Protein Structures and Protein–Protein Interactions." *Mass Spectrometry Reviews* 25, no. 4 (July 2006): 663–82. https://doi.org/10.1002/mas.20082.

144. Smock, Robert G., Mandy E. Blackburn, and Lila M. Gierasch. "Conserved, Disordered C Terminus of DnaK Enhances Cellular Survival upon Stress and DnaK in Vitro Chaperone Activity." *Journal of Biological Chemistry* 286, no. 36 (September 2011): 31821–29. https://doi.org/10.1074/jbc.M111.265835.

145. Smythe, C.V. "THE REACTION OF IODOACETATE AND OF IODOACETAMIDE WITH VARIOUS SULFHYDRYL GROUPS, WITH UREASE, AND WITH YEAST PREPARATIONS." *Journal of Biological Chemistry* 114, no. 3 (July 1936): 601–12. https://doi.org/10.1016/S0021-9258(18)74789-3.

146. Snijder, Joost, Rebecca J. Rose, David Veesler, John E. Johnson, and Albert J. R. Heck. "Studying 18 MDa Virus Assemblies with Native Mass Spectrometry." *Angewandte Chemie International Edition* 52, no. 14 (April 2, 2013): 4020–23. https://doi.org/10.1002/anie.201210197.

147. Soltermann, Fabian, Eric D. B. Foley, Veronica Pagnoni, Martin Galpin, Justin L. P. Benesch, Philipp Kukura, and Weston B. Struwe. "Quantifying Protein–Protein Interactions by Molecular Counting with Mass Photometry." *Angewandte Chemie International Edition* 59, no. 27 (June 26, 2020): 10774–79. https://doi.org/10.1002/anie.202001578.

148. Sonn-Segev, Adar, Katarina Belacic, Tatyana Bodrug, Gavin Young, Ryan T. VanderLinden, Brenda A. Schulman, Johannes Schimpf, et al. "Quantifying the Heterogeneity of Macromolecular Machines by Mass Photometry." *Nature Communications* 11, no. 1 (April 14, 2020): 1772. https://doi.org/10.1038/s41467-020-15642-w.

149. Spence, J, A Cegielska, and C Georgopoulos. "Role of Escherichia Coli Heat Shock Proteins DnaK and HtpG (C62.5) in Response to Nutritional Deprivation." *Journal of Bacteriology* 172, no. 12 (December 1990): 7157–66. https://doi.org/10.1128/jb.172.12.7157-7166.1990.

150. Srikakulam, Rajani, and Donald A. Winkelmann. “Chaperone-Mediated Folding and Assembly of Myosin in Striated Muscle.” *Journal of Cell Science* 117, no. 4 (February 1, 2004): 641–52. https://doi.org/10.1242/jcs.00899.

151. Stege, G J J, K Renkawek, P S G Overkamp, P Verschuure, A F van Rijk, A Reijnen-Aalbers, and W C Boelens. “The Molecular Chaperone aB-Crystallin Enhances Amyloid Neurotoxicity.” *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* 262, no. 1 (1999). https://doi.org/10.1006/bbrc.1999.1167.

152. Szabo, Alexander, and F Ulrich Hartl. “The ATP Hydrolysis-Dependent Reaction Cycle of the Escherichia Coli Hsp7O System-DnaK, DnaJ, and GrpE.” *Proc. Natl. Acad. Sci. USA*, 1994.

153. Taverner, Thomas, Nathan E. Hall, Richard A.J. O’Hair, and Richard J. Simpson. “Characterization of an Antagonist Interleukin-6 Dimer by Stable Isotope Labeling, Cross-Linking, and Mass Spectrometry.” *Journal of Biological Chemistry* 277, no. 48 (November 2002): 46487–92. https://doi.org/10.1074/jbc.M207370200.

154. Teter, Sarah A, Walid A Houry, Debbie Ang, Thomas Tradler, David Rockabrand, Gunter Fischer, Paul Blum, Costa Georgopoulos, and F Ulrich Hartl. “Polypeptide Flux through Bacterial Hsp70: DnaK Cooperates with Trigger Factor in Chaperoning Nascent Chains,” 1999. https://doi.org/10.1016/S0092-8674(00)80787-4.

155. Thompson, Andrea D., Steffen M. Bernard, Georgios Skiniotis, and Jason E. Gestwicki. “Visualization and Functional Analysis of the Oligomeric States of Escherichia Coli Heat Shock Protein 70 (Hsp70/DnaK).” *Cell Stress and Chaperones* 17, no. 3 (May 2012): 313–27. https://doi.org/10.1007/s12192-011-0307-1.

156. Thulasiraman, V. “In Vivo Newly Translated Polypeptides Are Sequestered in a Protected Folding Environment.” *The EMBO Journal* 18, no. 1 (January 4, 1999): 85–95. https://doi.org/10.1093/emboj/18.1.85.

157. Todd, Matthew J., Paul V. Viitanen, and George H. Lorimer. “Dynamics of the Chaperonin ATPase Cycle: Implications for Facilitated Protein Folding.” *Science* 265, no. 5172 (July 29, 1994): 659–66. https://doi.org/10.1126/science.7913555.

158. Van Dijk, Aalt D. J., Rolf Boelens, and Alexandre M. J. J. Bonvin. “Data-driven Docking for the Study of Biomolecular Complexes.” *The FEBS Journal* 272, no. 2 (January 2005): 293–312. https://doi.org/10.1111/j.1742-4658.2004.04473.x.

159. Van Zundert, G.C.P., J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, A.S.J. Melquiond, M. Van Dijk, S.J. De Vries, and A.M.J.J. Bonvin. “The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes.” *Journal of Molecular Biology* 428, no. 4 (February 2016): 720–25. https://doi.org/10.1016/j.jmb.2015.09.014.

160. Vandermarliere, Elien, Elisabeth Stes, Kris Gevaert, and Lennart Martens. “Resolution of Protein Structure by Mass Spectrometry.” *Mass Spectrometry Reviews* 35, no. 6 (October 2016): 653–65. https://doi.org/10.1002/mas.21450.

161. Vogel, Markus, Matthias P. Mayer, and Bernd Bukau. “Allosteric Regulation of Hsp70 Chaperones Involves a Conserved Interdomain Linker.” *Journal of Biological Chemistry* 281, no. 50 (December 2006): 38705–11. https://doi.org/10.1074/jbc.M609020200.

162. Wandinger, Sebastian Karl, Klaus Richter, and Johannes Buchner. “The Hsp90 Chaperone Machinery.” *Journal of Biological Chemistry* 283, no. 27 (July 2008): 18473–77. https://doi.org/10.1074/jbc.R800007200.

163. Wegele, Harald, Sebastian K. Wandinger, Andreas B. Schmid, Jochen Reinstein, and Johannes Buchner. “Substrate Transfer from the Chaperone Hsp70 to Hsp90.” *Journal of Molecular Biology* 356, no. 3 (February 2006): 802–11. https://doi.org/10.1016/j.jmb.2005.12.008.

164. Whitesell, Luke, and Susan L. Lindquist. “HSP90 and the Chaperoning of Cancer.” *Nature Reviews Cancer* 5, no. 10 (October 2005): 761–72. https://doi.org/10.1038/nrc1716.

165. Wiech, Hans, Johannes Buchner, Richard Zimmermann, and Ursula Jakob. “Hsp90 Chaperones Protein Folding in Vitro.” *Nature* 358, no. 6382 (July 1992): 169–70. https://doi.org/10.1038/358169a0.

166. Wu, Ching-Chung, Vankadari Naveen, Chin-Hsiang Chien, Yi-Wei Chang, and Chwan-Deng Hsiao. "Crystal Structure of DnaK Protein Complexed with Nucleotide Exchange Factor GrpE in DnaK Chaperone System." *Journal of Biological Chemistry* 287, no. 25 (June 2012): 21461–70. https://doi.org/10.1074/jbc.M112.344358.

167. Wu, Yunkun, Jingzhi Li, Zhongmin Jin, Zhengqing Fu, and Bingdong Sha. "The Crystal Structure of the C-Terminal Fragment of Yeast Hsp40 Ydj1 Reveals Novel Dimerization Motif for Hsp40." *Journal of Molecular Biology* 346, no. 4 (March 2005): 1005–11. https://doi.org/10.1016/j.jmb.2004.12.040.

168. Wysocki, Vicki H., Katheryn A. Resing, Qingfen Zhang, and Guilong Cheng. "Mass Spectrometry of Peptides and Proteins." *Methods* 35, no. 3 (March 2005): 211–22. https://doi.org/10.1016/j.ymeth.2004.08.013.

169. Xu, Zhaohui, Arthur L. Horwich, and Paul B. Sigler. "The Crystal Structure of the Asymmetric GroEL–GroES–(ADP)7 Chaperonin Complex." *Nature* 388, no. 6644 (August 1997): 741–50. https://doi.org/10.1038/41944.

170. Young, Jason C., Nicholas J. Hoogenraad, and F.Ulrich Hartl. "Molecular Chaperones Hsp90 and Hsp70 Deliver Preproteins to the Mitochondrial Import Receptor Tom70." *Cell* 112, no. 1 (January 2003): 41–50. https://doi.org/10.1016/S0092-8674(02)01250-3.

171. Zhang, Bojie, Ming Cheng, Don Rempel, and Michael L. Gross. "Implementing Fast Photochemical Oxidation of Proteins (FPOP) as a Footprinting Approach to Solve Diverse Problems in Structural Biology." *Methods* 144 (July 2018): 94–103. https://doi.org/10.1016/j.ymeth.2018.05.016.

172. Zhu, Xiaotian, Xun Zhao, William F. Burkholder, Alexander Gragerov, Craig M. Ogata, Max E. Gottesman, and Wayne A. Hendrickson. "Structural Analysis of Substrate Binding by the Molecular Chaperone DnaK." *Science* 272, no. 5268 (June 14, 1996): 1606–14. https://doi.org/10.1126/science.272.5268.1606.

173. Zybailov, Boris L. "Large Scale Chemical Cross-Linking Mass Spectrometry Perspectives." *Journal of Proteomics & Bioinformatics* 01, no. S2 (2013). https://doi.org/10.4172/jpb.S2-001.

Appendix

A

```
>sp|P11142|HSP7C_HUMAN Heat shock cognate 71 kDa protein OS=Homo sapiens OX=9606 GN=HSPA8 PE=1 SV=1
MSKGPAVGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVA
MNPTNTVFDAKRLIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGETKSFYPEEVS
SMVLTKMKEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAA
IAYGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDTHLGGEDFDNRMVNH
FIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGIDFYTSITRA
RFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRIPKIQKLLQDFFNGKELN
KSINPDEAVAYGAAVQAAILSGDKSENVQDLLLLDVTPLSLGIETAGGVMTVLIKRNTTI
PTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLGKFELTGIPPAPRGVPQIEVTFDI
DANGILNVSAVDKSTGKENKITITNDKGRLSKEDIERMVQEAEKYKAEDEKQRDKVSSKN
SLESYAFNMKATVEDEKLQGKINDEDKQKILDKCNEIINWLDKNQTAEKEEFEHQQKELE
KVCNPIITKLYQSAGGMPGGMPGGFPGGGAPPSGGASSGPTIEEVD
```

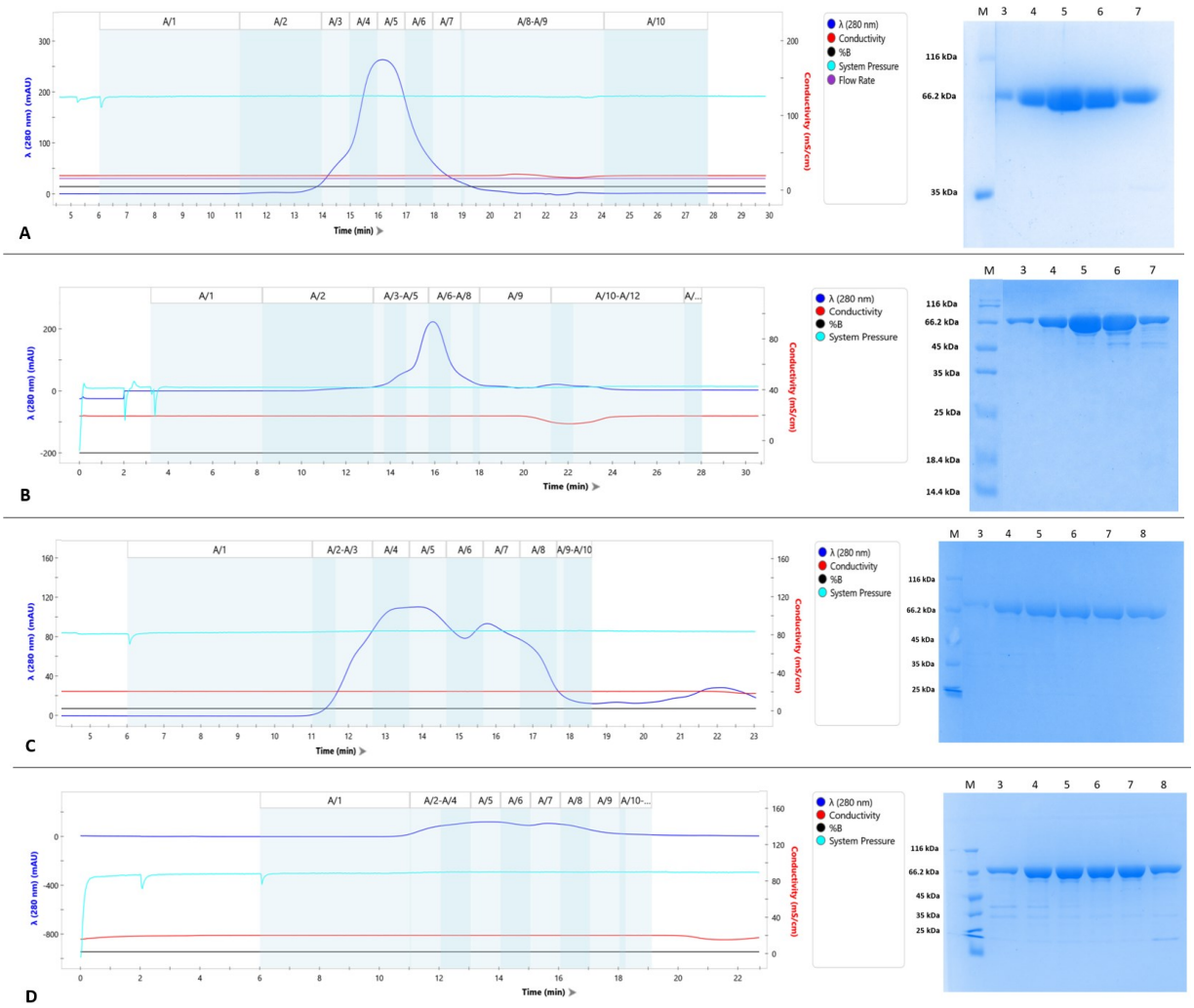
B

```
>HSP7C_HUMAN Heat shock cognate 71 kDa protein SBD-mutant V438F OS=Homo sapiens
MSKGPAVGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVA
MNPTNTVFDAKRLIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGETKSFYPEEVS
SMVLTKMKEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAA
IAYGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDTHLGGEDFDNRMVNH
FIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGIDFYTSITRA
RFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRIPKIQKLLQDFFNGKELN
KSINPDEAVAYGAAVQAAILSGDKSENVQDLLLLDVTPLSLGIETAGGVMTVLIKRNTTI
PTKQTQTFTTYSDNQPGFLIQVYEGERAMTKDNNLLGKFELTGIPPAPRGVPQIEVTFDI
DANGILNVSAVDKSTGKENKITITNDKGRLSKEDIERMVQEAEKYKAEDEKQRDKVSSKN
SLESYAFNMKATVEDEKLQGKINDEDKQKILDKCNEIINWLDKNQTAEKEEFEHQQKELE
KVCNPIITKLYQSAGGMPGGMPGGFPGGGAPPSGGASSGPTIEEVD
```

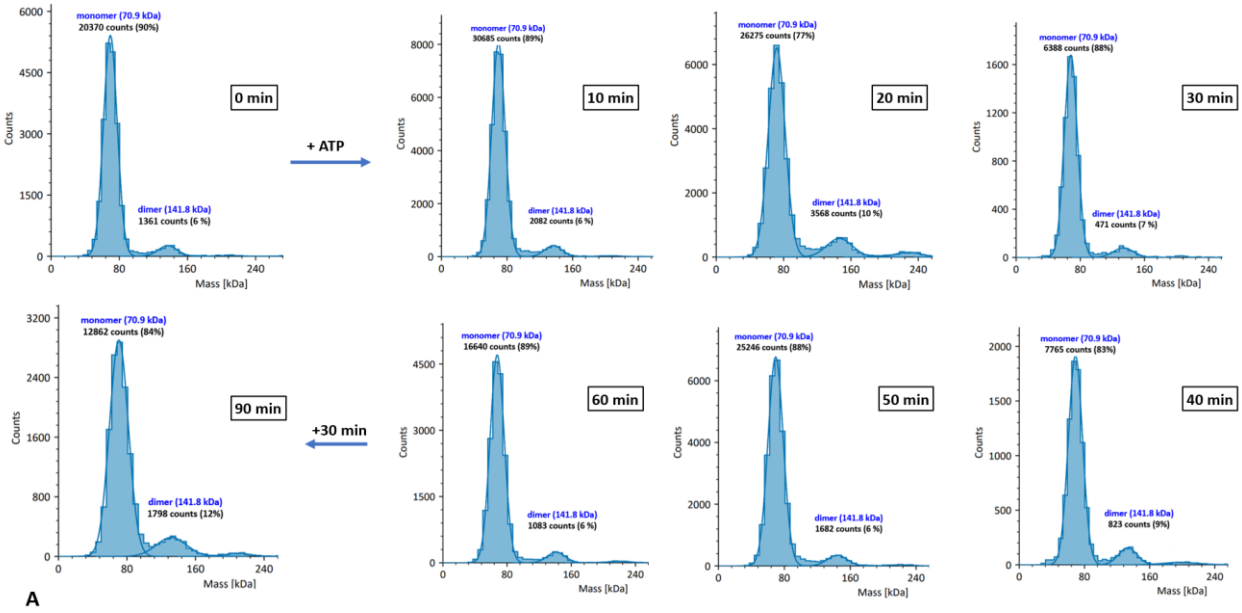
C

```
>HSP7C_HUMAN Heat shock cognate 71 kDa protein linker-mutant L392E,L394E OS=Homo sapiens
MSKGPAVGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVA
MNPTNTVFDAKRLIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGETKSFYPEEVS
SMVLTKMKEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAA
IAYGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDTHLGGEDFDNRMVNH
FIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGIDFYTSITRA
RFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRIPKIQKLLQDFFNGKELN
KSINPDEAVAYGAAVQAAILSGDKSENVQDLELEDVTPLSLGIETAGGVMTVLIKRNTTI
PTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLGKFELTGIPPAPRGVPQIEVTFDI
DANGILNVSAVDKSTGKENKITITNDKGRLSKEDIERMVQEAEKYKAEDEKQRDKVSSKN
SLESYAFNMKATVEDEKLQGKINDEDKQKILDKCNEIINWLDKNQTAEKEEFEHQQKELE
KVCNPIITKLYQSAGGMPGGMPGGFPGGGAPPSGGASSGPTIEEVD
```

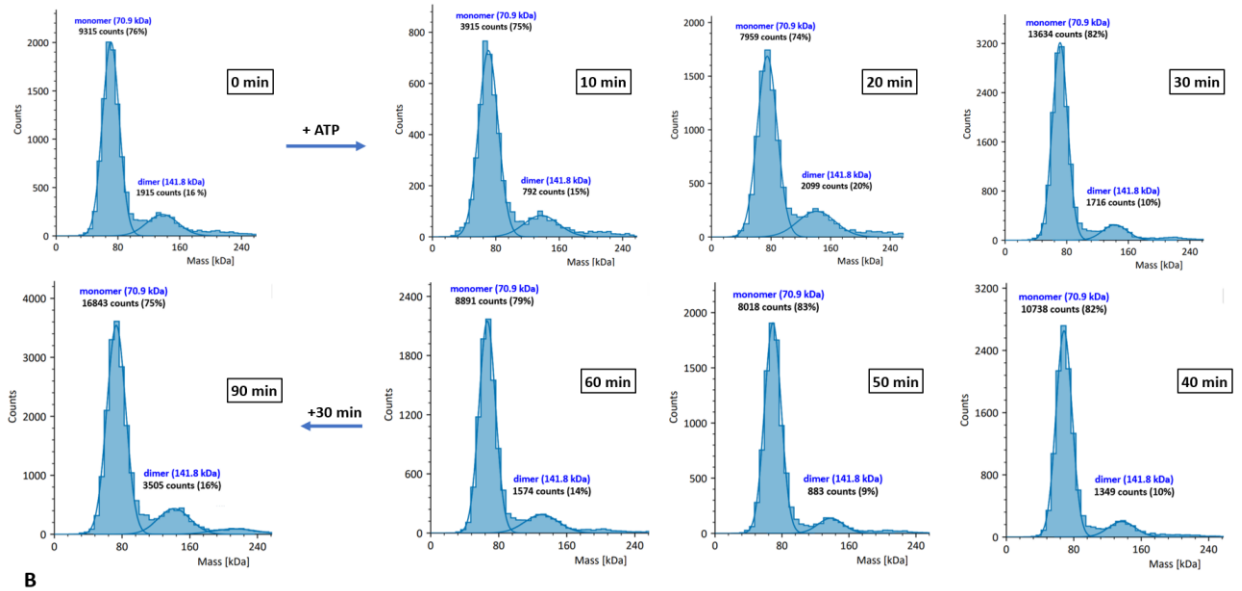
Appendix 1. The sequences of the studied proteins. **A.** HSC70 WT. **B.** HSC70 SBD-mutant (V438F); **C.** HSC70 linker-mutant (L392E, L394E).



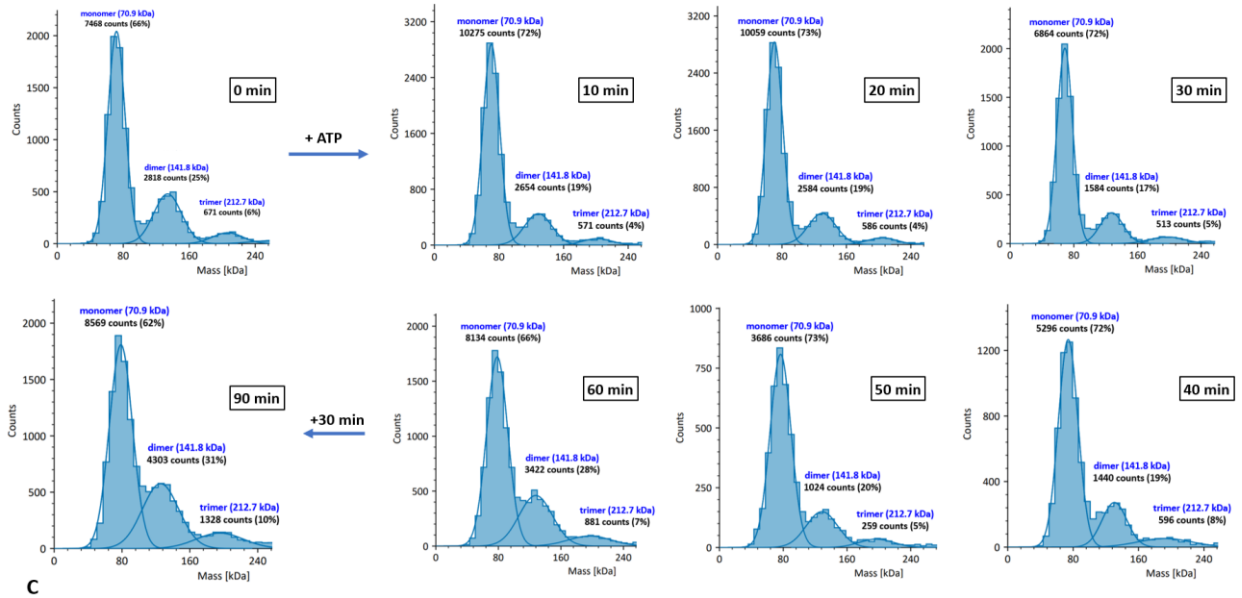
Appendix 2. The SEC chromatograms and SDS-PAGE gels of the collected fractions. **A.** ^{14}N-SBD-mutant; **B.** ^{15}N-SBD-mutant; **C.** ^{14}N-linker-mutant; **D.** ^{15}N-linker-mutant.


0 min
monomer (70.9 kDa)
20370 counts (90%)
dimer (141.8 kDa)
1361 counts (6 %)
+ ATP
10 min
monomer (70.9 kDa)
30685 counts (89%)
dimer (141.8 kDa)
2082 counts (6 %)
20 min
monomer (70.9 kDa)
26275 counts (77%)
dimer (141.8 kDa)
3568 counts (10 %)
30 min
monomer (70.9 kDa)
6388 counts (88%)
dimer (141.8 kDa)
471 counts (7 %)
40 min
monomer (70.9 kDa)
7765 counts (83%)
dimer (141.8 kDa)
823 counts (9%)
50 min
monomer (70.9 kDa)
25246 counts (88%)
dimer (141.8 kDa)
1682 counts (6 %)
60 min
monomer (70.9 kDa)
16640 counts (89%)
dimer (141.8 kDa)
1083 counts (6 %)
+30 min
90 min
monomer (70.9 kDa)
12862 counts (84%)
dimer (141.8 kDa)
1798 counts (12%)
Counts
Mass [kDa]

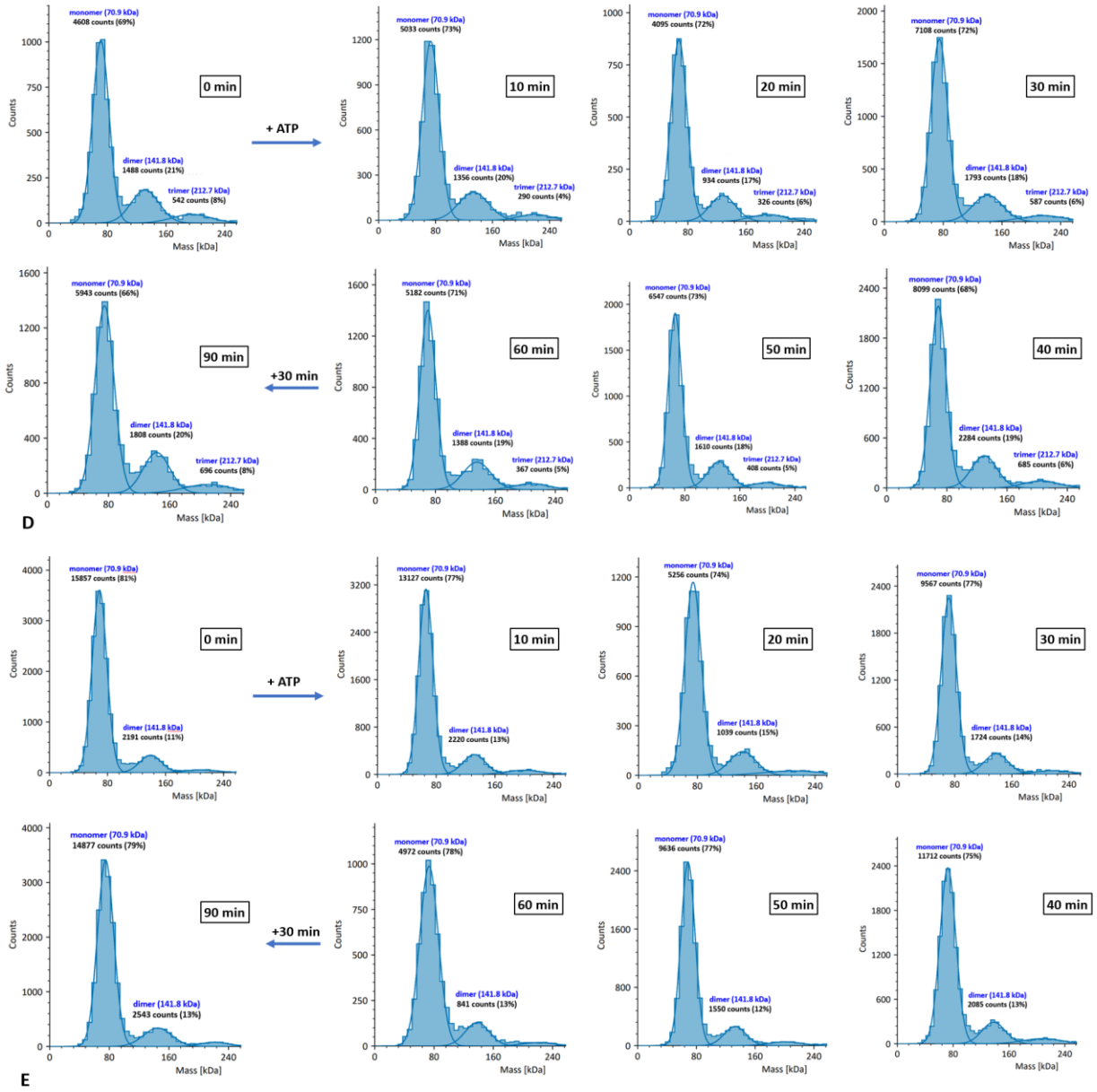
A


0 min
monomer (70.9 kDa)
9315 counts (76%)
dimer (141.8 kDa)
1915 counts (16 %)
+ ATP
10 min
monomer (70.9 kDa)
3915 counts (75%)
dimer (141.8 kDa)
792 counts (15%)
20 min
monomer (70.9 kDa)
7959 counts (74%)
dimer (141.8 kDa)
2099 counts (20%)
30 min
monomer (70.9 kDa)
13634 counts (82%)
dimer (141.8 kDa)
1716 counts (10%)
40 min
monomer (70.9 kDa)
10738 counts (82%)
dimer (141.8 kDa)
1349 counts (10%)
50 min
monomer (70.9 kDa)
8018 counts (83%)
dimer (141.8 kDa)
883 counts (9%)
60 min
monomer (70.9 kDa)
8891 counts (79%)
dimer (141.8 kDa)
1574 counts (14%)
+30 min
90 min
monomer (70.9 kDa)
16843 counts (75%)
dimer (141.8 kDa)
3505 counts (16%)
Counts
Mass [kDa]

B


0 min
monomer (70.9 kDa)
7468 counts (66%)
dimer (141.8 kDa)
2818 counts (25%)
trimer (212.7 kDa)
671 counts (6%)
+ ATP
10 min
monomer (70.9 kDa)
10275 counts (72%)
dimer (141.8 kDa)
2654 counts (19%)
trimer (212.7 kDa)
571 counts (4%)
20 min
monomer (70.9 kDa)
10059 counts (73%)
dimer (141.8 kDa)
2584 counts (19%)
trimer (212.7 kDa)
586 counts (4%)
30 min
monomer (70.9 kDa)
6864 counts (72%)
dimer (141.8 kDa)
1584 counts (17%)
trimer (212.7 kDa)
513 counts (5%)
40 min
monomer (70.9 kDa)
5296 counts (72%)
dimer (141.8 kDa)
1440 counts (19%)
trimer (212.7 kDa)
596 counts (8%)
50 min
monomer (70.9 kDa)
3686 counts (73%)
dimer (141.8 kDa)
1024 counts (20%)
trimer (212.7 kDa)
259 counts (5%)
60 min
monomer (70.9 kDa)
8134 counts (66%)
dimer (141.8 kDa)
3422 counts (28%)
trimer (212.7 kDa)
881 counts (7%)
+30 min
90 min
monomer (70.9 kDa)
8569 counts (62%)
dimer (141.8 kDa)
4303 counts (31%)
trimer (212.7 kDa)
1328 counts (10%)
Counts
Mass [kDa]

C



Appendix 3. Mass photometry measurements in time after the addition of ATP. **A.** SBD-mutant only; **B.** linker-mutant only; **C.** equimolar WT + SBD-mutant; **D.** equimolar WT + linker-mutant; **E.** equimolar SBD-mutant + linker-mutant.